A STUDY ON SENTIMENTAL ANALYSIS AND TRUST IN COVID-19 PANDEMIC SITUATION

Thesis Submitted for the Award of the Degree of

DOCTOR OF PHILOSOPHY

in

(Computer Science & Engineering)

By

Aseem Kumar

Registration Number: 41801005

Supervised By

Dr. Arun Malik

Computer Sciences & Engineering

Lovely Professional University

Transforming Education Transforming India

LOVELY PROFESSIONAL UNIVERSITY, PUNJAB

2024

DECLARATION

I, hereby declared that the presented work in the thesis entitled "A Study on Sentimental Analysis and Trust in Covid-19 Pandemic Situation" in fulfilment of degree of **Doctor of Philosophy (Ph.D.)** is outcome of research work carried out by me under the supervision Dr. Arun Malik, working as **Professor** in the Computer Science & Engineering of Lovely Professional University, Punjab, India. In keeping with general practice of reporting scientific observations, due acknowledgements have been made whenever work described here has been based on findings of other investigator. This work has not been submitted in part or full to any other University or Institute for the award of any degree.

(Signature of Scholar)

Aseem Kumar
Registration No.: 41801005
Computer Science & Engineering
Lovely Professional University,
Punjab, India

CERTIFICATE

This is to certify that the work reported in the Ph. D. thesis entitled "A Study on Sentimental Analysis and Trust in Covid-19 Pandemic Situation" submitted in fulfillment of the requirement for the reward of degree of **Doctor of Philosophy (Ph.D.)** in the Computer Science & Engineering, is a research work carried out by Aseem Kumar, Registration No. 41801005, is bonafide record of his/her original work carried out under my supervision and that no part of thesis has been submitted for any other degree, diploma or equivalent course.

(Signature of Supervisor)

Name of supervisor:

Designation:

Department/school:

University:

ABSTRACT

This doctoral research presents a comprehensive computational study of sentiment and trust dynamics within digital societies during the COVID-19 pandemic, focusing on the unique challenges of analyzing Hinglish, a blend of Hindi and English widely used in social media communications. This work addresses critical gaps in understanding how trust emerges, fluctuates, and evolves in online interactions under crisis conditions, emphasizing the complexities introduced by multilingual, code-mixed discourse in diverse digital communities.

The study is grounded in an extensive review of existing literature encompassing digital trust models, sentiment analysis techniques, and natural language processing of mixed languages. Through this review, significant limitations were identified in prior approaches, particularly the absence of robust mathematical and computational frameworks capable of accurately quantifying trust relationships in linguistically diverse online environments experiencing extraordinary circumstances such as a global pandemic. This foundational analysis motivated the development of novel models designed to better capture the multifaceted nature of digital trust expressed through complex linguistic and social signals.

The research employed a rigorous and multi-phased methodology covering data collection, algorithm development, and model validation within a COVID-19 contextual framework. Data gathering focused on the critical pandemic period, capturing Indian social discourse through various phases—from the initial outbreak to prolonged lockdowns—with global data for broader contextualization. Comprehensive datasets were utilized, including geo-located COVID-19 lockdown tweets from India, a vast multilingual Twitter dataset exceeding 500 million tweets, topic-specific collections identifying hashtag-country relationships worldwide, sentiment-annotated second wave conversations, and trending coronavirus discourse data. These diverse datasets ensured a rich, representative sample of social media communication patterns, enabling a thorough exploration of trust dynamics.

Corpus construction employed advanced filtering protocols using Hinglish keywords to represent an extensive range of emotional expressions and trust-related indicators.

Terms ranged from simple affective states such as "sveekaar" (acceptance) and "vishvaas" (trust) to more nuanced sentiments like "aashavaadee" (optimistic) and "niraashaavaad" (pessimistic). This granular keyword inclusion allowed capturing subtle linguistic cues within the code-mixed tweets, which reflect nuanced social trust or distrust signals. Data acquisition utilized Twitter's developer APIs with authentication secured via designated developer accounts, applying systematic hydration procedures to retrieve full tweet metadata comprehensively. The methodology ensured strict adherence to platform usage policies and augmented the corpus's relevance and quality for subsequent computational analyses.

The approved research objectives included developing algorithms to understand the dynamics of trust within digital societies such as Twitter. Two innovative computational models were designed to meet the study's objectives. The Equation-based Digital Social Trust Model (E.D.S.T.M.) introduced a mathematically grounded, linear regression framework incorporating seven pivotal variables: demographic alignments (age, gender, ethnicity, geographical origins), shared interests (hobbies), interaction metrics (frequency of communication), and linguistic sentiment signals (count of positive sentiment utterances). This model offered a structured and interpretable approach to quantifying trust as a weighted function of these foundational social and behavioral indicators. Each coefficient captures the relative impact of corresponding variables, providing a clear mechanistic understanding of trust formation in digital interactions. Recognizing that linear models may insufficiently capture the complexity and dynamism of real-world social trust phenomena, the Advanced Digital Society Interaction and Trust Model (A.D.S.I.T.M.) was developed to overcome these limitations. This sophisticated model incorporates non-linear interdependencies, emotional variability, and cognitive perception influences, alongside adaptive mechanisms responding to external disruptive events such as the COVID-19 pandemic. Employing clustering and segregation algorithms, the model segregates digital society members into nuanced groups based on emotional profiles, trust levels, and the intensity of pandemic-related impacts. This enables the model to reflect the heterogeneous and evolving nature of trust correlations more realistically across complex, large-scale, and culturally diverse digital social networks.

Additionally, the study employed content analysis techniques—specifically Guided Content Analysis (G.C.A.) and the Topic Content Analysis Algorithm (T.C.A.A.)—combined with Speech Act Theory (S.A.T.) to analyze semantic meaning, contextual nuances, and communicative intentions within a corpus of COVID-19-related tweets. Model construction and evaluation utilized a mixture of secondary, primary, and simulated Twitter data. Empirical evaluation of the proposed models demonstrated clear advantages of the A.D.S.I.T.M. over the E.D.S.T.M. across a suite of performance metrics. The advanced model consistently achieves a higher recall range (0.75–1.0) than the equation-based approach (0.5–1.0), reflecting its enhanced ability to accurately identify positive trust instances. A.D.S.I.T.M also has better precision range (0.93-1.0) over the E.D.S.T.M (0.5–1.0).

The A.D.S.I.T.M. demonstrates a better accuracy range (0.91–0.99) compared to the E.D.S.T.M. (0.48–0.99), indicating superior overall accuracy. It consistently outperforms the E.D.S.T.M. in model fit, with R^2 values ranging from 0.513 to 0.776 across iterations, whereas the E.D.S.T.M.'s R^2 values vary between 0.48 and 1.0. Regarding mean absolute error (MAE), the E.D.S.T.M. exhibits a wider error range (0.55–1.0) compared to the significantly lower and narrower range of the A.D.S.I.T.M. (0.085–0.220), suggesting the latter's greater precision in aligning predicted and observed values. The A.D.S.I.T.M. also shows superior forecasting accuracy, evidenced by consistently lower root mean squared error (RMSE) values (0.111–0.189) relative to the E.D.S.T.M.’s broader range (0.48–0.99). Numerically, the A.D.S.I.T.M. maintains greater stability across supplementary performance metrics, consistently delivering reliable outcomes despite input data variability. In contrast, the E.D.S.T.M. exhibits moderate numerical stability, indicating slightly higher sensitivity to fluctuations. Moreover, the A.D.S.I.T.M. offers enhanced flexibility and adaptability, enabling it to effectively handle diverse scenarios and respond dynamically to changing conditions. Both models possess a moderate level of interpretability, meaning their operations and results can be reasonably understood and explained. However, the E.D.S.T.M. is comparatively less complex, favoring simplicity and transparency, whereas the A.D.S.I.T.M. incorporates higher complexity due to its advanced features. Overall, there is strong evidence that the A.D.S.I.T.M. surpasses the E.D.S.T.M. in key performance aspects including recall, R^2, MAE, RMSE, numerical stability, flexibility, and adaptability, while both maintain a balanced degree of interpretability suited to their respective complexities.

The research makes impactful theoretical and practical contributions by synthesizing sociological trust constructs with computational linguistics and machine learning to forge comprehensive analytical frameworks for multilingual digital societies, especially under crisis conditions. The focus on Hinglish communication uniquely addresses challenges in code-mixed language analysis, offering significant insights for comparable multilingual settings worldwide. Finally, during this study, the third objective was pursued by utilizing algorithms such as NMFF, PLSI, and LDA, whereby keywords were extracted, and tweets were classified according to the principles of speech act theory. Findings reveal a dominance of persuasive communication tactics in pandemic-related discourse, with frequent use of terms such as "kindly," "please," and "retweet" that function to influence online social behaviors and trust dynamics. The study validates seminal components of Austin and Searle's speech act theory in this contemporary digital and multilingual context, while highlighting the extensive presence of COVID-19 discourse encompassing factual data, scientific assertions, conspiracy narratives, and advisory exchanges.

Overall, this comprehensive investigation advances the understanding of how digital trust operates and transforms during global health crises, providing rigorously tested computational tools and theoretical lenses applicable to future pandemic responses and multilingual social media studies. The adaptable models and methodologies developed herein pave the way for enhanced real-time monitoring and nuanced interpretation of evolving trust and sentiment landscapes in increasingly complex, multilingual, and interconnected digital ecosystems.

PREFACE/ACKNOWLEDGEMENT

I look back on the long and rewarding journey of my Ph.D., I am filled with gratitude for the incredible people who have stood by me, their support shaping every step of this achievement.

First and foremost, my deepest thanks go to my mother, Smt. Sumitra Devi. Your endless love, unwavering faith, and constant encouragement laid the foundation for everything I have accomplished. From my earliest days, your determination and hard work inspired me to start and persist in my schooling. This success belongs to you as much as it does to me.

My heartfelt tribute is also to my late father, Dr. Mohal Lal Sukhija. It was his dream that at least one of his children would earn a doctorate, and pursuing this dream has been my guiding force throughout. I am sure he would have been proud to see this day. His blessings have accompanied me always, from school days until now.

To my beloved wife, Chavi, thank you for believing in me, for your boundless patience and encouragement even in the hard times. Your love has been a steady light guiding me forward.

My precious daughter, Myra, your smiles and laughter have been a source of endless joy and motivation. You gave me the strength to keep going when my energy waned. This achievement is dedicated to you. I am deeply grateful to my sisters, Shivani, Shweta, and Surbala, whose constant encouragement and faith in me lifted my spirits when I needed it most. A special thanks to Surbala — without your help, this would not have been possible.

To my respected family members Mr. Akashdeep Doda, Mr. Vijay Kukkar, and Mr. Naresh Kukkar, thank you for your words of encouragement, trust, and guidance which have been invaluable on this journey. The love and joy from my nephews Abhinav, Adhiraj, and Abhigyan, along with my nieces Dr. Samridhi (Dolly) and Shivanya, were uplifting and encouraging throughout my research.

I want to sincerely thank my uncle, Mr. Ravi Kant Doda, and my aunt, Seema Doda, who supported me in every way throughout my childhood. Uncle's loving discipline—especially his gentle but firm push to get me to school when I was a stubborn, naughty kid—helped set me on the right path. I am also thankful to my maternal uncles and their families who stood by us through thick and thin, providing strength and support whenever needed.

My faithful furry friend Sydney, my beloved dog, has been a constant source of comfort and joy throughout this journey. Your unwavering loyalty and quiet presence were a soothing balm during the long hours of research and writing, reminding me of the simple joys in life.

I owe a profound debt of gratitude to my esteemed guide, Dr. Arun Malik. Your expertise, patient mentorship, and unwavering support have been instrumental to my research and this thesis. I am also thankful to all my teachers and professors who have supported me in this academic journey. Finally, I thank God for blessing me with strength, opportunities, and inspiration throughout this endeavor. Your divine guidance has carried me through every step.

To all of you—family, mentors, friends, and companions—this achievement is a shared success. Your presence, support, and love have made this possible, and for that, I am deeply grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Organization of the Thesis

This thesis work focuses on the computational analysis of Hinglish, a hybrid language that combines Hindi and English. The research explores the importance of understanding Hinglish and its impact on trust sentiment, particularly considering the influence of the pandemic.

Chapter 1 introduces the topic, emphasizing the significance of computational analysis of Hinglish and discussing the effect of the pandemic on people's trust sentiment.

Chapter 2 presents a comprehensive review of the literature. It covers various aspects such as computing trust in digital interactions, pre-processing techniques, rule-based algorithms, machine learning algorithms, post-processing techniques, ensemble methods, and sentiment analysis based on aspects. The chapter concludes with a summary of the literature survey and a gap analysis.

Chapter 3 describes the methodology employed in the research. It outlines the workflow, including a comparative study of trust models, corpus collection, data collection process, and the design of a novel model called the Equation-based Digital Social Trust Model (E.D.S.T.M). The chapter also explains the construction of algorithms using E.D.S.T.M and introduces an advanced model called the Advance Digital Society Interaction and Trust Model (A.D.S.I.T.M). Furthermore, it presents content analysis techniques, including Guided Content Analysis (G.C.A) and the Topic Content Analysis Algorithm (T.C.A.A).

Chapter 4 presents the results and discussions. It provides dataset statistics, evaluates the trust models, and analyzes the performance using various evaluation metrics. The chapter also examines the results of the Digital Society Interaction and Trust Model (D.S.I.T.M) and discusses the findings of the Guided Content Analysis (G.C.A). Additionally, it includes a comparative analysis.

Chapter 5 concludes the thesis by discussing the social and technological impact of the study. It summarizes the key findings and inferences drawn from the research.

Overall, this thesis work contributes to the understanding of computational analysis of Hinglish and the impact of the pandemic on trust sentiment. The findings and insights gained from this research can have significant implications for social interactions and technological advancements.

Chapter - 1

INTRODUCTION

No doubt, the proliferation of information has led to a change in the way we communicate. Communication has become extremely quick, crisp and human contact devoid of it, especially when the whole world is suffering from covid-19 pandemic [1]. Via facial expressions or body language, confidence and trust could be judged earlier during the times of face-to-face contact. However, contact has become increasingly faceless now with the use of emails, SMS and chat bots and it is thus becoming increasingly difficult to determine trust in relationships [2]. Trust is visible through facial gestures, body language, tone, and tenor of voice, which form the foundation of relationships, and the social context is intangible. All these criteria are being driven out of the human mode of communication gradually. Trust is important to the stability and development of relationships and society [4]. Because most of the contact happens online because traditional communication is lacking, it is desirable to know the degree of trust through posts, e- messages and tweets.

With over 1 billion unique visitors per month, the social networking sites such as twitter, Facebook, Pinterest etc., are engines of communication and business. The web traffic data for Facebook, a social networking site, shows a million pieces of data transmitted every month. Each user's profile is of trustworthy users and is of vital importance to their participation on the web in online communities whose primary purpose is social networking, such as Twitter, Pinterest, Facebook, and LinkedIn. They start by developing a profile for example aseem@twitter.com when individuals join social networking sites, then create connections to established friends as well as those they meet via the web. Sharing photographs, birthdays, events, hometowns, faith, race and personal interests may also be included.

Members send a "friend" message to communicate with others, which must be approved by the other party to create a connection [2]. Another member of "friends" allows them access to profiles [5], adds them to your social network and vice versa. As the number of users increases, there is also an increase in complexity in understanding the digital world. It becomes hard to trust and be intimate on the web.

There are a lot of advantages to such mediums of communication, and these platforms continue to give a lot of opportunities for everyone to remain in touch with each other day and night. Which is important, especially when there is a pandemic such as COVID-19 is going on. But these possibilities have not come without a cost, as malicious participants are increasingly targeting these groups in an attempt to manipulate the perceived social links inherent in the management of community- based knowledge. For example, the presumed social link between users may be manipulated by malicious users to increase the likelihood of disseminating disinformation, pushing participants to the seedy side of the Internet (e.g. to sites hosting malware), and other threats to the quality of knowledge centred on the group.

In current Web 2.0 applications, some of these risks have already been observed, including impersonated (or fraudulent) digital identities, targeted malware distribution, phishing enhanced by social networks, and corrupt metadata (or tags) created by users. It is challenging to detect and mitigate this social spam and social deception, particularly as opponents continue to change their strategies and strategies [6]. With these issues in mind, there is a need for systems that compute the trustworthiness of a particular transaction in question. In India, people communicate with a new form of language known as Hinglish [7].

A language that is a mixture of Hindi and English [8]. The rampant usage of this language for communication in social media is not only because of comfort that one feels but the difference in trust that one feels while expressing this post-modern form of Hindi and English in the Indian subcontinent. Hence, any system that needs to compute the mathematics of trust, love and intimacy in an Indian context must be able to work with this language known as "Hinglish". Figure 1.1 gives examples of Hinglish being used in advertisements [9]. It can be observed from the advertisements that Bilingual Creativity is at its best in Indian advertisements currently [10].

Figure 1.1: Example of Hinglish

1.1 Importance of Computational Analysis of Hinglish:

The widespread usage of Hinglish has forced many companies and people to rethink building software with local language support. In fact, its study of its usage have become significant in a variety of natural language processing applications such as machine translation (MT) and speech-to - speech translation with its widespread use in social media such as Whatsapp, Twitter and blogs posts. In addition to Hinglish, a language known as "Benglish" is quite popular these days [11] [12]. It is a term used in academic papers to describe a mixture of Banglish (Bengali language [13]) and English in academic papers. Therefore, building a sentimental analysis application based on such a mixed language is imperative to the current research work in the field of Natural language processing. There are already multiple apps that help to detect the Hinglish language and translate them in the desired language. Keyboards that support Hinglish are already on the market. In this next section, we discuss the impact that covid 19 pandemic had on the trust sentiment of the people.

1.2 The Evolution of Trust Dynamics during Pandemics

Pandemics have historically [See table 1.1] had profound effects on the levels and types of trust within societies. Trust, both interpersonal and institutional, plays a crucial role in how communities respond to crises [15]. This section explores how different pandemics have influenced trust dynamics, cultural practices, and daily interactions, highlighting the shifts in trust across various phases.

Table 1.1: Chronology of the epidemics occurred in India

S. No	Year	Pandemic
1	Sixth cholera (1910-1911) [17]	The sixth cholera epidemic began in India and then spread to the Middle East, North Africa, Eastern Europe and Russia. Sixth cholera (1910-1911). 8, 00,000 people were killed in the outbreak.
2	The Spanish flu (1918–1920). [18]	In India, the Indian soldiers who fought in World War I were the ones who carried the messages that led to the death of many people. About 17 and 18 million Indians died of influenza, taking between 50 and 100 million lives.
3	Smallpox Epidemic (1974) [19]	Sixty per cent of the worldwide smallpox cases are in India.
4	Surat Plague (1994) [15]	The plague hit Surat in September 1994, and a large population migrated to other parts of India.
5	SARS (2002–2004):	SARS was the first serious disease that was communicable through sneezing or coughing from person to person after the 21st century.
6	Dengue and Chikungunya, (2006) [20][27][28][29]	Mosquito diseases affect individuals throughout India and costs India a whopping $5.71 billion. Predominantly seen in 21–30 years age group. The actual reported cases were far lesser than the original count.
7	Swine influenza flu epidemic (2014-2015) [21]	Gujarat, Rajasthan, Delhi, Maharashtra, and Telangana states were affected, and about 2000 people died.
8	Nipah Virus Outbreak (2018) [21]	An infection in Kerala that became epidemic within a few days was caused by fruit bats.

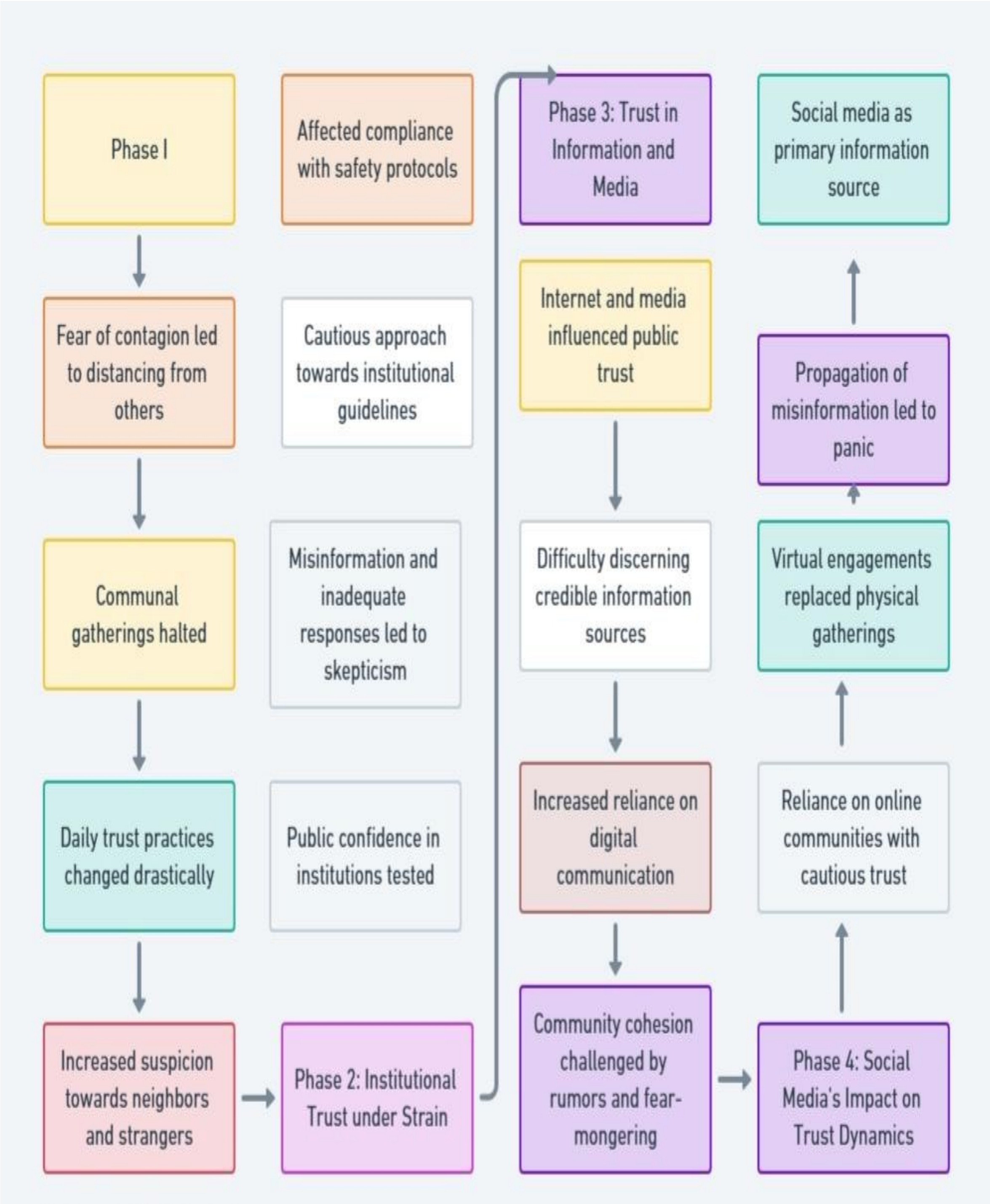

Figure 1.2: The evolution of trust dynamics during pandemic

Phase 1: Erosion of Interpersonal Trust and Social Intimacy

Associated Pandemics: Sixth Cholera Pandemic (1910-1911) and Spanish Flu (1918-1920)

During the early 20th century, pandemics like the sixth cholera outbreak and the Spanish flu led to a significant decline in interpersonal trust and intimacy. The fear of contagion caused people to distance themselves from others, even close friends and family. Cultural practices that involved communal gatherings, such as festivals and religious ceremonies, were halted. Daily trust practices changed drastically as individuals avoided physical contact and public places. Suspicion towards neighbours and strangers became prevalent, undermining the social fabric of communities.

Phase 2: Institutional Trust under Strain

Associated Pandemics: Smallpox Epidemic (1974) and Surat Plague (1994)

In the mid to late 20th century, outbreaks like the smallpox epidemic and the Surat plague shifted the focus towards institutional trust. The public's confidence in governmental and healthcare institutions was tested. Misinformation and inadequate responses led to skepticism about official communications and guidelines. Cultural practices were adjusted as people questioned the effectiveness of public health measures. Daily interactions were characterized by a cautious approach towards information from authorities, affecting compliance with recommended safety protocols.

Phase 3: Trust in Information and Media

Associated Pandemics: SARS (2002-2004), Dengue and Chikungunya (2006), Swine Influenza (2014-2015)

The early 21st-century pandemics highlighted the role of media in shaping trust. With the advent of the internet and 24-hour news cycles, people were inundated with information, both accurate and misleading. Trust levels fluctuated as individuals struggled to discern credible sources. Cultural practices evolved with increased reliance on digital communication, altering daily trust practices as face-to-face interactions decreased. Social trust was challenged by the spread of rumors and fear-mongering, affecting community cohesion.

Phase 4: Social Media's Impact on Trust Dynamics

Associated Pandemic: Nipah Virus Outbreak (2018)

The Nipah virus outbreak underscored the significant influence of social media on trust. Platforms like Facebook and Twitter became primary sources of information and interaction. While social media facilitated connectivity, it also propagated misinformation, leading to panic and mistrust. Cultural practices adapted with virtual engagements replacing physical gatherings. Daily trust practices were redefined as people relied on online communities for support yet remained wary of the authenticity of information and interactions.

Current Trends: Rebuilding Trust in the Digital Age

In the context of the COVID-19 pandemic and beyond, trust dynamics continue to evolve:

- **Types of Trust**: Emphasis on rebuilding both interpersonal trust (among individuals) and institutional trust (in governments and health organizations).
- **Levels of Trust**: Efforts to enhance trust through transparent communication and consistent public health messaging.
- **Cultural Practices**: Adoption of new rituals that balance safety and social connection, such as virtual celebrations and socially distanced gatherings.
- **Daily Trust Practices**: Increased vigilance in verifying information sources, promoting media literacy to combat misinformation.

Key Observations

- **Decline in Social Trust**: Fear of disease transmission has historically led to reduced trust among individuals, impacting social cohesion.

- **Mistrust of Authorities**: Inadequate or delayed responses from institutions can erode public trust, hindering effective crisis management.
- **Role of Media**: The proliferation of information channels necessitates critical evaluation to maintain trust in shared information.
- **Cultural Adaptations**: Societies modify cultural practices to navigate trust issues, reflecting resilience and adaptability.
- **Need for advances in Sentiment Analysis** Technology are warranted now.

1.3 Advancement Phases in Sentiment Analysis in Response to Pandemics [15 to 34 references]

The evolution of sentiment analysis has been closely tied to technological advancements and the demands of pandemic responses. This section outlines the key phases of sentiment analysis development, highlighting how each stage corresponded with the tools and methods available during different health crises

Phase 1: Manual Content Analysis and Basic Statistical Methods

Associated Pandemics: Sixth Cholera Pandemic (1910-1911) and Spanish Flu (1918-1920)

In the early 20th century, sentiment analysis was purely manual. Researchers conducted meticulous content analyses of newspapers, journals, and government reports. The focus was on basic statistical methods, such as frequency counts, to track disease spread rather than to interpret public sentiment. The limitations of this period included a lack of sophisticated tools to analyze sentiment patterns comprehensively.

Phase 2: Emergence of Survey Research and Early Computer-Based Analysis

Associated Pandemics: Smallpox Epidemic (1974) and Surat Plague (1994)

During this phase, survey research methods became prominent. Structured surveys and questionnaires were utilized to gauge public sentiment, aiding governments in formulating responses. The introduction of early computer-based analysis allowed for basic text processing, marking the transition from purely manual methods to computational assistance in analyzing sentiment data.

Phase 3: Introduction of NLP and Machine Learning Algorithms

Associated Pandemics: SARS (2002-2004), Dengue and Chikungunya (2006), and Swine Influenza (2014-2015)

The early 21st century saw the integration of Natural Language Processing (NLP) into sentiment analysis. Rule-based NLP systems facilitated basic sentiment extraction from textual data. Concurrently, machine learning algorithms, such as Support Vector Machines (SVMs), enhanced sentiment classification accuracy. These advancements allowed for more nuanced analyses of public reactions and sentiments during pandemics.

Phase 4: Adoption of Deep Learning Techniques and Real-Time Social Media Monitoring

Associated Pandemic: Nipah Virus Outbreak (2018)

Deep learning techniques revolutionized sentiment analysis in this phase. Models like Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs) significantly improved the extraction and interpretation of sentiments from large datasets. Real-time social media monitoring became crucial, enabling immediate tracking of public sentiment and facilitating rapid governmental responses.

Current and Emerging Trends in Sentiment Analysis

- **Hybrid Approaches**: The combination of machine learning and rule-based methods is enhancing accuracy and robustness in sentiment analysis models.
- **Transfer Learning**: Utilizing pre-trained models is becoming widespread, allowing for more efficient and effective sentiment analysis across different domains and languages.
- **Real-Time Analytics**: Cloud-based infrastructures are enabling rapid analysis and dissemination of sentiment data, which is essential for timely pandemic responses.

The progression of sentiment analysis technologies has profoundly impacted how pandemics are managed:

- **Early Detection**: Advanced sentiment analysis tools facilitate the early identification of shifts in public opinion, crucial for prompt interventions. This includes analyzing multilingual and code-mixed language data to capture diverse public sentiments.
- **Targeted Interventions**: Real-time analytics enable the development of strategies tailored to specific public concerns, enhancing the effectiveness of responses.
- **Collaborative Efforts**: Interdisciplinary collaboration among technologists, policymakers, and healthcare professionals is vital for maximizing the benefits of sentiment analysis in pandemic responses.

By understanding these developmental phases, researchers and policymakers can better leverage sentiment analysis tools to formulate informed and effective strategies for future health crises [32].

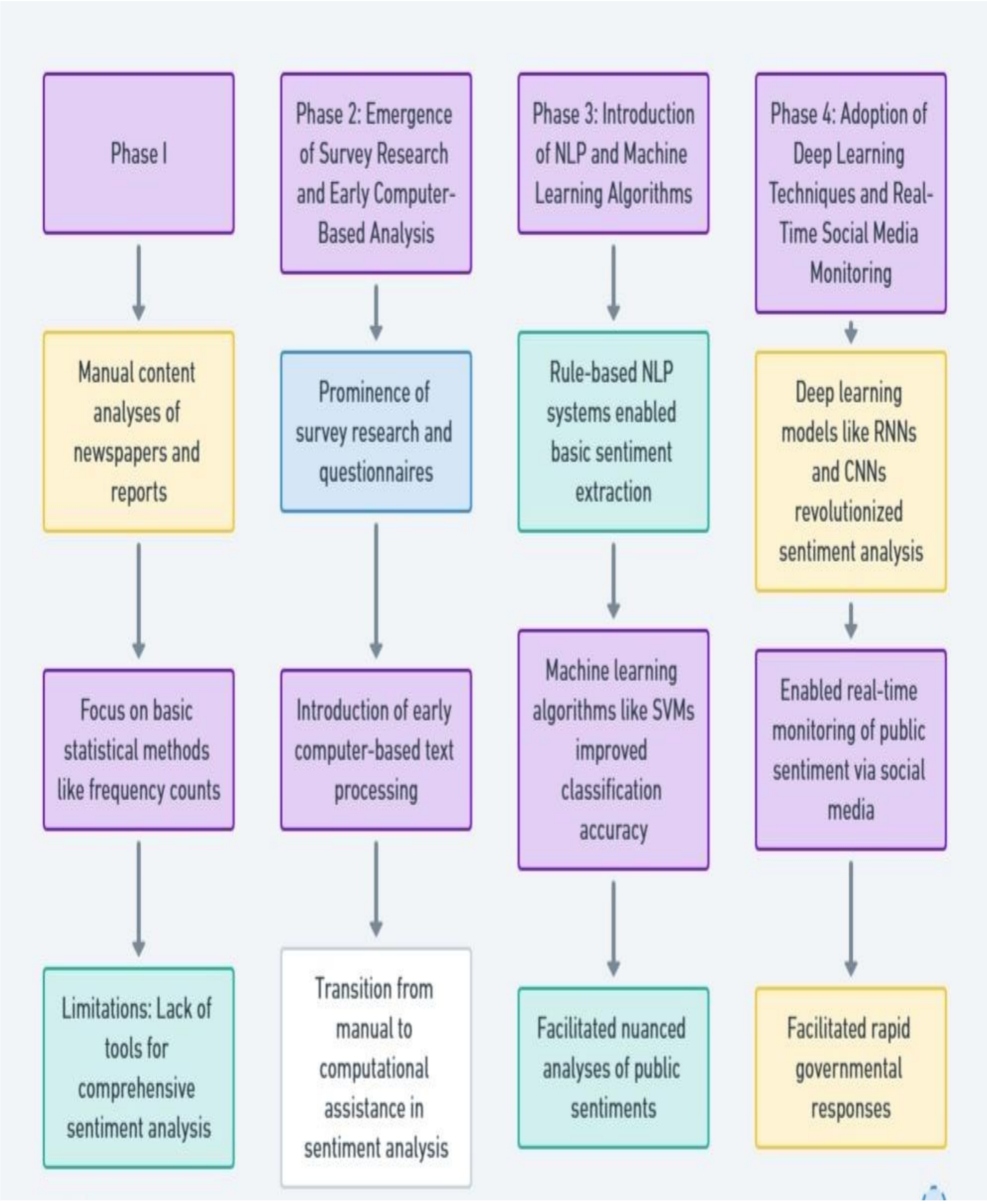

Figure 1.3 Advancement phases in sentimental analysis in response to pandemics

1.4 Sentimental Analysis in Code-Mixed Languages:

In the context of developing a solution of sentiment analysis that uses Hinglish text, there is a need to understand the rules of the English language as well as the rules that govern Hindi.

Given the Hinglish text, for example.

Eg. for eg. "It's not a nice eating place: Itna ghatiya khaana to kabhi nahi khaya"

That means, "That restaurant is not healthy. In my life, I have never had such bad food". Hence, there is a need to have consensus of the “context” the sentence is conveying in both the languages. This makes the task challenging. In context of the topic undertaken in this research, we need to extract meaningful information from the tweets, messages etc., that would help us compute the level of trust between groups of individuals that interact with each other during a pandemic, and their communication is entirely digital in nature. The tone and tenor of the language used in the tweets, their duration etc. [33], may help to measure opinions against a specific event or occurrence. The type of language used, contact frequency, tweet length are measures of the degree of confidence and trust that communicating partners have among themselves. All this can be done under the concept of sentimental analysis [32].

This approach of studying the behaviour of digitally aware individuals may be termed as Aspect based Sentiment analysis [30]. Here, the aspect is “trust”. Machine Learning and Deep learning algorithms are already making news in the field of sentiment analysis [31].

Chapter - 2
REVIEW OF LITERATURE

Outline and Introduction

As explained in the introduction chapter that code-mixed language involves speaking or writing in multiple languages. Since social media, multilingual speakers do this. Hinglish is a Hindi-English mixture. Sentiment analysis examines a text's feelings, thoughts, and beliefs. Sentiment analysis of code-mixed languages like Hinglish has grown in popularity.

As mentioned in the introduction chapter, life was not digital earlier when a series of pandemics occurred and only later on life and analysis became digital form. Hence, we discuss the digital landscape initially and evolution of trust as sentiment during the breakout of pandemics.

The chapter further covers and discusses the recent research works done in the domain of the sentiment analysis of code-mixed languages like Hinglish, including methodologies, algorithms, difficulties, and challenges.

The chapter begins with the works that describe the fundamentals and basic workflow of sentiment analysis in various cases, especially in the context of doing sentiment analysis in code mix text.

Later on the chapter devolves into the intricate aspects of sentiment analysis that includes methods of pre-processing, algorithms that improve the quality of the textual data, methods that help us build.

2.1 Digital Life

From the current literature in this context, it can be observed that trust is central to our lives, even if we meet in digital space. Furthermore from multiple studies it can be inferred that it is impossible to develop strong and lasting relationships without trust.

One of the greatest obstacles to digital trust is the problem of security. In recent years, cyber-attacks, data breaches, and identity theft have grown increasingly prevalent.

According to Satya Nadella, CEO of Microsoft, "cybersecurity is the greatest threat to the digital economy today." To battle this threat, firms are investing extensively in cybersecurity measures to safeguard the security of their customers' data. Privacy is another crucial feature of digital trust. As we share more and more personal information online, the importance of privacy increases. According to Apple CEO Tim Cook, privacy is a fundamental human right, and companies are increasingly emphasising privacy as a means of building customer confidence. Apple's devices, for instance, are designed with privacy in mind, and the corporation has made its commitment to protecting customer data clear. Transparency is yet another essential component of digital trust. According to Sundar Pichai, CEO of Google, "transparency is essential to gaining and retaining the trust of our users." This means that businesses must be upfront and honest about how they gather, utilise, and share data. In addition, their rules and procedures for processing user data must be transparent. Trust is an indispensable aspect of our digital life. As we continue to conduct a greater proportion of our personal and professional concerns online, the importance of trust increases. To establish and keep customers' confidence, businesses must prioritise cybersecurity, privacy, and openness. According to Satya Nadella, "trust is the basis of everything we do."

2.2 Trust Models in Digital Interactions

According to authors [34], under the framework of Twitter, social trust is an essential component in the process of constructing and sustaining relationships between users. The foundation of social trust is the conviction that other people will act in a manner that is in line with our expectations and that they will protect us from any harm they may do. Interactions between users on Twitter, such as retweets, likes, and comments, contribute to the development of a sense of social trust in the platform. Users develop a sense of social trust in one another when they engage in productive interactions with one another, such as by exchanging beneficial content or offering constructive comments. On the other hand, unfavourable encounters such as harassment or trolling have the potential to damage social trust.

According to many authors, Twitter (X) is also a place where social trust and economic trust both play a significant role, and both must be maintained. The foundation of economic trust is the conviction that other people will live up to their end of the bargain in business dealings, such as making payments for the goods or services they receive. The purchase and sale of goods and services on Twitter, including advertising, sponsored content, and affiliate marketing, all contribute to the development of an atmosphere of economic trust on the platform. When users engage in financial transactions on Twitter, they are required to have faith in the other party and trust that they will meet their obligations.

Many scholars consider political trust as an important factor to consider in relation to Twitter. The foundation of political trust is the conviction that governmental institutions and officials will act in a way that is to the benefit of the populace as a whole. On Twitter, political trust is developed by interactions between users and political actors like politicians and government agencies. These exchanges take place between users. Those who interact with political figures on Twitter need to have faith that these political figures will behave in a responsible and accountable manner.

It can be observed that in this context there are a variety of trust models that can be applied to digital interactions, such as hierarchical trust models, risk-based trust models, and decentralised trust models. In general, there are various models of trust that can be applied. Further, from established researcher’s work it was found that the concept that trust may be developed through the establishment of hierarchical relationships between persons and organisations forms the foundation of hierarchical trust models. In the context of Twitter, hierarchical trust models may be applicable to the interactions that take place between users and large companies or institutions, such as corporations or government agencies. Hence, the fundamental tenet upon which decentralised trust models, such as the web of trust model, are predicated is the notion that trust is cultivated by means of decentralised networks comprising both individuals and institutions. When applied to the setting of Twitter, decentralised trust models may be applicable to the interactions that take place between users and less significant businesses or individuals, such as bloggers or social media influencers.

The premise upon which risk-based trust models are based is the assumption that the degree of risk associated with a certain contact or transaction determines the amount of trust that can be placed in that interaction or transaction. In the context of Twitter, risk-based trust models may be applicable in the interactions that take place between users and strangers, which involve a higher amount of risk than other types of interactions.

It can safely be said that the trust models are an essential foundation for gaining an understanding of and successfully navigating the digital interactions that take place on Twitter and other social media platforms. Specifically, trust models in social, economic, and political contexts are essential to the process of constructing and sustaining relationships between users. Users can make educated judgments about how to connect with other people on Twitter and other social media platforms if they have a solid understanding of the various trust models and how these models apply to interactions that take place online.

2.3 How to compute trust in digital interactions

Recent work on mathematical and computational modelling show that systematically modelling the variable of trust model is a complex task. Computer scientists and linguists utilise a variety of strategies to model and compute trust in digital interactions. These models and techniques take into consideration a wide range of criteria, including user behaviour, content, network structure, and intensity of the emotions and polarity of the sentiment.

The authors [35] have used the social network analysis (SNA) model. It is one of the models that is utilised almost universally for the purpose of computing trust in digital interactions. This model determines the trustworthiness of users on the network based on the structure of the network rather than the users themselves. In order to determine which users in the network are the most trustworthy and influential, the model takes into account the relationships between users as well as the amount of followers, retweets, and mentions.

The content-based model is yet another approach to computing trust. The authors [37] have used this model to examine the content of the tweets to ascertain whether the user can be trusted. When determining whether a user can be trusted, this model takes into account not only the words, tones, and sentiments expressed in the tweets, but also other characteristics such as the user's posting frequency.

Another method that computer scientists and linguists use to compute trust in digital interactions is called the sentiment analysis model. There are many researchers that have dedicated a large part of their research time on this topic. This model determines whether or not a user can be trusted by analysing the sentiment of their previous tweets. The natural language processing methods are utilised by the sentiment analysis model in order to extract the sentiment contained within the tweets and assess whether or not they are good or negative. In addition to the models described above, linguists also employ methods from the field of discourse analysis and topic analysis to calculate trust in online interactions. The goal of discourse analysis is to determine the author's intent and the meaning that is intended to be conveyed by the text by analysing the structure of the language that is used in the tweets. In order to establish whether or not the user can be trusted, this model looks at the user's contributions in terms of how coherent, cohesive, and contextual they are.

As per the authors [5] the best way to model trust using twitter is the use of equations. The equation model is a mathematical model developed with the purpose of quantifying the level of trust that exists between two users of Twitter. The model is built on a series of equations that take into consideration a variety of characteristics that can affect a user's level of trust, such as the number of followers a person has and the frequency with which they engage with other users. The equation model is a straightforward approach that has proven to be useful in a variety of contexts, including the analysis of trust in Twitter conversations. However, few authors use the twitter data for conducting Link Analysis.

The mathematical model known as the link analysis model uses graph theory in order to investigate the trust links that exist between users of Twitter. According to this representation, every Twitter user is a node in a graph, and the connections that exist between the nodes stand in for the interactions that take place between users. The link analysis model does a study of the graph in order to determine the users' level of confidence in one another using a number of different techniques. This technique is particularly helpful for studying large datasets and discovering trends in trusting relationships.

The level of trust that exists between two different Twitter users can be calculated with the help of a probabilistic model called the Bayesian network model. This model takes into account a variety of different aspects that can influence a person's level of trust, including the quantity of followers a user has, the nature of the tweets that are posted, and the regularity with which users interact with one another. In order to calculate the likelihood that the two users may trust one another, the Bayesian network model makes use of Bayesian inference. Currently, the most frequently used method of building models is machine learning models. The machine learning model is a type of mathematical model that does an analysis of the trust ties that exist between users of Twitter by employing machine learning techniques. The machine learning algorithm in this model is trained on a dataset of Twitter conversations, and it uses this training data to find patterns in trust relationships. The Twitter dataset is used to train the algorithm. The machine learning model is particularly good for evaluating big datasets and discovering subtle patterns in trust relationships, both of which may be difficult to spot using other models. It is particularly useful for these tasks because it was developed specifically for these purposes. The random walk model is a mathematical model that analyses the trust ties between users of Twitter by employing an algorithm called the random walk. According to this representation, every Twitter user is a node in a graph, and the connections that exist between the nodes stand in for the interactions that take place between users. The random walk model employs a random walk algorithm in order to explore the graph and ascertain the degree to which users can trust one another. This technique is particularly helpful for studying large datasets and discovering trends in trusting relationships.

A number of different mathematical models of trust can be utilised to perform an analysis of the digital interactions that take place on Twitter. These models take into account a variety of aspects that can influence a user's level of trust, such as the number of followers a user has, the nature of the tweets that are posted, and the regularity with which users engage with one another. Researchers use these models to detect trends in trust connections and to build new methods for evaluating digital interactions on Twitter. Each model possesses its own unique strengths and shortcomings, and academics utilise these models in their work. When we have a better knowledge of these models, we will have a better understanding of the dynamics of trust in digital interactions, which will allow us to develop new tactics for the betterment of society.

The use of different languages in the course of a single conversation, also known as code-mixing, has become increasingly commonplace in the world of digital communication. As a consequence of this, there is an increasing demand for approaches of sentiment analysis that are able to deal with data that contains code- mixed languages. In the next piece, we will talk about the strategies that are used most frequently in code mix language-based sentiment analysis, including both pre- processing and post-processing methods.

Pre-Processing Techniques

Techniques that are used in pre-processing are extremely important to the process of sentiment analysis since they assist in cleaning the data and get it ready for analysis [36]. The following is a list of some of the pre-processing approaches that are utilised most frequently in code-mix language-based sentiment analysis:

Tokenization refers to the process of separating a string of text into its component parts, which may be words, phrases, or symbols. Tokenization is a technique that helps separate words from several languages for the purpose of code-mix language- based sentiment analysis.

Elimination of Stop Words Stop words are terms that are used frequently in a language but do not have a meaningful significance in that language. The elimination of stop words is one technique used in code-mix language-based sentiment analysis. This technique helps to improve the analysis's accuracy by decreasing background noise.

Stemming and lemmatization are two methods that can be utilised to cut down on the number of letters in a word and get it back to its fundamental form. Stemming and lemmatization are two processes that help improve the accuracy of code-mix language-based sentiment analysis. These processes contribute to lowering the dimensionality of the data and making the analysis more precise.

Tagging Each Word in a document With Its Corresponding Part of Speech Part-of-Speech (POS) tagging is the process of marking each word in a document with the part of speech that corresponds to it. POS tagging is useful in code-mix language- based sentiment analysis because it helps disambiguate words that could have multiple meanings depending on the language, they are used in.

The development of algorithms for the analysis of code-mixed language sentiment

After the data have been pre-processed, several different techniques can be used to conduct an analysis of the sentiment of code-mixed language data. The following are some of the algorithms for code-mixed language sentiment analysis that are used most frequently:

Tagging or attaching each word in a document with its category of entity such as Noun, Adverb, Adjective, Verb and so on is process that is referred to as Part of Speech Part-of-Speech (POS), this is many times the focus of many researchers who want to understand the structure of the language. It becomes critical in cases where the language is going into a lot of changes as in the case of Code Mix Language. According to researchers, POS tagging is useful in code-mix language-based sentiment analysis because it helps disambiguate words that could have multiple meanings depending on the language they are used in.

Further, when we study the methods or the workflow of the advanced sentiment analysis work, it can be observed that there are following categories of algorithms in this domain that are most frequently used.

Rule-Based Algorithms

Rule-based algorithms analyse a piece of text by comparing it to a set of guidelines from which they derive their conclusions. Rule-based algorithms may be utilised in code-mixed language-based sentiment analysis in order to give individual words or phrases a score depending on how they are perceived to convey sentiment [37].

Algorithms Used in Machine Learning

In order to train a model on a big dataset that contains labelled data, machine learning algorithms are utilised. Machine learning techniques can be utilised in code-mixed language-based sentiment analysis in order to discover the patterns and relationships that exist between words and the sentiment scores that are associated with them [39].

Hybrid algorithms combine the advantages of rule-based algorithms and machine learning algorithms to obtain a higher level of accuracy in sentiment analysis. In code-mix language-based sentiment analysis, hybrid algorithms can be used to incorporate linguistic rules and patterns into machine learning models [38]. This type of analysis is called code-mix language-based sentiment analysis.

Post-Processing Techniques

Post-processing methods are utilised to further develop and improve the findings of sentiment analysis. The following is a list of some of the post-processing approaches that are utilised most frequently in code-mix language-based sentiment analysis:

Sentiment lexicons are compilations of words and phrases that have been linked to a specific sentiment score. Sentiment lexicons can be utilised in code-mix language-based sentiment analysis to provide sentiment scores to terms that were not included in the training data.

Ensemble Methods

To improve the accuracy of sentiment analysis, ensemble methods aggregate the results of numerous algorithms into a single analysis. The usage of ensemble methods allows for code-mix language-based sentiment analysis to obtain a higher level of accuracy by combining the beneficial aspects of several different algorithms [40].

Analysis of Sentiment Based on Aspects

This is a method that Analyses the Sentiment of Individual Aspects of a Product or Service Aspect-based sentiment analysis is a method that analyses the sentiment of individual aspects of a product or service. Aspect-based sentiment analysis is a type of language-based sentiment analysis that can be utilised in code-mix language-based sentiment analysis to determine the tone conveyed by specific words or phrases within a body of text [41].

Table 2.1: Additional Tabular Summary of literature Survey

Reference(s)	Year	Indexing	Main Findings
43. Applied Sciences, 10(8), 2881. doi:10.3390/app10082881, "Predicting reputation in the sharing economy with twitter social data"	2020	Scopus & SCI	1. Trust is an essential component for social sharing-based applications. 2. Reputation can be predicted using digital artefacts. 3. New Dataset was formed.
44. Human communication Research. 2020, "Why do people share ideologically extreme, false, and misleading content on social media? A self-report and trace data–based analysis of counter media content dissemination on Facebook and Twitter"	2020	Scopus	The findings indicated that posting counter media content on Facebook is correlated with ideological extremism and with low confidence in the mainstream news media.
45. Journal of Medical	2020	Scopus &	This work demonstrated how

Internet Research 22(11) "Twitter Discussions and Emotions About the COVID-19 Pandemic: Machine Learning Approach "		SCI	Twitter data and machine learning methods can be used to conduct an infodemiology study, allowing for the investigation of changing public discourse and emotions during the COVID-19 pandemic.
46. Published in IEEE International Conference on Trust, Security and Privacy in Computing and Communications,2020, "Trust and Believe-Should We? Evaluating the Trustworthiness of Twitter Users"	2020	Scopus	The authors used random forest and support vector machine classifiers to identify political Twitter users as trustworthy or untrustworthy. To define any unlabeled, undefined records in our dataset, we used an active learning model.
47. Journal of medical Internet research. ,2021, "Artificial Intelligence–Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the United States: Observational Study"	2021	Scopus & SCI	In the United Kingdom, overall strong, negative, and neutral sentiments averaged 58 percent, 22 percent, and 17 percent, respectively, compared to 56 percent, 24 percent, and 18 percent in the United States. Concerns about vaccine safety, economic feasibility, and corporate control were found alongside public excitement about vaccine production, efficacy, and trials. We compared our

			results to those of national surveys conducted in both countries and discovered a strong correlation.
48. Published in IEEE Transactions on Network and Service Management ,2021, “Critical impact of social networks infodemic on defeating coronavirus COVID-19 pandemic: Twitter-based study and research directions”	2021	SCOPUS	One of the study's findings was that manipulation of the CO19 crisis was present while redirecting readers to irrelevant topics is not. Further data analysis proved the need for geographic diversity in a pressing concern during the current crisis. Several observations and studies have been contributed in the sense of making computation and examining possible solutions as well as for research on social networks in times of turmoil.
49. Plos One, 2021, “Don’t put all social network sites in one basket: Facebook, Instagram, Twitter, TikTok, and their relations with well-being during the COVID-19 pandemic.”	2021	Scopus & SCI	Active Twitter use was connected to higher life satisfaction, and lower social comparison, which was negatively correlated with stress. Ultimately, as I had anticipated.
50. International Journal of Advanced Research in Engineering and Technology (IJARET,	2021	Scopus	The research finds the flow of information on Twitter during the corona virus outbreak. Sentiment analysis and

2021, “Informational flow on Twitter–Corona virus outbreak–topic modelling approach”.			subject modelling was used to analyse tweets about #coronavirus following post-processing with Latent Dirichlet Allocation. The study found that the knowledge flow about the corona virus outbreak was accurate and credible, with little misinformation. LDA analysis established the most pertinent and reliable topics related to the corona virus outbreak, and Sentiment analysis verified the existence of both negative and positive emotions such as fear. Governments, healthcare agencies, and organisations successfully used Twitter to disseminate accurate and reliable information.
51. International Journal of Information Management, 2021, “Pre-and post-launch emotions in new product development: Insights from twitter analytics of three products.”	2020	SCI & Scopus	The article explains how social media analytics can be used to aid in the creation of new products (NPD). The emotions of consumers prior to and following the launch of three new products – a pizza, a car, and a

			smartphone – are compared in order to extract insights for new product growth. Negative emotions expressing disappointment with new products are consistent with business success. Trust and joy were prevalent emotions during the pre-launch period for pizza, joy for the car, and trust for the phone. In the post-launch era, anger and disgust were directed at pizza, joy and confidence were directed at the car, and joy was directed at the phone.
52. ACM Transactions on Information Systems (TOIS)., 2020, "Emotion dynamics of public opinions on twitter".	2020	Scopus & SCI	In this research, it was found that the social dynamics of emotion expressed in user opinions in order to better understand how users' emotions against a social problem change over time, (ii) the effect of collective emotions on individual emotions, and (iii) the causes of shifting opinion due to social factors. The researchers

			found that users' emotion dynamics over a series of 17.65M tweets from 69.36K users and discovered that 63% of users are likely to change their emotional state about the subject in subsequent tweets. The fact that tweets originated from the member group demonstrates a greater capacity for influence than other community outlets. Additionally, it is found that retweets have a greater impact on users than hashtags, mentions, and replies.
53. Government Information Quarterly, 2020, "Linguistic analysis of municipal twitter feeds: Factors influencing frequency and engagement"	2020	Scopus & SCI	The article investigates the relationship between linguistic factors and municipal Twitter interaction and frequency of tweets. Utilizes tools to analyse the linguistic style of Twitter feeds from the 100 most populated cities in the United States. Engaging accounts were more nervous, emotionally diverse, and affiliative in tone.

			The style of the more involved accounts was more contemporary, casual, complex, and feminine. Tweet behaviour was found to be negatively correlated with citizen participation.
54. Nature communications, 2021, “Cognitive reflection correlates with behaviour on Twitter.”	2021	Scopus & SCI	The authors discover that individuals who scored higher on the Cognitive Reflection Test—a commonly used indicator of analytical thinking—were more selective in their social media use, as shown by the styles and number of accounts followed, as well as the reliability of the news outlets posted. Additionally, a network study suggests that the concept of echo chambers, in which dialogue is more likely to occur with like-minded others, is not exclusive to politics: individuals with lower cognitive reflection scores appeared to adopt a collection

			of accounts that individuals with higher cognitive reflection scores avoid. According to these researchers, their findings shed light on the factors that influence activity on social media sites and call into question intuitionist ideas that critical thinking is unimportant for everyday judgement and decision-making.
55. Proceedings of the Royal Society A. ,2021, “A trust model for spreading gossip in social networks: a multi-type bootstrap percolation model “	2020	Scopus & SCI	The authors implement a multi-type bootstrap percolation model, dubbed T-Bootstrap Percolation (T-BP), and use it to investigate knowledge dissemination in social networks. A social network is represented in this model by a graph G, each vertex of which has a label indicating the type of position the individual plays in the network (e.g. a student, an educator etc.). Once an initial set of vertices in G is randomly chosen to carry gossip (e.g., to be infected), the gossip propagates to a

			new vertex if it is spread by a minimum threshold of vertices with distinct labels. We investigate the T-BP model's various properties through numerical simulations and discuss their consequences for rumour spread, false news. and marketing strategies.
56. Computing, 2021,A propagation trust model in social networks based on the A* algorithm and multi-criteria decision making.	2021	Scopus & SCI	Research results show that trust metric value propagates through trust distances with varying average path lengths while maintaining the accuracy of inferred trust values between each unconnected pair of nodes. The evaluations were conducted on Facebook and Twitter networks with varying topologies, and the results were compared to those obtained using the TidalTrust and MoleTrust algorithms.
57. Information Fusion 2020, A trust-similarity analysis-based clustering method for large-scale group decision-making under a social network.	2020	Scopus & SCI	The aim of this research was to establish a trust-similarity analysis (TSA)-based clustering approach for managing the clustering process in LSGDM events

			that occur within a social network context. The authors created a trust- similarity matrix to collectively define the decision details. Second, the values of all measurement attributes are mapped to a trust-similarity plot, from which the joint threshold can be determined.
58. Advances in Intelligent Systems and Computing, vol 1287. Springer, Singapore,2021 "A Novel Approach for Sentiment Analysis of Hinglish Text"	2021	SCI	Hinglish is a written language that combines Hindi and English words, phrases, and slang. This work proposes a novel way to Hinglish sentiment analysis. Preprocessing the data using methods like Stemming, Levenshtein distance, and Soundex index is proposed. The experiment shows that the proposed method works well for sentiment analysis of Hinginglish text.
59.ArXiv:2102.12149,2021 "Sentiment Analysis of Code-Mixed Social Media Text"	2021	Scopus	In this paper, author discussed various results obtained for different techniques applied for performing the sentiment analysis of social media (Twitter) codemixed text

			written in Hinglish and data was processed through various stages such as data consolidation, cleaning, transformation, and modelling. Data was transformed using count vectorizer, one hot vectorizer, tf-idf vectorizer, doc2vec, word2vec and fasttext embeddings. The models were created using various machine learning algorithms such as SVM, KNN, Decision Trees, Random Forests, Naïve Bayes, Logistic Regression, and ensemble voting classifiers. The models created were evaluated using the F1-score (macro). The best F1-score of 69.07 was achieved using ensemble voting Classifier.
60. Proceedings of the LREC 2020 4th Workshop on Computational Approaches to Code Switching, 2020," Sentiment Analysis for Hinglish Code-mixed	2020	Scopus	Study investigates the use of unsupervised cross-lingual embeddings to decode code-mixed social media content. Embeddings can be used to train a sentiment model in one language and assess it in

Tweets by means of Cross-lingual Word Embeddings”.			another language projected in the same space. We use these embeddings to perform sentiment analysis on Hinglish Tweets, which combine English and Hindi (transliterated).
61. Indian Journal of Science and Technology, 2020, “Annotated corpus creation for sentiment analysis in code-mixed Hindi-English (Hinglish) social network data”.	2020	Scopus	Many applications now require sentiment analysis of tweets, blogs, comments, and postings. We offer an annotated corpus of code-switched social media text in Hindi, English and Hinglish. The analysis is based on word polarity (positive, negative, or neutral) and the accuracy can be improved by considering code-mixed text. The proposed corpus can be used for market research, customer behaviour, polling, brand monitoring, etc.
62. ArXiv:2010.11019, 2020, “LT3 at SemEval-2020 Task 9: Cross-lingual embeddings for sentiment analysis of hinglish social media text”	2020	Scopus	Hinglish Sentiment Analysis combines Hinglish and pre-trained English FastText word embeddings in the same area. This approach has the best results, with an F1-score of 70.52 percent on the held-out test data. Pre-trained

			English embeddings are gradually retrained using Twitter compare hinglish tweets.
63. Proceedings of the International Conference on Innovative Computing & Communications (ICICC) 2020," Current State of Hinglish Text Sentiment Analysis"	2020	SCI, Scopus	The key issue is categorizing aspect-based sentiment in Hinglish. Choosing the right categorization model is crucial. The properties of emotion text data and their extraction methodologies are explained. Text and textual features of several ML approaches for real word analysis were also analyzed
64. Information Sciences Volume 508, January 2020, "measuring trust in social networks based on linear uncertainty theory."	2020	Scopus, SCI	Uncertainty theory is a mathematical approach that investigates expert belief. It provides a new way to measure social network trust. This paper applies uncertainty theory to social network modelling. The recommended trust value is determined from direct trust values. To reduce secondary uncertainties generated by subjective weighting in multi-node, multi-path chains, two weighted trust aggregation operators are introduced.

65. Advances in Cybernetics, Cognition, and Machine Learning for Communication Technologies, 2020, “Lexical Analysis and Mathematical Modelling for Analysing Depression Detection of Social Media Reviews”	2020	SCI and Scopus	As a result, social media has emerged as the primary source of information for identifying the physical and mental illnesses of such individuals. A multidimensional depression detection model is proposed in order to examine how different kinds of social contract influence the interaction of a user with other people on social networks. Patients' circumstances worsen as a result of the fact that the vast majority of those suffering from depression do not seek medical assistance in the early stages of the illness. In the meanwhile, many individuals are seeking assistance from social media in order to express their emotions and share their deeds.
66. Hindawi Research Article | Open Access, Volume 2020, “A Discrete Mathematical Modeling and Optimal Control of the Rumor Propagation in	2020	Scopus	In this research work, proposed a new approach based on the cholera model to consider the expert pages specialised in the dissemination of rumours

Online Social Network”			from an existing IRCSS model. Also authors recommended an optimal control strategy to fight against the spread of the rumors on social media and emphasised upon the three optimal controls which minimise the number of spreader users, fake pages, and corresponding costs. To demonstrate the theoretical results obtained, authors proposed a numerical simulations for several scenarios applying the forward-backward sweep method (FBSM) to solve our optimality system.
67. Modern information technologies and IT education,2019 “Mathematical Modelling of the News Spreading Process in Social Networks”	2019	Scopus	A mathematical model of news propagating from social network posts. Epidemiological modelling is used to create mathematical models of news spread. For rumour control, a redesigned model is provided incorporating media awareness as a control approach. People can't tell if these posts are true or false.

			This study presents a mathematical foundation for these issues.
68. Doctoral dissertation, Queen Mary University of London),2015, “Mathematical modelling of the statistics of communication in social networks.”	2015	Queens Mary London University, PhD Thesis	As a result, social media has emerged as the primary source of information for identifying the physical and mental illnesses of such individuals. A multidimensional depression detection model is proposed in order to examine how different kinds of social contract influence the interaction of a user with other people on social networks. Patients' circumstances worsen as a result of the fact that the vast majority of those suffering from depression do not seek medical assistance in the early stages of the illness. In the meanwhile, many individuals are seeking assistance from social media in order to express their emotions and share their deeds. A mathematical model of news propagating from social network posts.

			Epidemiological modelling is used to create mathematical models of news spread. For rumour control, a redesigned model is provided incorporating media awareness as a control approach. People can't tell if these posts are true or false. This study presents a mathematical foundation for these issues. Chat rooms are one of the most interactive internet places, attracting a lot of social network researchers. But genuine social networks are dynamic, and RWT changes over time. Our research shows that the RWT distribution is multi-scale, which significantly alters current ideas about RWT. We also identify elements that could reduce response waiting time and improve friendship connections. These findings imply that the time context or environment has a significant impact on users' RWT.

69. ACM Computing Surveys. Vol. 45, No. 4A "survey of trust in social networks, 2020" A survey of trust in social networks "	2013	Scopus	The authors looked at trust from the perspectives of sociology, psychology, and computer science. This study classified trust as calculative, relational, emotional, cognitive, institutional/ systemic, and dispositional in nature. The self-reinforcing and event-sensitive components of trust have gotten less attention. A social network's most important feature is its social capital, or collective value. Social capital can be used to build social network trust. From this perspective, the authors found two aspects of trust in social networks: sociological and computational. On compares the literature and identifies areas for social trust research.

2.4 Summary of the Literature Survey

Code-mixed language involves speaking or writing in multiple languages. Since social media, multilingual speakers do this. Hinglish is a Hindi-English mixture. Sentiment analysis examines a text's feelings, thoughts, and beliefs. Sentiment analysis of code-mixed languages like Hinglish has grown in popularity. This chapter covers sentiment analysis of code-mixed languages like Hinglish, including methodologies, algorithms, difficulties, and challenges.

This chapter begins with research on sentiment analysis in code-mixed languages. Lack of labelled data makes code-mixed language sentiment analysis difficult. Machine learning algorithms need labelled data, but code-mixed languages are hard to label. Crowdsourcing, active learning, and transfer learning have helped researchers solve this problem. Researchers in this domain also work on embeddings, sentiment lexicons, and part-of-speech tags to improve sentiment analysis algorithms. This chapter's second section covers sentiment analysis methodologies. To create topic modelling and sentiment analysis algorithms, tokenization, stemming, and stop-word removal are needed. Pre-processing can reduce data dimensionality and improve algorithm accuracy. N-grams and feature selection can also improve sentiment analysis systems. Machine learning algorithms for sentiment analysis are covered in the third section. Sentiment analysis in code-mixed languages uses Naive Bayes, Support Vector Machines, and Deep Learning. Researchers have used supervised and unsupervised learning to improve these algorithms. Convolutional and Recurrent Neural Networks have been utilized for code-mixed language sentiment analysis with encouraging results. Content analysis methods conclude this chapter. Content analysis examines a text's emotions, views, and attitudes. Speech act theory divides speech acts into locutionary, illocutionary, and perlocutionary categories. Speech act theory can help sentiment analysis by analysing a text's emotions, views, and attitudes. In nutshell, code-mixed language sentiment analysis is difficult but possible with the correct approaches and algorithms. Addressing the absence of labelled data and language complexity are issues. Crowdsourcing, active learning, and transfer learning have helped researchers overcome these obstacles. To develop algorithms for topic modeling and sentiment analysis, it is essential to incorporate tokenization, stemming, and the removal of stop words. Naive Bayes, Support Vector Machines, and Deep Learning have also shown promise. Speech act theory can help sentiment analysis. Sentiment analysis of code-mixed languages can be used for social media monitoring, market research, and customer feedback analysis with further development. In this chapter, we touched on multiple aspects of building models of digital social interactions that involve trust or lack trust between the people involved. Further, the literature can be summarized with the following key points:

i. Social networking has become a staple of social interaction and has opened up new and potential doorways for cyber-crime to almost all age groups worldwide. This was generally seen in the period of lockdowns. Many people were forced to change their dealings on digital platforms. Hence, there is a need to build systems that can help to compute trust and build signals of confidence between the parties. Most people are interacting informally on digital platforms and are using mixed languages such as Hinglish.

ii. There is a need to develop new capabilities for processing Hinglish languages in terms of natural language processing, machine translation, and sentiment analysis.

iii. Improving digital forensic capabilities is the need of the hour, especially when everyone is working digitally.

iv. There is a need for building standardised forensics APIs and models of digital social interactions.

v. In contemporary literature, limited work can be found in the context of building models of digital social interactions that help to map trust deficits in digital societies.

2.5 Research Gap Analysis

From the literature survey's findings, it can be safely said that the computational analysis of mixed languages is quite a challenge due to cross embeddings and variation in the meaning of the words with respect to the local language. From the literature survey, the following gaps have been identified:

i. Demographic Variation: There is no standard per se "Hinglish" language. There are a lot of variations that happen due to geographic regions. Such variations increase the complexity of understanding the elements of the language. Phonetic variation is another challenge that needs to be addressed for doing this research work.

ii. Rule based algorithms will find it hard to find logic for extraction of the elements that give hints about a particular sentiment. This is because Hinglish has no fixed rules of grammar. The connotations of the grammar rules applicable to a particular sentence require tracing or finding similarity with the Hindi or English grammar rules.

iii. Spellings: There is little agreement on the spellings of the words used in the Hinglish. Hence, the development of a new open-ended dictionary is necessary and a tedious job.

iv. Due to all the above-mentioned problems, it is clear that there is a need for building custom parts of speech algorithms, tokenizers, stemmers, etc.,

v. Public availability of datasets that can help to compute trust between individuals is not available. Hence, there will be a need to conduct a two-level analysis of social media data. First, there will be a need to find people interacting in their close circles using some social media vehicle, and second, to develop a crawler that would get their publicly available "hinglish" content.

vi. Limited machine learning or deep learning models of sentiment analysis can be found that focus on collecting and analysing the trust deficit between the people interacting digitally. Moreover, there is no agreed definition of trust and a trust deficit in contemporary literature. This makes the work quite challenging.

2.6 Research Objectives

Relationships make social media social and successful. However, online social media treats all users the same way: as trusted friends or total strangers, with little or nothing in between. Relationships fall everywhere along this spectrum—from total strangers to close friends. Trusted friends and family members can affect emotional health and often join together to lead organizations through tough times such as the COVID-19 pandemic. Since relationships are now becoming predominately online, It is therefore of interest to investigate how well online social media data can predict trust among individuals. This is especially important as individuals are not in physical proximity to each other and yet there are lives are influenced by the communication they have online. In this study, we have present a predictive model that maps social media due to trust. Regarding the research work, the following objectives have been finalised:

1. To study contemporary models of trust and sentiment analysis algorithms.
2. To collect the corpus of Hinglish text messages that indicates trust, intimacy, and intensity in social relationships and to analyse the content of the crawled corpus using NLP techniques.
3. To design and develop a novel model for social interactions, trust and trustworthiness.
4. To compare and validate the proposed model of social interactions and trust with the state of the art.

Problem Statement

The problem revolves and lies in accurately modeling trust in digital interactions, especially in multilingual, code-mixed environments like Hinglish. Despite the exponential growth of sentiment analysis models and tools, there remains a gap in trust modeling that combines linguistic analysis, interaction frequency, and user network structure. The challenge is to develop a robust model that quantifies trust in code-mixed social media data by integrating content attributes and interaction metrics through equation-based models and machine learning approaches.

Mathematically, the problem is represented as followings:

1. **Computation of Trust Score (T)**: The dependent variable representing the computed trust level between users.
    - **T=f(S,E,A,N)T = f(S, E, A, N)T=f(S,E,A,N)**, where:
        - **S**: Sentiment score derived from NLP sentiment analysis of tweets.
        - **E**: Engagement metrics (e.g., likes, retweets, comments).
        - **A**: Algorithm-specific adjustments (e.g., machine learning enhancements via active learning).

- **N**: Network structure and position within the social graph, using elements like follower count and mutual interactions.

2. **Computation of the Sentiment Score (S)**: Derived from the NLP content analysis, which examines linguistic attributes, including sentiment, topic, and tone.
3. **Equation-Based Modeling (EBM)**: Utilizes mathematical representations (e.g., probabilistic and link analysis) to structure the relationship between trust and interaction dynamics on Twitter.
4. **Machine Learning Enhancements (ML)**: Incorporates machine learning models for further advanced sentiment and trust prediction, adaptable across multilingual and code-mixed data.

Chapter - 3
METHODOLOGY

This chapter outlines the methodologies used to achieve the research objectives of the study. It describes in detail the comparative analysis of various methods, techniques, algorithms, pipelines, datasets, and other relevant aspects that were examined to build social trust models for digital social research.

The fundamental design of the social trust models and the various mathematical constructs used to develop them are explained in separate sections of the chapter. The models were developed with the aim of enhancing social trust between individuals and communities in the digital realm.

The chapter also includes a detailed explanation of the content analysis of the collected corpus. However, before diving into the methods and results, it is important to establish the assumptions that were made for this study.

Firstly, the data for building the corpus was collected during the period of the COVID-19 pandemic and lockdowns in India, spanning from February 2020 to December 2020 and any extended period if applicable. The purpose of focusing on this time frame was to capture the unique and significant changes in social interactions brought about by the pandemic and lockdown measures.

Secondly, the qualification criteria for the building corpus included two conditions: the text must be written in Hinglish, a combination of Hindi and English languages, and participants must be actively engaged in social interactions and have an interactive circle. These conditions were set to ensure that the collected data accurately reflects the social dynamics of the target population.

It is important to note that these assumptions lay the foundation for the methods and results presented in this chapter. The chapter provides a thorough analysis of the data collection and analysis process, including the various steps taken to ensure the reliability and validity of the findings. For better understanding the workflow followed in this research work block Figure 3.1 may be referred to.

3.1 Workflow of Research

In this study, the research workflow is designed to investigate social trust specifically within the context of the COVID-19 pandemic. The unique social dynamics during the pandemic, spanning from February to December 2020, provide an ideal setting to observe how trust develops, fluctuates, and adapts in response to widespread social disruptions. Each phase of the research, from data collection to modelling, has been structured to capture trust-related signals within this pandemic-specific timeframe, ensuring that the insights gained are deeply connected to the social context influenced by COVID-19. While the models and methods presented may be adaptable for broader applications, this study is purposefully anchored in pandemic-era data to address the research objectives within this critical period.

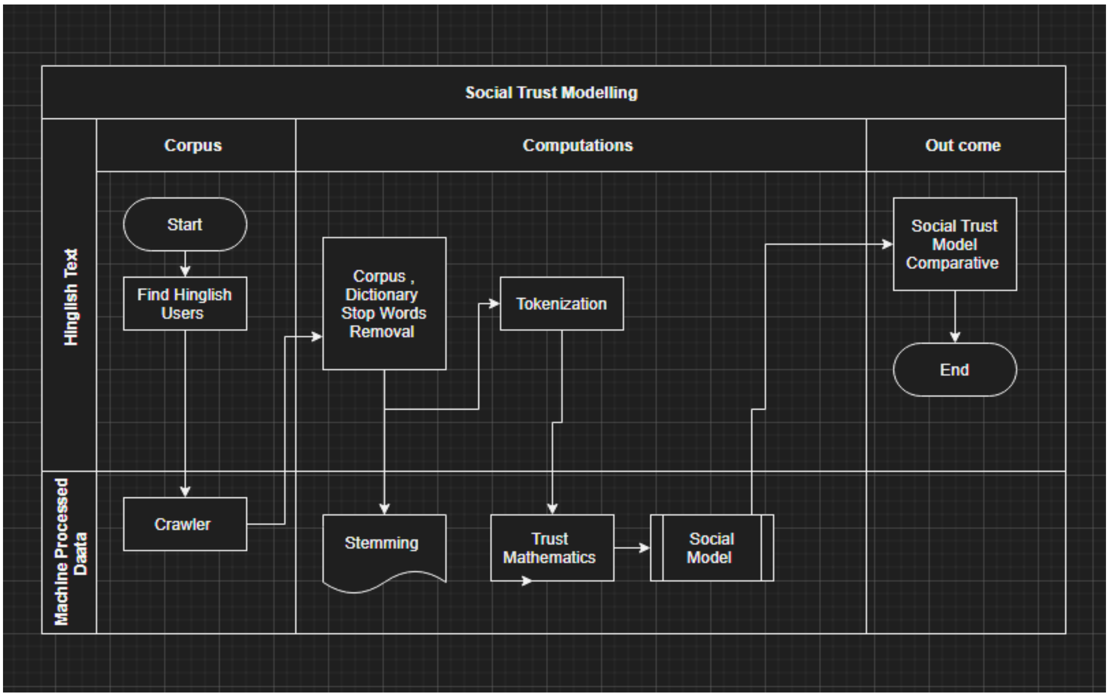

Figure 3.1: Workflow of the Proposed Methodology

The second step involves linguistic content analysis of the obtained corpus. This step involves three parts: annotation, abstraction, and analysis. Annotation involves understanding the different parts of speech and components of the speech. Abstraction involves mapping terms in the scheme of words that makes sense for a particular objective, and analysis involves the process of analysing data statistically. The third step is mathematical computational modelling. This step will use a mathematical formula and derivations or computing trust. The collected data will undergo graph theory-based analysis to extract and compute many parameters for building regression, classification, and agent-based models of trust. The corpus analysis will involve checking the distribution of word frequencies in the corpus. A dictionary of words representing these emotions will be created for the purpose of classifying these emotions into one of the three categories. The fourth step is the establishment of models of interactions and trust. Formalization of digital trust mathematics and models will be done, which may involve the construction of simulation and prediction models. The final step is the validation of models and work. These steps will help to obtain the approved research objectives by performing a comparative study of various methods, techniques, algorithms, pipelines, datasets, and other aspects related to building social trust models that can be employed in digital social research.

Table 3.1: Frame for Collection of Hinglish Corpus.

Why /Purpose	To collect the corpus of Hinglish text messages that indicates trust, intimacy, and intensity in social relationships and to analyse the Content of the crawled corpus using NLP techniques.
What	Highlish words and phrases that convey trust and intimacy levels
Where	Historical data and live steam data from India Geographic area
Who	Everyone who is talking about a corona and relationship etc.
How	Filtering using custom code logic
Language	Hindi + English, Hinglish
Outcome	Data Mining and exploratory analysis of conversation between online communities after the creation of the dataset.

Using these principles and guidelines, a creation of the corpus is underway. The next step is tokenization and noise suppression. This is done so that frequency and occurrences of keywords extracted are ready for exploratory and comparative analysis and modelling.

3.2 Comparative Study of Trust Models

In this section, we cover work done for the attainment of the first objective i,e. to conduct a comparative study of trust models.

From contemporary literature it is amply clear that 'Trust' is a fundamental feature of human connection and is gaining significance in a variety of domains, including the social, political, emotional, economic, psychological, and computer sciences. Models of trust are utilised to quantify and analyse the level of trust between individuals or entities. In this comparative study, numerous trust models, including social, political, emotional, economic, psychological, computer science, and hybrid models, will be examined. [See table 3.2]

Table 3.2: Comparative Analysis of Models of Trust

S. No	Model	Description	Limitation
1	Social Model	Based on social relationships and interactions between agents.	Data sparsity and imbalance, especially in capturing trust expressions specific to a pandemic setting.
2	Political Model	Models trust based on power and authority relations between agents.	Variability in trust expression across platforms can impact consistency and generalizability of model outcomes.
3	Emotional Model	Focuses on emotional connections and relationships between	Sentiment analysis limitations due to code-mixed language (Hinglish) reduce accuracy in

		agents.	emotion-based trust metrics.
4	Economic Model	Relies on economic transactions and exchanges between agents.	Temporal constraint limits findings to pandemic-specific behaviour, affecting generalizability to non-crisis contexts.
5	Psychological Model	Examines cognitive and psychological factors influencing trust.	Challenges arise in detecting non-verbal trust cues, limiting insights into nuanced trust dynamics in digital text.
6	Computer Science Model	Uses computational algorithms and data analysis techniques for trust evaluation.	Automated and proxy accounts introduce noise and skew trust metrics, complicating accurate measurement.
7	Hybrid Models	Combines multiple models to enhance trust modelling accuracy.	Complexity in combining multiple domains may increase computational cost without necessarily enhancing insights.

From Table 3.2 it can be inferred that each trust model has its own advantages and limitations, and hybrid models can provide more accurate and robust results. The selection of a trust model or combination of models depends on the specific application and the nature of the relationships being modelled. Mathematically, some of the important models that are noteworthy are as follows:

Social trust models can also be represented mathematically. For example, the Expectation

Confirmation Theory (ECT) model can be expressed as:

$$ECT = \delta 1E + \delta 2\sum i{=}1n(\alpha iVi - Ci) + \varepsilon$$

Where, E is expectation, V is perception of system value, C is confirmation of expectations, n is the number of confirmation items, and δ1, δ2, αi, and ε are coefficients.

Political trust models can also be represented mathematically, such as the calculus-based model of political trust, which is derived from rational choice theory:

$$PT = \beta 1 + \beta 2PS + \beta 3PF + \beta 4MP + \varepsilon$$

where PT is political trust, PS is perception of security, PF is perception of freedom, MP is perception of material prosperity, and β1, β2, β3, β4, and ε are coefficients.

Emotional trust models, such as the Trust in Automation model, can be expressed mathematically using a combination of Bayesian statistics and computational models:

$$TA = P(A|O) = P(O|A)P(A) / P(O)$$

Where TA is trust in automation, A is the automation system, O is the observed data, and P(A), P(O|A), and P(O) are the prior probability of A, the likelihood of observing O given A, and the marginal probability of observing O, respectively.

Economic trust models, such as the Agency Theory model, can also be expressed mathematically:

$$AT = f(\omega i, \gamma, \theta, \varepsilon)$$

Where AT is agency trust, ωi is the effort level of the agent, γ is the incentive provided by the principal, θ is the risk aversion of the agent, and ε is the unobserved error term.

Psychological trust models, such as the Dispositional Trust model, can also be expressed mathematically:

$$DT = \alpha 1 + \alpha 2OS + \alpha 3NN + \alpha 4FE + \varepsilon$$

Where DT is dispositional trust, OS is openness to experience, NN is neuroticism, FE is agreeableness, and α1, α2, α3, α4, and ε are coefficients.

Computational trust models, such as the Peer-to-Peer Reputation System, can be expressed mathematically as:

$$CR = f(Ri, Tj, Wk)$$

Where CR is the computational trust rating, Ri is the reputation of the individual, Tj is the trust value of the interaction partner, and Wk is the weight of the feedback.

To create these models, it was found that the most frequently used computer language is Python, and libraries such as Scikit-learn, TensorFlow, PyTorch, and others are extensively used for this purpose. These libraries provide various tools for data analysis, machine learning, and deep learning, which can be used to create mathematical models for trust. By utilising these tools, one can also analyse and compare the performance of different models and select the most suitable one for a particular application. In this next subsection, we tabulate our findings from the comparative analysis of multiple models of trust that get impacted due to some kind of stimuli such as covid 19 pandemic events.

Table 3.3: Mathematical Representation of Models

Model	**Mathematical Expressions**	**Advantages**	**Limitations**
Reputation-based	Trust = f (reputation)	Scalability, adaptability	Vulnerability to attacks
Social-based	Trust = f (social network)	Flexibility, better discrimination	Inaccurate information
Cognitive-based	Trust = f (perceptions)	Good discrimination, high-level of control	Highly subjective
Emotion-based	Trust = f (emotions)	Intuitive, subjective, fast decision-making	Limited application, interpretability
Game Theory-based	Trust = f (rationality)	Strong theoretical basis, good for strategic interactions	Limited to rational behaviour
Hybrid-based	Combination of multiple models	Improved accuracy, flexibility	Complexity, computational cost

Table 3.3's comparison analysis reveals that all of the trust models covered in this study are complex functions of several variables or factors. Identification and

selection of pertinent variables are vital to the development of any particular trust model. The variables included in the model will vary according to the type of model being produced and the application situation.

Demographic characteristics, for instance, play a key impact in the development of social trust models. Among the demographic variables are age, gender, race, and degree of education. These variables can have a substantial impact on a person's trust in others and their decision-making processes.

Similarly, variables such as facial expressions, body language, and tone of voice are significant in emotional trust models. These characteristics are indicative of emotions and can be used to determine a person's level of trust for another.

In addition, the development of trust models requires the use of numerous theoretical frameworks. The Speech Act Theory is one such paradigm that seeks to comprehend the complexities of digital written utterances. This theory acknowledges that communication is a social act and that the circumstances in which a message is conveyed can affect its meaning. The Speech Act Theory can be utilised to create models that capture the nuances of digital communication and enhance the precision of trust assessments.

Moreover, a number of emotion theories can be utilised to the development of trust models. The Appraisal Theory, for instance, posits that individuals analyse their environment and assess its possible effect on their well-being. This idea can be used to construct models that anticipate an individual's level of trust based on their emotional response to a certain event.

Table 3.4: Modelling Approaches

S. No	Model /Approach	Description	Purpose
1	Schelling [33] : Model for understanding racial and Income segregation or on the basis of some kind of rule, premise or principle.	It is a well-established fact; the people at large are generally segregated on the basis of income, ethnicity, race and based on different cults.	Trust travels at different speeds with complex dynamics in such societies. When such societies are digitally enabled the behaviour and attitude towards the spread and control of the epidemics need careful observation and research mechanism
2	Granovetter [34], [35] : Models for understanding collective behaviour of groups based on threshold/criteria or some kind of rule	For understanding the collective behaviour equilibrium. Research shows that based on threshold or criteria large groups unintentionally take decisions that overall impact the characteristics of a group/society.	In the event of trust being won or lost, many people accept it unknowingly, which alters the feature of collective trust invested in specific premises, according to research.
3	Standing Ovation Model [36]	Models that help to understand the peer pressure and effects.	People lose or gain confidence and trust with respect to issues in life based on evaluation of their peers.
4	Identification Models (Dmitry Sivaev et al. 2020; Liming Zhang et al. 2018)	Basis on which the individual's behaviour and collectives act	Besides many categories of human race, colour and education. People also identify themselves with ideas such as Anti-Vaccination or Pro-Vegetarian. Trust in such cases can be assessed by building identification models
5	Equation Based Models [39]	Model based on identified mathematical relationships (input variables and output)	Trust can simply be modelled using mathematical equations by considering trust as a function of other factors.

6	Rational Actors Model/Aggregation Models [40], [41]	The rational choice theory assumes that individual actions lead to aggregate social behaviour. Individually, the theory predicts the agent will choose the action (or consequence) they favour.	The degree of trust an individual feels ultimately aggregates for the sum total behaviour of the society.
7	Percolation models /Susceptible (SIS) [42]Models/Tipping points	Models that help to understand the spread of phenomenon such as deadly disease in an area, state of war etc.,	Useful in the context of the current corona -virus kind of situation and its impact on the social fibre.
9	Markov Models [43], [44]	The Markov Model attempts to explain a random process that is dependent on the present event but not previous events. We employ the Markov Chain Model for fully observable dynamic systems and the Hidden Markov model for partially observable systems. Probabilistic or stochastic models seek to predict the behaviour of a randomised independent process. Markov Models, on the other hand, attempt to explain non-random processes.	Trust in the digital society can be modelled using the Markov principles.

There are other classes of models that are also useful for understanding the dynamics of various systems and subsystems such as society. The table 3.4, gives information on such models.

Schelling's approach is useful because it assists in comprehending how segregation based on money, ethnicity, race, and cults influences the transmission and control of epidemics in digitally-enabled society. This model does not account for all of the elements that contribute to the spread of epidemics, including cultural beliefs and practices.

Understanding the collective behaviour of groups based on a threshold/criteria or rule is facilitated by Granovetter's models. This model facilitates comprehension of the equilibrium of collective behaviour. This model does not account for the subtleties of individual behaviour and trust.

The Standing Ovation Model is important for comprehending the influence of peer pressure on an individual's confidence and trust. Understanding the influence of peer pressure on individual behaviour is facilitated by this paradigm. Nevertheless, this model is limited in that it only accounts for the influence of peer pressure and ignores other elements that contribute to trust.

Assessments of trust based on identification models benefit from the use of identification models. These models are useful for comprehending how individuals define themselves and how this affects their level of trust. Unfortunately, these models are limited since they do not account for every element that influences trust.

Equation-based models are effective for modelling the relationship between trust and other variables. These models are useful for comprehending how other factors influence trust. However, these models have limitations because they only account for the variables included in the equation.

The Rational Actors Model/Aggregation Model is important for comprehending how individual acts influence the aggregate level of trust in society. These models are useful for comprehending the connection between individual and group behaviour. Nonetheless, these models are limited since they presuppose rational decision-making and do not account for subtleties.

On the basis of the table's contents [3.4], it can be inferred that there are numerous models and approaches accessible for comprehending and modelling trust in a digital society. Each model has its own benefits and drawbacks, and their implementation depends on the environment and research problems at hand.

For instance, models such as Schelling and Granovetter's threshold models are helpful for comprehending the collective behaviour of groups, but the Standing Ovation Model is excellent for comprehending the influence of peer pressure. Equation-based models and Markov models are helpful for modelling trust as a function of other factors, but percolation models are good for comprehending the propagation of events such as sickness.

A combinational method that analyses several variables from various disciplines may create a more accurate and exhaustive model of trust in a digital world. The choice of model or method should be based on the unique research issue and context, with its constraints taken into consideration. Therefore, the correct selection of variables and theoretical frameworks is crucial for the development of reliable trust models. The application of demographic data, speech act theory, and emotion theories can considerably increase the validity of trust assessments. By utilising the capabilities of machine learning algorithms and programming languages like Python, it is possible to construct strong and accurate trust models that can be used for a variety of digital social interactions. Hence, in the section [3.2], we explain the development of the novel social trust model that simulates the conditions of covid-19 events. But, before that a section [3.3] dedicated to the explanation of the dataset creation is given.

3.3 Collection of Corpus

This section outlines the process of collecting a corpus of Hinglish tweets from February to December 2020, a period defined by the COVID-19 pandemic and its impacts on social interactions. The decision to focus on pandemic-era data was made to capture how significant societal disruptions influence trust dynamics in digital interactions. Hinglish, as a common language for Indian social media users, reflects a natural linguistic blend that conveys nuanced social cues, particularly relevant in times of crisis. The data collection was further refined using filters specific to COVID-19 themes, such as keywords related to health, lockdown, community support, and pandemic responses, ensuring that the dataset represents social trust within the pandemic context. Further, using speech act theory, the following instructions were followed to acquire tweet data linked to COVID-19 attitudes and emotions.

We established the exact COVID-19-related emotions and sentiments for which you wish to collect data. This allowed us to design precise search queries in order to discover relevant tweets.

We select a data-gathering platform, such as Twitter's API, as well as certain third-party applications, such as Tweepy and Twint. These tools let us to access Twitter data and obtain tweets in response to certain search queries. The queries were built using the bag of words/dictionary of terms.

Use search phrases tailored to the feelings and sentiments associated with COVID-19. For instance, "COVID-19 AND anxiety", "COVID-19 AND despair", "COVID-19 AND hope", etc. were used to collect the data.

The speech act hypothesis was applied to your search queries. Speech act theory investigates the significance of the words used in spoken and written communication. You can search for tweets with phrases such as "I hope," "I'm scared," "I'm afraid," "I'm excited," etc.

We filtered out irrelevant tweets from our search results. This was accomplished through the use of exclusion terms unrelated to COVID-19 or the precise emotions and sentiments you're interested in.

Export the filtered search results to a CSV file in order to collect the data.

Lastly, we evaluated the data using sentiment analysis methods to determine the overall tone of tweets about COVID-19 as well as the specific emotions expressed.

Here's a summary of the steps in tabular form:

Table 3.5: Steps to Collect Data

Step	Description
1	Determine specific emotions and sentiments related to COVID-19
2	Choose a data collection platform
3	Use specific search queries related to COVID-19 and emotions
4	Apply speech act theory to search queries
5	Filter search results to remove irrelevant tweets

6	Collect data by exporting filtered search results
7	Analyse data using sentiment analysis tools

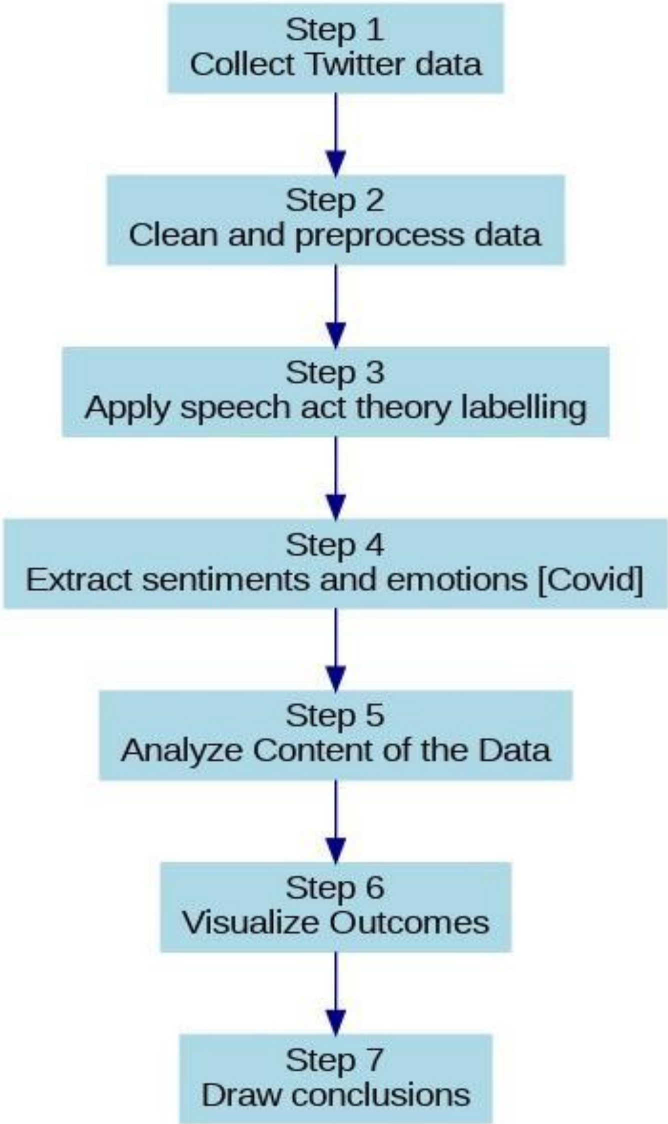

Figure 3.2: Steps to Apply Speech Act Theory

3.4 Data Collection Process

Following steps were followed to obtain the second objective.

1) Definition of the research questions: We defined the research question or problem that you are trying to solve. In our case, the research question is to collect tweets related to COVID-19 sentiments and emotions using the speech

act theory based vocabulary. Hence, a dictionary of search filters was developed for sending search queries to the twitter search engine for downloading tweets. The development of search filters necessitates the compilation of a list of English and Highlish keywords that may be used to mine tweets from either live twitter streams or historical datasets. In the process of building the custom search filters, the following keywords have been found so far and were used for collecting tweets from the existing datasets and live streams of tweets.

Table 3.6: Keywords (Search Filters) for Collecting Tweets

sveekaar	laaparavaah	udaasee	niraasha	jaadoo
prashansa	laaparavaah	prasann	naastikata	gul khilana
aaraadhana	dekhabhaal	hatotsaahit	ayogy karaar diya	prabuddh
sneh	daan puny	udaas	asahajata	virakti
dara hua	mukhar	ichchha	asantosh	utsaah
vyaakulata	utsaah	niraasha	asantusht	daah
yantrana	klosterophobiya	nirdhaarit	ghrna	ahasaas
aakraamak	balapoorvak	sandeh	niraash	utsaah
alaarm	aaraamadaayak	bhay	naapasand	hataash
chintit	vishvaas hai	chalaaya hua	bechainee	utsaah
alagaav kee bhaavana	ulajhan	hakka - bakka rah jaana	gumaraah	ummeed
vismay	ninda	utsukata	maayoos	sammohan
duvidha	saamagree	paramaanand	aprasannata	dar
manoranjan	saahas	utsaah	vyaakulata	paratadaar
gussa	raad	sharmindagee	sankat	dhyaan kendrit
peeda	kroorata	sahaanubhooti	bindh daalee	anuraag
naaraaj ho	jigyaasa		pramukh	udaaseen
aashanka	kutilata	aalasy	pyaar	sunn
chintit	ghabadaaya hua	betaab	havas	jiddee
mitrata	majaboor	udaaseenata	paagal	apamaanit
Bhay	ghar ke baahar rahane se khinn	krodhit	udaasee	udaaseen
hataash hokar	aasha	aasakti	dukhee	aashaavaadee
Rosh	niraashaajanak	vyathit	krpanata	ullanghan

ullaas	bhayaatur	asuraksha	milaaya hua	abhibhoot
udaas	mehamaananavaaz	vyaavahaarik	sheel	bhagadad mach
chamak	niraadar	apamaan	udaaseen	pairaanoyad
prati aabhaar	vinamrata	byaaj	apamaanit	junoon
laalach	chot	intriguaid	chakkar mein pad	dheeraj
Shok	histeeriya	kheeja hua	bura	talleenata
aabhaar	daah karana	prthak	michalee	vikal
aparaadh	prasannata	pasand	nakaaraatmak	zabaradast
khushee	harsh	ghrna	upeksha	niraashaavaad
napharat	aanandotsav	akela	shaktiheen	daya
ghrna	meharabaan	laalasa	hairaan	krodh
sakaaraatmak	maalikaana	deevaana	gaurav	jaldabaaj
pareshaan	isteepha	ghin aana	krtagyata	anishchitata
Khed	bechainee	aatm-dekhabhaal	romaanchit	kam aanka
asveekrt	tabadeelee	aatm dayaalu	thaka hua	bechainee
Dheel	kroor	khud-etamaad	sahanasheelata	aprasann
kaary mukt	udaasee	sankochee	peeda	sakate mein aa
anichchhuk	santushti	aatm mahatvapoorn	vijayee	asthir
aatma glaani	dara hua	aatm ghrna	tang kiya	anishchit
naaraazagee	schhadainfraiudai	aatm prerit	vishvaas	pareshaan
pareshaan	dukh	svany par daya	aashchary	taamasik
Khed	baavajood	svaabhimaanee	duvidha	shaatir
atak gaya	par bal diya	svayan samajh	sandehajanak	satarkata
Vinamr	majaboot	bhaavukata	sahaanubhooti	chapet mein
Peeda	ziddee	shaanti	komalata	kamazor
Maaliny	tanaav	aatank	Chintit	haay

It is commonly believed that people with wide emotional vocabularies are emotionally healthier than those with a more limited vocabulary. According to linguistic principles, the richness or diversity of a person's actively employed emotional vocabulary may connect with his or her typical emotional experiences. It is evident that the dialogue between these persons has the potential to develop into a long-term and trustworthy friendship. However, its dynamics alter throughout time and in response to external conditions. In this study, we aim to determine how the dynamics of the Internet shift in the setting of covid. Inputs from the Emotion Circumplex Model, Plutnik's Wheel of Emotions, and Discrete Emotion Theory were used to create a list of keywords that reflect positive or negative aspects of trust and intimacy. This partial list was developed using the Google transliteration application to create a Hinglish-based keyword store. Over time, this will be expanded even higher. To retrieve the tweets and create a corpus, a formal request for permission of a Twitter developer account was filed. After receiving approval for my application, we developed APIs (as seen in the screenshots below) in our Twitter developer account to retrieve tweets from the whole Twitter ecosystem. So far, the whole technique for retrieving tweets is operating flawlessly.

2) Determination of the data source: For our research work Twitter was the primary source of data.

3) Obtain API access: To collect tweets from Twitter, we obtained API access by creating a Twitter developer account and creating an app.

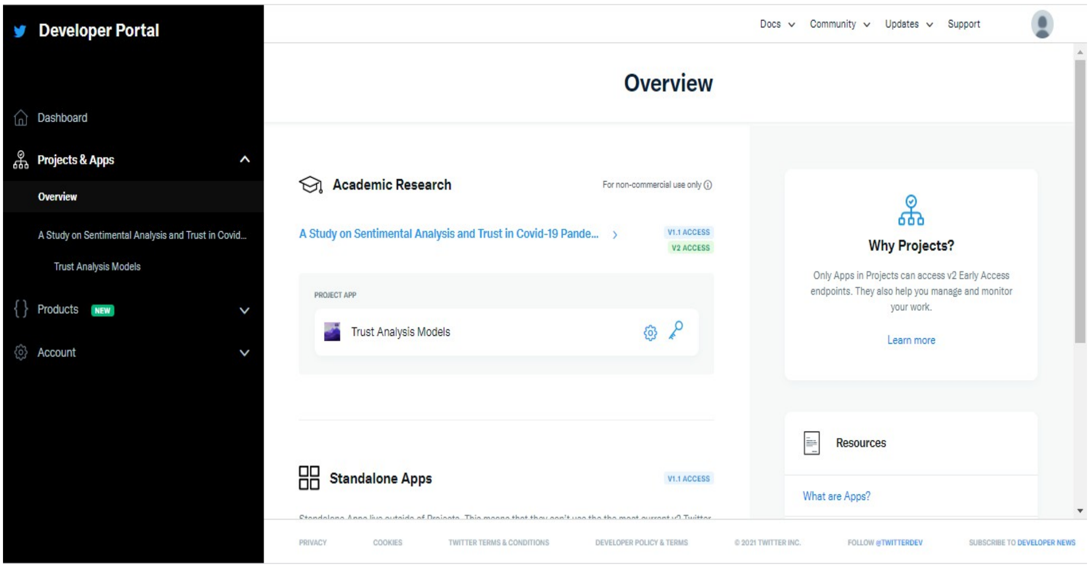

Figure 3.3: Screen Short of Overview Dashboard

4) Installation of Python libraries: To interact with the Twitter API using Python, we installed the necessary libraries. Including Tweepy and Python Twitter Tools (PTT).

5) Authenticate API access: To authenticate API access, we used the credentials from the app you created in step 3.

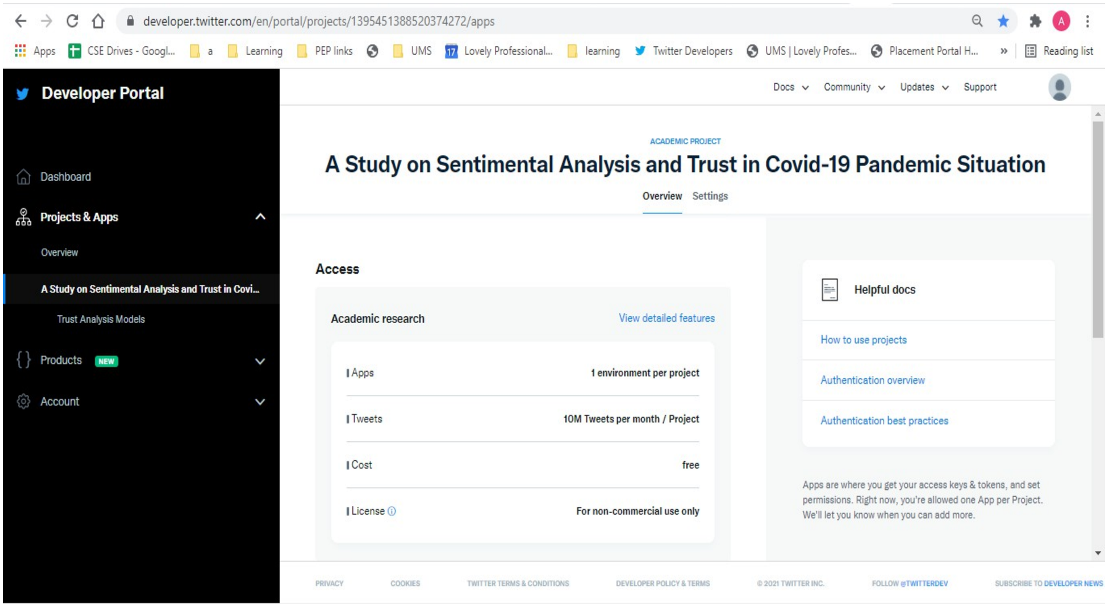

Figure 3.4: Screen Short of Key Access

6) Set search parameters: We fixed and determined the keywords, hashtags, usernames, locations, and other parameters that you will use to search for relevant tweets. For this we constructed a bag of words /Dictionary.

7) Collect tweets: Use the API and search parameters to collect the relevant tweets. You can specify the number of tweets to collect and the time frame in which to collect them.

8) Store tweets: Stored the collected tweets in a file or database for further analysis.

9) Label tweets using speech act theory: To label tweets based on speech act theory, we fixed the type of speech act that each tweet represents (assertive, directive, commissive, expressive). We also used natural language processing (NLP) techniques and libraries such as NLTK to identify speech acts in text.

10) Validate labels: Validated the accuracy of the labelled tweets by checking a sample of them manually. Also adjusted the labelling process as necessary.

11) Analyse data: Analysed the labelled tweet data to identify patterns, trends, and insights related to COVID-19 sentiments and emotions based on speech act theory.

In the next section, we will examine the process of building trust models for the digital society utilising mathematical expressions and data gathered from a variety of sources. As we have seen, trust models are vital to the functioning of digital societies, and building an appropriate trust model involves a comprehensive grasp of the underlying aspects and variables.

3.5 Design of novel model of Social Interaction

Identifying the characteristics that influence trust in a particular digital society is the first step in developing a model of digital society trust. These variables may include demographic variables, cultural differences, psychological variables, economic issues, etc. Once these variables have been discovered, we must collect data pertaining to them from numerous sources, including surveys, online forums, social media platforms, etc. The next phase, following the collection of pertinent data, is to employ mathematical models to analyse the data and determine the correlations between different variables. This may entail the use of statistical methods such as regression analysis, correlation analysis, etc. This investigation seeks to identify and quantify the variables that have the greatest influence on trust in the digital society. Utilizing the outcomes of the data analysis, a mathematical model of trust in the digital society can be constructed. This model could take several shapes based on the nature of the data and the identified links. It could be a basic linear model, a complex nonlinear model, or a hybrid model that integrates many mathematical functions. Next, the model must be validated using various techniques, such as cross-validation, hypothesis testing, etc. This stage is essential for ensuring that the model is accurate and trustworthy and can be used to predict trust in the digital society.

3.5.1 Equation based Digital Social Trust Model (E.D.S.T.M.)

There are numerous ways for constructing models of trust. The equation-based digital social trust model is one such method. Each equation consists of the variables that can represent the components of a relationship that can be modelled mathematically. This section will elaborate on the equation-based model of digital social trust. It will discuss the many model components, the underlying mathematical equations, as well as the approach's benefits and shortcomings. In addition, we will investigate how to build such models using the Python programming language.

3.5.1.1 Construction

As per the objective of this research, we are primarily concerned with analysing "behavioural data," or textual utterances from digital platforms such as Twitter, to gain insights into the emotions and sentiments related to people's 'trust'. Table 3.8 in this section provides information on the possible types of trust models, useful parameters, and assumptions specific to the domain from which the computational model may take inspiration. To construct the mathematical model of trust from real- life data, certain conditions and assumptions need to be considered. Hence, Table 3.7 presents a set of statements that make up the problem. Using these statements, a computational parameter estimation process can be initiated to construct the sociological trust model.

Table 3.7: Observations on Social Trust model

S. No	Observations on Trust Emotion based on Sociological Aspects
1	The value of trust 't' between two parties is directly proportional to the frequency of their interactions.
2	The value of trust between two or more parties increases when they belong to the same community, place of origin, gender group, ethnicity aggregation and interests.
3	The value of trust between two parties is directly proportional to the frequency of their interactions in the physical and digital world.
4	The value of trust between two parties or more is directly proportional to the number of 'positive sentiments' expressed in their digital utterances.

By using the observations mentioned in Table 3.6, the variables (exogenous variables) that impact the overall trust (endogenous variable) of the society can be assessed.

Table 3.8: Exogenous variables that impact the 'Social Trust'.

Parameter	Definition	Data Type	Limits and Ranges
X1 = Age	Age of the digital society member	Integer	18 <= X1<100
X2= Gender	Gender of a person	Enum	Male, Female , Unmentioned
X3= Ethnicity	A state of belonging to a common regional or cultural tradition.	List of ethnic groups	String
X4= places of origins	Place where the person had lived or is living.	List of places of groups	5
X5= hobbies	Interests such as books, sports etc of the person,	List of interests	5
X6= Number of interactions	Number of times a person had conversations with the other party.	Integer	1 <= X6<100
X7 = 'Number of Positive Sentiment Utterances'	From 'X6' number of interactions, how many digital utterances were positive in nature.	Integer	X7 = {1,2,3,.,n}, where n= number of positive interactions

With the help of these seven variables, the following equations can be formulated between each independent variable (X1,X2,X3,X4,X5,X6,X7) and dependent variable (social trust). The X4 and X5 variables have a max range of 5, which implies that five pieces of evidence are required to determine the trust value.

3.5.1.2 Mathematical Construction of E.D.S.T.M.

Let f(t) define the "multiplier effect'' function that increases or decreases when the seven independent variables are changed due to the trigger of some event such as covid 19.

Let 'n' be the total number of members of digital society, each member 'm' having seven x1,x2,x3,x4,x5,x6,x7 computable and observable attributes as explained in table Table 3.9.

Table 3.9: Variable and Equations of Model E.D.S.T.M

Eq. No	Variable	Equation /Condition
1	X1 = Age	*if(xa1 in range(xa1)=true, then ts +1 , where xa=ith party, xb =i th+1 party , ts = social trust value,* ***If both parties are the same age, then the level of trust between them will be greater***
2	X2 = Gender	*f(ti)=if((xa2i xb2(i+1))= (xa2i | xb2(i+1))) then ts+1 xa=ith party, xb =i th+1 party , ts = social trust value ,* ***If both partners are of the same gender, their level of trust will be higher***
3	X3= Ethnicity	*f(ti)=if((xa3i xb3(i+1))= (xa3i | xb3(i+1))) then ts+1 xa=ith party, xb =i th+1 party , ts = social trust value ,* ***If both parties share the same ethnic background, their level of trust will be increased.***
4	X4= places of origins	*f(ti)=if((xa4i xb4(i+1))= (xa4i | xb4(i+1))) then ts+1 xa=ith party, xb =i th+1 party, ts = social trust value ,* ***If both partners were born or raised in the same location, they will have a higher level of trust compared to those. who were born or raised elsewhere***
5	X5= hobbies	*f(ti)=if((xa5i xb5(i+1))= (xa5i | xb5(i+1))) then ts+1 xa=ith party, xb =i th+1 party , ts = social trust value ,* ***If both people connect with one another and share similar interests, such as reading comparable types of books, their level of trust will increase.***
6	X6= Number of interactions	*f(ti)=if(unique(xia6, xb6(i+1))>Threshold, then ts+1 xa=ith party, xb =i th+1 party , ts = social trust value ,* ***If both parties interact on a digital network such as Twitter, they have incentive to trust one another.***
7	X7 = 'Number of Positive Sentiment	*f(ti)=if(unique(xia7, xb7(i+1))>PositiveThreshold, then ts+1 xa=ith party, xb =i th+1 party, ts = social trust value.*

	Utterances'	***As both sides connect on a digital network such as Twitter, they will have a motivation to trust one another, and their trustworthiness will rise if they exchange good speech acts.***

Using these variables and factors, we can construct the non-linear equation model as follows:

Trust $= \alpha + \beta 1 {*} X1 + \beta 2 {*} X2 + \beta 3 {*} X3 + \beta 4 {*} X4 + \beta 5 {*} X5 + \beta 6 {*} X6 + \beta 7 {*} X7 + \varepsilon$

Where,

- Trust is the dependent variable, representing digital trust in the digital society
- α is the intercept term
- $\beta 1$ to $\beta 7$ are the coefficients of the independent variables.
- X1 to X7 are the independent variables as defined above
- ε is the error term

The coefficients $\beta 1$ to $\beta 7$ represent the impact of each independent variable on the digital trust in the digital society. The model can be further analyzed to identify the most significant variables and factors contributing to digital trust.

3.5.1.2.1 Rationale for Variable Selection in E.D.S.T.M

The selection of variables in the Equation-based Digital Social Trust Model (E.D.S.T.M.) was informed by a combination of sociological insights and computational relevance. The model incorporates seven key variables representing factors that influence digital trust in online interactions. These include parameters such as communication frequency, response time, sentiment polarity, topic relevance, user engagement level, past interaction count, and lexical tone.

Each variable reflects a dimension of trust theory commonly referenced in sociology (e.g., trust-building through consistent interaction and emotional tone) and is computationally measurable through NLP and data mining techniques. For instance, sentiment polarity is widely accepted as a proxy for emotional alignment, while response time signifies user attentiveness and perceived reliability. The inclusion of these specific variables ensures the model captures both quantitative behavior patterns and qualitative trust signals that are central to digital social interactions. The combination was validated through iterative testing and performance evaluation on simulated and real Twitter datasets.

3.5.1.3 Rules of Social Interactions

Equation No: 1

Variable: X1 = Age

Equation/Condition: *if (xa1 in range (xa1)=true, then ts +1, where xa=ith party, xb =i th+1 party, ts = social trust value.*

If both parties are of the same age, their level of trust will be higher.

Equation No: 2

Variable: X2 = Gender
Equation/Condition: *f(ti)=if((xa2i xb2(i+1))= (xa2i | xb2(i+1))) then ts+1, where xa=ith party, xb =i th+1 party, ts = social trust value.*

If both partners are of the same gender, their level of trust will be higher.

Equation No: 3

Variable: X3= Ethnicity

Equation/Condition: *f(ti)=if((xa3i xb3(i+1))= (xa3i | xb3(i+1))) then ts+1, where xa=ith party, xb =i th+1 party, ts = social trust value.*

If both parties share the same ethnic background, their level of trust will be increased.

Equation No: 4

Variable: X4= places of origins

Equation/Condition: *f(ti)=if((xa4i xb4(i+1))= (xa4i | xb4(i+1))) then ts+1, where xa=ith party, xb =i th+1 party, ts = social trust value.*

If both partners were born or raised in the same location, they will have a higher level of trust compared to those who were born or raised elsewhere.

Equation No: 5

Variable: X5= hobbies

Equation/Condition: *f(ti)=if((xa5i xb5(i+1))= (xa5i | xb5(i+1))) then ts+1, where xa=ith party, xb =i th+1 party, ts = social trust value.*

If both people connect with one another and share similar interests, such as reading comparable types of books, their level of trust will increase.

Equation No: 6

Variable: X6= Number of interactions

Equation/Condition: *f(ti)=if(unique(xia6, xb6(i+1))>Threshold, then ts+1, where xa=ith party, xb =i th+1 party, ts = social trust value.*

If both parties interact on a digital network such as Twitter, they have an incentive to trust one another.

Equation No: 7

Variable: X7 = 'Number of Positive Sentiment Utterances'

Equation/Condition: *f(ti)=if(unique(xia7, xb7(i+1))>PositiveThreshold, then ts+1, where xa=ith party, xb =i th+1 party, ts = social trust value.*

As both sides connect on a digital network such as Twitter, they will have a motivation to trust one another, and their trustworthiness will rise if they exchange good speech acts and written utterances.

3.5.1.4 Updating the Digital Social Trust Model (E.D.S.T.M)

Using the rules given in the previous answer, we can construct a non-linear equation model for digital trust in digital societies such as Twitter. Let us denote the social trust value between two parties i and i+1 as ts and let the values of the variables X1 to X7 be denoted by xa1 to xa7 for party i and xb1 to xb7 for party i+1.

The non-linear equation model for digital trust can be expressed as:

ts = f(ti) = (xa1 in range(xb1)) + (xa1 == xb1) + ((xa2i != xb2(i+1)) and (xa2i != 'Threshold') and (xb2(i+1) != 'Threshold')) + ((xa3i != xb3(i+1)) and (xa3i != '') and (xb3(i+1) != '')) + ((xa4i != xb4(i+1)) and (xa4i != '') and (xb4(i+1) != '')) + (len(set(xa5i) & set(xb5(i+1))) > 0) + (len(set(xia6) & set(xb6(i+1))) > Threshold) + (len(set(xia7) & set(xb7(i+1))) > PositiveThreshold)

Here, the function f(ti) calculates the social trust value between two parties i and i+1 based on the rules mentioned earlier. The range function is used to check if the age of the two parties is within a certain range. The == operator is used to check if the age of the two parties is the same. The!= operator is used to check if the gender, ethnicity, place of origin, and hobbies of the two parties are not the same. The & operator is used to find the common interests and interactions between the two parties. Finally, the len() function is used to check if the number of common positive sentiment utterances is greater than the positive threshold. It must be noted that the Thresholds and Positive Thresholds are predefined values that can be adjusted based on the data and the specific needs of the model.

Computations:

3.5.1.5 Construction of Algorithms Using E.D.S.T.M

1. Initialise social trust value (ts) for each member in digital society to 0.

   For each member (m) in digital society:

   a. Compute ts for m using the following equations:

   i. If Equation 1 is true for m, *ts = ts + 1.*

   ii. If Equation 2 is true for m, *ts = ts + 1.*

   iii. If Equation 3 is true for m, *ts = ts + 1.*

   iv. If Equation 4 is true for m, *ts = ts + 1.*

   v. If Equation 5 is true for m, *ts = ts + 1.*

vi. If Equation 6 is true for m, *ts = ts + 1.*

2. If the sentiment analysis for m is greater than the Positive Sentiment_Threshold, *ts = ts + 1.*

3. Return ts for each member.

This work presents foundational notions for developing mathematical reasoning for the purpose of constructing realistic algorithms of 'social trust.' These seven equations and associated operations derive their inspiration from theories presented in recent literature in the fields of sociology, economics, psychology, and other areas. An illustration of how the social trust model/algorithm can be constructed has been provided, which would help aspiring researchers in this field develop higher order mathematical thinking skills. It must be noted that in terms of the purpose and scope of this (E.D.S.T.M.) model it is designed to quantify digital trust between individuals within a specific online community or network, such as Twitter. It achieves this by evaluating straightforward, observable characteristics (age, gender, frequency of interaction, etc.) that influence trust. The model provides a linear relationship between these characteristics and trust, which simplifies the computational complexity and makes it suitable for smaller, well-defined networks.

So basically, the E.D.S.T.M. has a limited scope to situations where trust interactions are relatively straightforward and where demographic and interactional data can effectively represent trust dynamics. It assumes linear interactions between variables, which may limit its accuracy in complex social environments where interactions are influenced by multiple, interdependent factors.

Due to its reliance on a linear equation, E.D.S.T.M. may not fully capture the nuanced or non-linear nature of trust dynamics in larger or more diverse social systems. Above all, it does not account for broader sociological or psychological factors that can impact trust, such as cultural influences, emotional variability, or the influence of major events. It fails to model how trust evolves over time in response to external shocks or societal changes. Social trust dynamics are often non-linear and influenced by complex interactions among multiple factors. For example, a small increase in negative sentiment during a pandemic can disproportionately erode trust. E.D.S.T.M. does not account for the dynamic nature of trust, especially during significant events like the COVID-19 pandemic. Hence, it is obvious to further enhance the model with more variables that cover more complex relationships for computing the trust.

Following is the scope and approach of Advanced Digital Society Interaction and Trust Model (A.D.S.I.T.M):

The ADSITM is technically a more comprehensive model that addresses the limitations of E.D.S.T.M. by incorporating a broader set of variables and capturing complex, non-linear relationships. Its purpose is to simulate digital trust within a large, heterogeneous digital society by considering additional sociological, psychological, and demographic factors. This model introduces non-linear dynamics, accounting for how trust evolves under the influence of external events (such as COVID-19), as well as group behaviour, cultural factors, and emotional intensity. In terms of scope, ADSITM is intended for larger, more diverse digital networks where trust interactions are influenced by a variety of factors beyond simple demographic and interactional characteristics. It is particularly useful for analysing digital trust in contexts where emotions, group identities, and external events play a significant role.

- **How It Overcomes Limitations**: A.D.S.I.T.M. overcomes the limitations of the simpler E.D.S.T.M. by:
    a. Introducing non-linear relationships to model complex interactions and feedback loops in trust dynamics.
    b. Including variables related to social identity, collective behaviour, and emotional states, which allow it to capture the diversity of interactions within a digital society.
    c. Employing a clustering approach that groups individuals based on similarities in trust and interaction profiles, providing a more nuanced view of trust dynamics.
    d. Incorporating adaptability to external events, enabling it to model changes in trust across the population in response to significant events.

In the next section, we discuss the steps taken for the construction of the Advance model.

3.5.2 Advance Digital Society Interaction and Trust Model (A.D.S.I.T.M):

While the E.D.S.T.M. offers a foundational approach to modeling digital social trust using linear relationships among basic demographic and interaction variables, it falls short in capturing the multifaceted and dynamic nature of trust within complex digital societies, especially during unprecedented events like the COVID-19 pandemic. Trust is not solely influenced by observable interactions but is deeply rooted in psychological, emotional, and cognitive processes that interact in non-linear ways. Hence, in this part, we explain in detail how the Digital Society Interaction and Trust Model (D.S.I.T.M.), a hybrid method to understanding social trust in digital societies, was constructed. To develop a comprehensive algorithm for comprehending social interactions and trust in digital societies, the model takes into account a variety of psychological, sociological, and demographic characteristics. The A.D.S.I.T.M presents a more realistic model of behaviour in digital societies, allowing for a more accurate explanation of how social trust develops in online environments, by incorporating factors from other domains. Altogether, the A.D.S.I.T.M represents a significant achievement in the field of digital sociology and serves as a great resource for scholars wishing to investigate the complex dynamics of social trust in digital societies. Following are the assumptions and logical statements that make up the A. D.S.I.T.M.

1. Digital society is an 'institution' that operates on the internet with dynamic rules 'r' that guides the thinking and behaviour of individual agents.

2. Dynamic rules/laws are a cognitive process through which individuals assimilate how to interpret and respond to the world around them.

3. The population 'p' has its own collective cognitive apparatus, and the rules/laws at the individual level may not be directly observable but can be inferred through data mining analysis.

4. The collective cognitive apparatus of the population 'p' forms perceptions ('pc') with anchored sentiments/emotions such as trust ('tr') that influence the actions of individuals.

5. Perception is the application of the assimilated rules/laws in the mental mind

and can be assessed based on the content/topics/subjects an individual responds to on the internet.

6. At a given time 't', it is assumed that the digital society maintains an equilibrium of observable laws/rules that result in a neutral perception ('pc') and trust emotion ('tr = neutral'), and the cognitive structure of the digital society is stable. The individual's perceptual apparatus is tacit knowledge that guides thoughts and actions and impacts emotions, such as trust.

7. Emotions are the reasons why rules/laws are followed and acted upon, and they can be linguistically expressed and empirically testable.

8. Emotions can have different intensities or activations, such as low, medium, or high trust, and they provide a sense of urgency or motivation to act in response to the environment.

9. Emotion analysis is different from sentiment analysis, as emotions are complex raw states while sentiment is an organised form of emotions expressed as positive, negative, or neutral.

10. The focus of this research is to study the organised state of emotions and perceptions in a digital society during events like COVID-19, using linguistic metal maps of individuals and sentiment dictionaries.

11. The equilibrium of the digital society changes as the dynamic rules/laws change due to the occurrence of events ('ev') of types Mediocristan and Extremistan, which bring structural and systematic changes to society.

12. Mediocristan has a thin-tail graph and affects individuals independently, while Extremistan affects a large number of individuals and has systemic consequences.

13. To understand the intensity of a particular sentiment, such as trust ('tr'), in a population ('p'), a consequentialist reasoning can be constructed using linguistic metal maps of sentiments represented by a bag of words ('sdic') and distance metrics.

14. The bag of words represents a specific sentiment dictionary ('sdic') and can be

used to assess the state of sentiment in the digital society in relation to a specific linguistic metal map of current sentiment.

The statements describe the assumptions, axioms, and constructs related to the digital society as an institution operating on the internet. They discuss the dynamic rules/laws

that guide the thinking and behaviour of individual agents, the collective cognitive apparatus of the population, the role of perceptions and emotions in shaping behaviour, and the impact of events on the equilibrium of the digital society. The statements also highlight the difference between emotions and sentiment, and propose the use of linguistic metal maps and sentiment dictionaries for analysing sentiments in the digital society.

Table 3.10: Variables used in Model II

Variable Name	Description	Possible Ranges/Conditions
nx	Number of members in the digital society (institution)	Positive integer
r	Set of rules for segregation algorithm	List of rules or conditions
Emotions (e)	Different levels of emotions impacted by events such as covid-19	Categorical variable (e.g. low, medium, high)
Trust (t)	Different levels of trust impacted by events such as covid-19	Categorical variable (e.g. low, medium, high)
Metal Map 'm'	Cognitive apparatus impacted by emotions, trust, and other factors	Complex data structure (e.g. matrix, graph)
Covid-19 Impact	Impact of covid-19 event on members of the digital society	Categorical variable (e.g. low, medium, high)
Clusters	Segregated groups of members based on rules, emotions, trust, and other factors	List of clusters or groups
Cluster Centers	Central points or representatives of each cluster	List of cluster center points
Distance Metric	Measure of similarity or dissimilarity between members in the digital society	Function or algorithm for calculating distance
Initialization Method	Method for initializing cluster centers	Function or algorithm for initializing cluster centers

Convergence Criteria	Criteria for determining convergence of the algorithm	Threshold or stopping condition
Cluster Assignment	Assignment of members to clusters based on distance metric and cluster centers	Function or algorithm for assigning members to clusters
Update Cluster	Update of cluster centers based on	Function or algorithm for
Centers	assigned members	updating cluster centers
Repeat Steps	Number of times to repeat the cluster assignment and update steps	Positive integer
Final Clusters	Segregated groups of members after convergence	List of final clusters or groups

This table outlines the various variables and conditions used in a digital society analysis, particularly in relation to the impact of events such as Covid-19 on trust and emotions within the society. The first variable is "nx", which represents the number of members in the digital society or institution. The second variable is "r", which is a set of rules or conditions used in the segregation algorithm for segregating members into different clusters based on factors such as emotions and trust levels. The third and fourth variables are "e" and "t", which represent different levels of emotions and trust respectively, and are impacted by events such as Covid-19. These variables are categorical and may range from low to high. The "m" variable is a complex data structure, such as a matrix or graph, representing the cognitive apparatus impacted by emotions, trust, and other factors within the society. The "Covid-19 Impact" variable represents the impact of the Covid-19 event on the members of the digital society and is also a categorical variable. The remaining variables, including clusters, cluster centers, distance metric, initialization method, convergence criteria, cluster assignment, update cluster centers, and final clusters, all relate to the clustering algorithm used to group members based on various factors. These variables involve functions or algorithms that assign members to clusters, update the clusters, and determine convergence criteria. Overall, this table provides a comprehensive overview of the various factors and variables involved in analysing the impact of Covid-19 on a digital society.

Algorithm : Advance Digital Society Interaction and Trust Model (A.D.S.I.T.M)

Step 1: Define the members of the digital society (Twitter) members =
['member1', 'member2', 'member3', ...] # List of members

# **Step 2**: Define the emotions and trust levels for each member
emotions = {'member1': 'emotion1', 'member2': 'emotion2', 'member3': 'emotion3', ...} #
Dictionary of emotions trust_levels = {'member1': 'trust1', 'member2': 'trust2', 'member3': 'trust3', ...} #
Dictionary of trust levels

# **Step 3:** Define the impact of the event (COVID-19) for each member
event_impact = {'member1': 'impact1', 'member2': 'impact2', 'member3': 'impact3', ...} #
Dictionary of event impacts

# **Step 4:** Define the rules (r) for segregation
rules = {'rule1': 'value1', 'rule2': 'value2', 'rule3': 'value3', ...} # Dictionary of rules

Step 5: Define a function to compute similarity between members based on emotions, trust levels, and event impact

```
def compute_similarity(member1, member2):
 similarity = 0
 if emotions[member1] == emotions[member2]:
 similarity += rules['emotion_similarity']
 if trust_levels[member1] == trust_levels[member2]:
 similarity += rules['trust_similarity']
 if event_impact[member1] == event_impact[member2]:
 similarity += rules['event_impact_similarity'] return
 similarity
```

Step 6: Initialize a dictionary to store the similarity scores for each member pair

similarity_scores = {}

 # Step 7: Compute similarity scores for each member pair

```
for i in range(len(members)):
 for j in range(i + 1, len(members)):
 similarity_scores[(members[i], members[j])] = compute_similarity(members[i],
members[j])
```

 # Step 8: Sort the similarity scores in descending order

```
sorted_similarity_scores = sorted(similarity_scores.items(), key=lambda x: x[1],
reverse=True)
```

 # Step 9: Initialise a dictionary to store the clusters

```
clusters = {member: [member] for member in members}
```

 # Step 10: Merge clusters based on similarity scores

```
for pair, similarity in sorted_similarity_scores:
member1, member2 = pair
 if similarity >= rules['similarity_threshold']:
 for member in clusters[member1] + clusters[member2]:
 clusters[member1].extend(clusters[member2])
 clusters[member2].extend(clusters[member1])
 clusters.pop(member2, None)
```

 # Step 11: Initialize a dictionary to store the segregated clusters

```
segregated_clusters = {}
```

 # Step 12: Assign each member to a segregated cluster

```
for member in members:
 for cluster in segregated_clusters.values():
 if member in cluster:
```

```
# Step 13: Merge clusters based on
rules for rule, value in rules.items():
 if rule.startswith('merge_'):
 clusters_to_merge =
 value.split(',') for cluster in
 clusters_to_merge:
 cluster_members =
 cluster.split(':')
 main_cluster =
 segregated_clusters[cluster_members[0]] for member
 in cluster_members[1:]:
 main_cluster.extend(segregated_clusters[member])
 segregated_clusters.pop(member, None
```

3.5.2.1 Explanation of the A.D.S.I.T.M. algorithm

1. Initialize the digital society with 'nx' number of members.

2. Assign each member a level of emotions ('e') and trust ('t') based on their individual characteristics, such as their behaviour, interactions, and responses to events like Covid-19.

3. Define the set of rules ('r') that will be used to segregate the members of the digital society based on their emotions and trust levels.

4. Create a metal map that represents the cognitive apparatus of each member, taking into account their emotions, trust levels, and other relevant factors.

5. Apply the segregation algorithm, using the set of rules ('r') to sort and group the members based on their metal maps.

6. Start the segregation process by identifying members who exhibit extreme emotions ('e') in response to the event, such as high anxiety, fear, or anger.

7. Separate members with extreme emotions into clusters based on their trust levels ('t'). For example, those with high trust may form one cluster, while those with low trust may form another.

8. Within each trust level cluster, further segregate members based on other factors, such as their cognitive apparatus, which includes their beliefs, thoughts, and perceptions.

9. Identify members who have moderate emotions and trust levels, and segregate them into appropriate clusters based on similar factors as in step 8.

10. Repeat the process for members who exhibit low emotions ('e') in response to the event, but have different trust levels ('t') and other factors affecting their cognitive apparatus.

11. Continue segregating members based on emotions, trust, and other

relevant factors until all members are grouped into distinct clusters.

12. Evaluate the effectiveness of the segregation algorithm by analyzing the resulting clusters and their characteristics, such as the distribution of emotions, trust levels, and cognitive apparatus within each cluster.

13. Refine the algorithm if needed, based on the evaluation results, to improve the accuracy and reliability of the segregation process.

14. Monitor the dynamics of the digital society and update the segregation algorithm periodically to account for changes in emotions, trust levels, and other relevant factors.

15. Use the segregated clusters of members to design targeted interventions or policies aimed at addressing the specific needs and challenges faced by each group in the digital society, considering the impact of the event (Covid-19) and other relevant factors.

3.5.2.2 Visualisation of the steps for the construction of D.S.I.T.M. algorithm

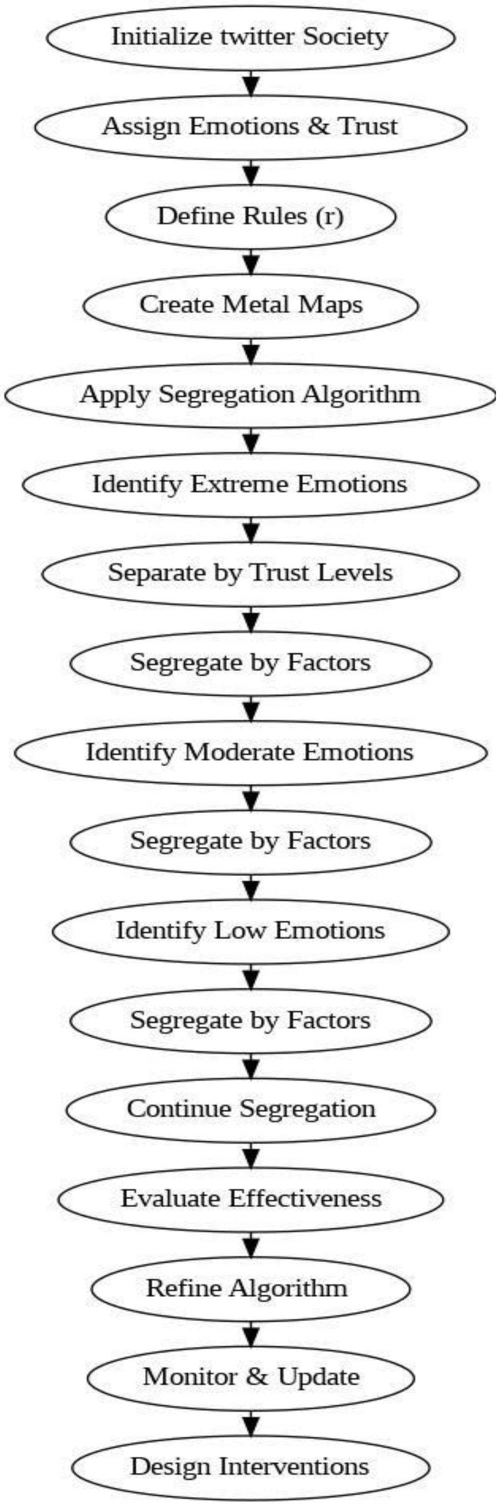

Figure 3.5: Block diagram of Constructing Model (Digital Society Interaction and Trust Model (D.S.I.T.M)

Scenarios and Cases Addressed by A.D.S.I.T.M.

The development of A.D.S.I.T.M. is essential to overcome the limitations inherent in E.D.S.T.M. By embracing complexity and incorporating a broader range of influential factors, A.D.S.I.T.M. provides a more accurate and relevant model of digital social trust. This enhanced model is particularly valuable in understanding trust dynamics during critical periods like the COVID-19 pandemic, where traditional models may fail to capture the nuanced shifts in trust influenced by psychological and emotional factors. Here are the scenarios and cases addressed by it.

- **Emotional Variability Impact:**
    - E.D.S.T.M. Limitation: Does not account for emotional intensity variations.
    - A.D.S.I.T.M. Solution: Models different emotion activation levels (low, medium, high) and their effect on trust.
- **Cognitive Perception Influence:**
    - E.D.S.T.M. Limitation: Lacks consideration of individual and collective cognitive processes.
    - A.D.S.I.T.M. Solution: Incorporates mental models (m) and perceptions (pc) to understand how individuals interpret and respond to information, affecting trust.
- **Response to External Events:**
    - E.D.S.T.M. Limitation: Static and cannot adapt to significant external changes.
    - A.D.S.I.T.M. Solution: Integrates the impact of events (ev), allowing the model to simulate trust dynamics during crises like the COVID-19 pandemic.
- **More Demographic Data Integration:** The advantage is by including more nuanced demographic variables, A.D.S.I.T.M. We can identify trust patterns across different community segments. It also enhances the model's ability to predict trust variations among diverse groups during the pandemic.
- **Emotional State Consideration:** One more advantage is that it recognizes that emotions like fear, anxiety, or hope significantly influence trust levels.

The outcome is that it provides a more accurate depiction of trust dynamics as people react emotionally to pandemic developments.

- **Cognitive and Perception Factors:** The benefit is that the model how misinformation or information overloads affects individuals' trust. The result is that it helps in understanding the spread of distrust or trust restoration efforts in digital societies.

In nutshell, the following is the end purpose and scope should be expected from the construction of both these models.

E.D.S.T.M.: Best suited for simpler or smaller-scale trust modelling tasks, where linear relationships between demographic data and trust interactions are sufficient.

A.D.S.I.T.M.: An advanced model designed to handle the complexity of digital trust in large, diverse social environments, addressing E.D.S.T.M.'s limitations by accounting for non-linear, emotional, and sociocultural factors.

The Equation-Based Digital Social Trust Model (E.D.S.T.M.) and the Advanced Digital Society Interaction and Trust Model (A.D.S.I.T.M.) were developed and tested using a dataset specific to the COVID-19 pandemic. Both models are tailored to examine trust-related behaviours in response to the social and emotional impacts of the pandemic. The E.D.S.T.M. provides a foundational approach, utilising linear interactions among key demographic and interaction variables to model trust during this period of heightened digital engagement. However, recognizing the complexity and evolving nature of trust in the context of a public health crisis, the A.D.S.I.T.M. extends this approach by incorporating additional sociocultural and emotional factors.

While these models may be applicable in other contexts, their design and calibration in this study are based on the unique social dynamics of the pandemic. The insights and parameters derived from pandemic-specific interactions are essential to understanding how trust manifests and shifts under collective stress and uncertainty, offering a focused view of digital trust within a major global event.

3.6 Content Analysis of the Crawled Corpus:

Most linguists scientifically assert that a public expression addressed to a large number of people has a "performative function" in society. This implies that these utterances play some role in changing the dynamics of society. For example, a speech by Martin Luther produced a dramatic change in American society. To cut a long story short, a series of tweets has caused dramatic changes in the sentiments and fabric of society. Hence, in this research, the work is carried out based on the assertion that "tweets have a "performative function" in society. The function is to bring about change.

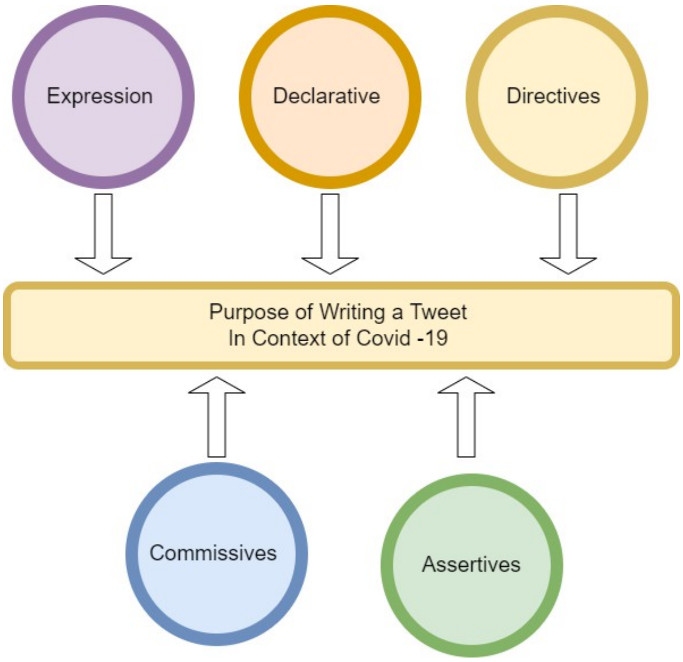

Figure 3.6: The Purpose of Writing a Tweet

To determine the mood expressed in tweets, a process known as' discourse integration' is required. Finally, pragmatic analysis must be used to identify the desired outcome through the use of a set of standards that define cooperative dialogues/tweets/ utterances. Each of these activities lays the groundwork for content analysis. Recently, the speech act theory has been viewed as the most established branch of pragmatics. However, it can also be observed this theory has been disregarded when attempting to comprehend statements in recent literature also. In the present context, it must be understood that semantics is concerned with the literal meaning of words and their relationships, whereas pragmatics is concerned with the inferred meaning perceived by readers and listeners. The study of meaning is known as semantics, or more accurately, the study of the link between language phrases and their meanings. Pragmatics is the study of context, or perhaps more particularly, how context affects our comprehension of linguistic utterances. However, in this research, an attempt will be made to apply some of the principles that were introduced by J.L. Austin and Sarte [39] to understand the nature of written expressions (tweets). According to the speech act theory [39], speech or utterances of words (tweets) are more than just a means of conveying information; they are also a call to some sort of action or act. Figure 3.6 gives a framework of speech act theory applied to written digital utterances (tweets). Hence, with the help of topic modeling-based algorithms and the speech act framework, the following questions will be addressed.

Hence, in this section, an explanation of the methods for collecting social media data and its pre-processing steps for making it suitable for meaningful analysis is provided. The section begins with the explanation of the process of acquiring datasets for the said content analysis and thereafter explains the process of cleaning the datasets and making them useful for sentiment machine learning models, data mining operations, and deep learning models. The twarc api, twitter api, and tweepy libraries were used to access live tweet streams and download existing tweets with help of the Hydration [40] process . For the "content" analysis of COVID and TRUST tweets, the Spycap python package was used, and the structure of each tweet is shown using python visualisation libraries.

Table 3.11: COVID tweet datasets used for Content Analysis (C.A)

S. No	Dataset	Short Description
1	Tweets Originating from India During COVID-19 Lockdowns [Ref] https://doi.org/10.21227/k8gw-xz18	This database includes tweets sent from India during the first week of each of the four different phases of the Indian administration's state wide lockdowns.
2	COVID-19 Twitter Dataset https://doi.org/10.5683/SP2/PXF2CU	The real-time Twitter feed is scanned for coronavirus related tweets using more than 90 different keywords and hashtags associated with the pandemic.
3	Coronavirus (COVID-19) Tweets Sentiment Trend (Global) https://doi.org/10.21227/t263-8x74	This dataset provides a snapshot of the overall sentiment trend in the public dialogue on Twitter surrounding the COVID-19 outbreak.
4	Data from: GeoCoV19: A Dataset of Hundreds of Millions of Multilingual COVID-19 Tweets with Location Information. https://doi.org/10.21227/fpsb-jz61	This is a sizable Twitter dataset encompassing over 524 million multilingual tweets published during a 90-day period beginning February 1, 2020.
5	Place#Hashtag Twitter Dataset: COVID-19 Hashtags. https://doi.org/10.21227/aga4-fz72	The authors of this project analysed 300,000 [country, hashtag] relationships from 190 countries and territories, as well as 5055 distinct hashtags.
6	COVID-19 second wave tweets (with annotations). https://doi.org/10.21227/57fn-3905	This study examines the thoughts and feelings of those who used Twitter in the early days of the second wave.

All six of the abovementioned datasets were processed with the help of the "**Hydration**" process for retrieving the actual tweet data, and each tweet passed through a pre-processing routine [40] . The pre-processing routine cleans up each tweet to strip off unwanted parts such as special characters, links, emoji, etc. The next section explains the process of content analysis of these tweets, referred to as "guided content analysis, (G.C.A)" as the analysis process involves 'speech act theory' as the underlying logic for analysis.

3.6.1 Guided Content Analysis (G.C.A):

To answer both the research questions mentioned in the last section, extraction and analysis of the terms found by the three **NMFF, PLSI, LDA** algorithms were done. Terms such as politics, bureaucrats, service, police, defence, army, news reporters, medical professionals, doctors, film actors, and ethnic and cultural identities such as 'Punjabi' were identified. Famous influences in the context of India, Lockdown, and COVID were also identified in due course of the analysis. Furthermore, a fair degree of terminology, items, and vocabulary associated with COVID, and lockdown were collected. Dependency analysis of each word found by these algorithms was done so that the structure and form of the different types of tweets could be checked if they fit into the 'speech act theory' framework. The inferences and the tables [1] were drawn based on the random sampling of tweets drawn from the pool of six datasets mentioned in Table 3.11.

Further, each sample was categorized into one of the five main groups as per the speech act theory, and further, each sample was given a label as per the 'act' or 'intention' of the tweet content.

A random sampling of tweets was done for the reason of content analysis (C.A.). The reason for choosing random sampling over full sample analysis was a lack of resources. Secondly, a large number of accounts are proxy, bot accounts where the real human person is not there to interact like in real life. All these steps are referred to as "guided content analysis" because items found from these three algorithms were further used to identify tweets that may fit into one of the five types of "tweet acts" or "writing acts"' or simply "speech acts." The working of the three (**NMFF, PLSI and**

LDA) algorithms is explained in section 3.6.2. The selection of the three algorithms is based on their popularity and effectiveness in this field, as found from contemporary surveys in this context.

3.6.2 The Topic Content Analysis Algorithm (T.C.A.A.)

The problem is figuring out how to understand the hidden topics and insights within a large corpus of tweets. There is a need to cluster the data in a way that small clusters of concepts or topics emerge and Inferences can be drawn from them for further content analysis. Hence, two categories of algorithms were considered for this purpose.

3.6.2.1 Group 1: NMF algorithms

Each original document is created from a tiny set of hidden features, according to NMF algorithms. NMF creates them. Non-negative matrix factorization (NMF or NNMF) is a collection of techniques in multivariate analysis and linear algebra that factor a matrix V into (typically) two matrices W and H with no negative elements.

Steps

1) Tweets are an input matrix. Tweet Input Matrix has X rows and Y columns. Documents are in columns, not rows. X words index Y documents. InTweet Input Matrix, 'vc' denotes a document. Each original document is presumed to have hidden characteristics. NMF creates them. X features are extracted to create an X-by-Y feature matrix W.

2) "CM" matrix has X rows and Y columns. W * H = a matrix, same shapethe as Tweet Input Matrix. Check if the factorization is a good approximation of the Tweet_Input Matrix.

3) From the approach of matrix multiplication above, each column in the product matrix W*H is a linear combination of X column vectors in the features matrix W with coefficients from the coefficient matrix CM.

4) NMFF clusters input columns in this step. Earlier steps cause this. Error function minimization approximates by. By constraining to orthogonality, the aforementioned minimising becomes K-means clustering minimization. Kullback–Leibler divergence as an optimization function equals probabilistic latent semantic analysis.

5) Computed yields cluster membership, i.e., if all I k, the input data belongs to the cluster. The computed yields cluster centroids, i.e., the-th column delivers the-th cluster's centroid. Convex NMF improves this centroid's representation. Orthogonality and clustering hold when the orthogonality constraint is not explicitly imposed.

3.6.3 Latent Dirichlet Allocation Algorithm

1) Set in subjects for LDA to find. How many themes are ideal? We attempt different n values till we're satisfied with the outcomes. Or, if we're lucky, we can use additional dataset information to determine the ideal number of topics.

2) Every document's words should be categorised. This temporary topic will be random at first, then updated.

3) In this step the algorithm goes through every document (in our case tweet posts) and then every word in that document to calculate two values: the probability that this document belongs to a certain topic based on how many words (except the current word) from this document belong to the topic of the current word and the proportion of documents assigned to the topic of the current word because of the current word.

4) In the final stage, the algorithm repeats step 3 several times (established before beginning to run the algorithm). We'll look at each document, discover its most prevalent topic, and assign it to that topic.

In this section, we have carried out an ***exhaustive Content Analysis*** of the tweet data by making use of a variety of algorithms and methods in order to process and examine the data. We have outlined and provided a detailed description of each phase, including the particular algorithms that are utilised at each level of the process.

In the next chapter under "Results," we will show the outcomes and discoveries that were gained from these steps. We will also provide insights and interpretations based on the data that was analysed. The findings give light on the patterns, trends, and correlations that were uncovered during the research. This contributes to a greater understanding of the data and the implications it carries. In addition, we give visualisations and graphical representations to support a clear and concise presentation of the data, which enhances the comprehensibility and interpretability of the findings. This was done so as to facilitate a clear and succinct presentation of the results. In general, the content analysis and subsequent procedures that were carried out in this research contribute to the progress of knowledge in the field of digital society interactions and trust modelling by providing significant insights and providing valuable information. The information and findings that are presented in this section serve as a foundation for the ensuing debates and conclusions that are reached in the results chapter. As a result, the research that was carried out receives a major boost in both value and relevance.

3.6.4 Content Analysis Algorithms: Rationale, Strengths, and Limitations

In this study, three algorithms were chosen for topic content analysis: Non-negative Matrix Factorization (NMF), Probabilistic Latent Semantic Indexing (PLSI), and Latent Dirichlet Allocation (LDA). Each of these algorithms offers unique strengths for analysing topic structures within the pandemic-specific corpus, providing a comprehensive view of how different social trust themes emerge across digital conversations. The selection of these algorithms was guided by their ability to handle high-dimensional text data and their complementary approaches to topic discovery.

1. **Non-negative Matrix Factorization (NMF)**
    - **Rationale**: NMF is a matrix decomposition technique that is particularly effective for identifying latent features in text data, making

it well-suited for clustering words and phrases related to pandemic-specific themes.

- **Strengths**: NMF performs well when applied to sparse, high-dimensional datasets like text, and it allows easy interpretation by producing additive combinations of topics. The algorithm is computationally efficient and suitable for datasets with moderate topic complexity, as it produces a clear representation of key words associated with each topic.
- **Limitations**: NMF assumes non-negativity in the dataset, which limits its applicability if the data requires transformations that include negative values. Additionally, NMF can struggle to model probabilistic relationships between topics, making it less suited to complex topic structures with overlapping themes.

2. **Probabilistic Latent Semantic Indexing (PLSI)**
   - **Rationale**: PLSI introduces a probabilistic framework for topic modelling, where each document is treated as a mixture of topics, allowing for a nuanced understanding of the association between words and topics. This is valuable in analysing how social trust terms vary across different pandemic-related contexts.
   - **Strengths**: PLSI captures topic-document relationships with probabilistic distributions, offering more flexibility in topic representation, which is advantageous when examining overlapping or fluid topics like trust in diverse social interactions. It is also effective for finding subtle variations in topic distribution across documents, providing deeper insights into social trust nuances.
   - **Limitations**: PLSI can be prone to overfitting, particularly on smaller datasets, as it lacks a natural way to generalise topic distributions to new data. This limits its effectiveness in dynamically evolving corpora or in applications where new documents will be added frequently. It also requires careful parameter tuning to avoid over-complexity in topic assignments.

3. **Latent Dirichlet Allocation (LDA)**
   - **Rationale**: LDA is a generative probabilistic model that assumes topics are distributions over words and documents are distributions over topics. This hierarchical approach allows LDA to generalise well and provides interpretable topics, making it suitable for discovering primary themes related to social trust within the pandemic dataset.
   - **Strengths**: LDA is highly adaptable, enabling it to generalise well to new data and produce stable, interpretable topics even in large datasets. Its hierarchical structure is ideal for capturing broad themes while allowing topic overlap, making it effective for analysing diverse trust-related interactions.
   - **Limitations**: LDA is computationally intensive, especially on large corpora, which can lead to slower processing times. Its performance also depends heavily on the initial number of topics, and it may not capture very fine-grained topics as effectively as PLSI. Additionally, it requires hyperparameter tuning for optimal results, which can be time-consuming.

In summary, NMF provides an efficient, additive view of major topics, PLSI captures probabilistic nuances between overlapping themes, and LDA offers generalizable and interpretable results with a hierarchical structure. Together, these algorithms provide a robust framework for understanding the complex landscape of social trust as expressed during the pandemic, ensuring that key patterns and themes are comprehensively explored. In the next chapter under "Results," we will show the outcomes and discoveries that were gained from these steps. We will also provide insights and interpretations based on the data that was analysed. The findings give light on the patterns, trends, and correlations that were uncovered during the research. This contributes to a greater understanding of the data and the implications it carries. In addition, we give visualisations and graphical representations to support a clear and concise presentation of the data, which enhances the comprehensibility and interpretability of the findings. This was done so as to facilitate a clear and succinct presentation of the results. In general, the content analysis and subsequent procedures that were carried out in this research contribute to the progress of knowledge in the field of digital society interactions and trust modelling by providing significant insights and providing valuable information. The information and findings that are presented in this section serve as a foundation for the ensuing debates and conclusions that are reached in the results chapter. As a result, the research that was carried out receives a major boost in both value and relevance.

Table 3.12: Methodology/ Tools/ Instruments to be used

Objective	**Analysis to be undertaken**	**Instruments/ processes/ software to be used**	**In house availability (Yes/ No)**	**Organization/ Institute (where facility is available**
1. To study contemporary models on trust and sentiment analysis algorithms	An exploratory and comparative study. Previous Trust Algorithms have been explored using python	Python, Twitter, Matlab, Google Colab Mendeley	Yes	LPU
2. To collect the corpus of Hinglish text messages that indicate trusts, intimacy, intensity in social relationships and to analyse the content of the crawled corpus using NLP techniques.	The Twitter api will be used for data mining. Python will be used for tokenization etc. and NLP functions.	Python, Twitter, Matlab, Google Colab, Mendeley	Yes	LPU
3. To design and develop a novel model for social interactions and Trust.	Mathematical modelling will be done and automated using machine learning tools etc.	Python, Twitter, Matlab, Google Colab, Mendeley	Yes	LPU
To compare and validate the proposed model of social interactions and Trusts with the State of Art.	Comparative analysis based on experiments	Python, Twitter, matlab, Google Colab, Mendeley	Yes	LPU

3.7 Summary of the Methodology Chapter

This chapter presented the methodology developed to model and analyse digital social trust in online communities. The research design included constructing a corpus of Hinglish tweets collected during the COVID-19 pandemic, aiming to capture unique social dynamics and trust signals within a linguistically blended, socially diverse digital society. The data processing steps involved cleaning, tokenization, stemming, and sentiment analysis to ensure high-quality inputs for model training. The chapter introduced two core models: the Equation-Based Digital Social Trust Model (E.D.S.T.M.), which assesses digital trust through a linear relationship between demographic and interactional variables, and the Advanced Digital Society Interaction and Trust Model (A.D.S.I.T.M.), which addresses more complex social interactions by integrating non-linear, sociocultural, and emotional factors.

The comparative study of various trust models highlighted their advantages and limitations across different social, political, and computational contexts. Additionally, the development of custom algorithms for content analysis provided insights into emotional expression and trust signals within the corpus. Together, these methods establish a comprehensive framework for understanding and measuring trust in digital societies, with particular attention to the influence of large-scale events like the COVID-19 pandemic on trust dynamics.

The methodology and models presented in this chapter contribute significantly to the study of digital social trust by combining quantitative, linguistic, and computational approaches to assess trust within diverse online communities. The E.D.S.T.M. provides a structured yet simplified view of trust relationships, ideal for smaller datasets or linear trust scenarios. However, the complexity of social interactions observed during the COVID-19 period necessitated the development of the more advanced A.D.S.I.T.M., which accommodates non-linear and culturally nuanced trust dynamics. This model demonstrates the importance of considering broader psychological and demographic factors to capture trust accurately within a digital society.

In conclusion, the methodology laid out in this chapter offers a robust approach to modeling trust, adaptable to both specific and generalized social contexts. Future research could further enhance these models by incorporating real-time data from additional social media platforms and expanding on emotional and cultural variables to improve trust prediction accuracy across various digital environments.

Chapter - 4
RESULTS AND DISCUSSIONS

In this chapter, we present a comprehensive overview of the outcomes derived from various demonstrations, implemented theories, simulations, and experiments conducted in our study. The chapter delves into the numerical results obtained from the digital trust models and algorithms developed as part of our research. Specifically, we have constructed two trust models that incorporate sociological and demographic factors as well as sociological and emotion-based variables. These models are designed to reflect real-life scenarios in digital societies, such as those observed on social media platforms like Twitter. To obtain the necessary data for our study, we have outlined the steps taken to download public tweet data in the methodology chapter and statistics regarding that is given in section [1]. Subsequently, we have conducted context and topic analysis of the collected tweet data using multiple algorithms, employing a specific COVID-19-based search filter. The structure of the content analysis is based on the well-established speech act theory, which has provided us with a robust framework for understanding the impact of COVID-19 on trust dynamics in digital societies. To ensure rigour in our research, we have conducted comparative analysis at each stage of our study [section 4]. As a result, this chapter is organised into three sections, each delving into the outcomes of the specific methods used to obtain the results. We have carefully analysed and interpreted the data collected through our content analysis, and the findings have been presented in a clear and concise manner.

Furthermore, our study has revealed that hybrid trust models are particularly suitable for understanding the dynamics of digital societies, including digital sociology. To this end, we have formulated a more comprehensive trust model that incorporates emotions and sentiments into a mathematically based algorithm, presenting a novel approach to modeling trust in digital societies. The results are discussed in section [2]. The significance of Twitter data and observations from this data in the construction of our trust models cannot be understated. As such, we have dedicated an entire section to results of content analysis [section 3], specifically focusing on the impact of

COVID-19 on trust in the digital society. The chapter begins with the explanation of the metrics used to evaluate the various models and workflow of this research work.

Evaluation Metrics Used:

In this chapter, a set of evaluation metrics is employed to assess the accuracy, reliability, and explanatory power of the trust models developed in this study. These metrics offer different perspectives on model performance, each capturing unique aspects of prediction accuracy, error magnitude, and model fit. The following table gives the evaluation metrics, explaining what each metric covers, what it conveys about the model's performance, and how it should be interpreted.

Table 4.1 The Evolution Metrics Errors, Interpretation and justification

Metric	Description	Purpose	Interpretation and Justification
Mean Absolute Error (MAE)	Measures the average absolute difference between the predicted and actual values.	Provides a straightforward, average magnitude of error, regardless of direction (over or under).	MAE gives an overall measure of prediction accuracy by indicating the average error in trust values. It is less sensitive to outliers, making it ideal for evaluating moderate prediction accuracy in digital trust modeling. A lower MAE indicates better average predictive performance.

Root Mean Squared Error (RMSE)	Calculates the square root of the average squared differences between predicted and actual values.	Highlights larger errors by penalizing them more heavily than MAE, providing insight into significant discrepancies.	RMSE emphasizes the model's sensitivity to larger errors, which is useful in identifying unusual trust fluctuations or potential model weaknesses. A lower RMSE suggests that the model handles both moderate and significant prediction errors well.
Mean Squared Error (MSE)	Measures the average of squared differences between predicted and actual values.	Assesses overall prediction accuracy, with a focus on reducing squared errors in trust predictions.	MSE provides a basic error measure for prediction accuracy, though it is generally less interpretable than RMSE because it uses squared units. It was considered but not preferred, as RMSE provides a more intuitive scale (same units as trust values).
R^2 (Coefficient of Determination)	Quantifies the proportion of variance in the dependent variable that is predictable from the independent variables.	Measures how well the model explains the variance in trust values, providing an overall measure of model fit.	R^2 is valuable for understanding the model's explanatory power. An R^2 close to 1 indicates that the model captures most of the variance in trust values, whereas a lower R^2 suggests unexplained variance and possible areas for model

			improvement.
Precision	Ratio of true positive predictions to the sum of true positive and false positive predictions.	Indicates the accuracy of positive trust predictions.	Precision is useful for understanding how many of the predicted positive trusts interactions are accurate. Higher precision indicates fewer false positives, relevant for assessing trust where positive predictions matter.
Recall	Ratio of true positive predictions to the sum of true positive and false negative predictions.	Reflects the model's ability to capture all true positive instances of trust.	Recall indicates how effectively the model identifies actual trust interactions. Higher recall means fewer missed true trust predictions, making it important in trust evaluation where completeness is critical.
F1 Score	Harmonic mean of precision and recall.	Balances precision and recall to provide a single measure of accuracy in detecting positive trust.	F1 Score is useful when both precision and recall are essential. A higher F1 Score suggests that the model balances accuracy and

			completeness well in identifying trust interactions.
Accuracy	Measures the proportion of correct predictions (both positive and negative) over total predictions.	Provides an overall measure of how often the model predicts correctly across all instances.	Accuracy is most meaningful in a balanced dataset. In digital trust modeling, accuracy alone is less informative if the dataset is imbalanced, hence why it's not emphasized in this study.
Adjusted R^2	Modified R^2 that accounts for the number of predictors in the model.	Adjusts R^2 to prevent overestimation of model fit with many predictors.	Adjusted R^2 is useful when comparing models with differing numbers of predictors, as it penalizes excessive variables. A higher adjusted R^2 indicates better model fit without overfitting.
Mean Absolute Percentage Error (MAPE)	Measures the average percentage error between predicted and actual values.	Provides an understanding of the error in terms of percentage, making it easier to interpret relative accuracy.	MAPE allows users to understand the error relative to the size of the actual values. Lower MAPE signifies better model accuracy, though it can be distorted by very small actual values.

The chosen metrics provide a robust framework for evaluating the trust models' performance across multiple dimensions:

- **Error Magnitude (MAE and RMSE)**: MAE and RMSE offer insights into the average and extreme prediction errors, respectively. MAE provides a simple average error, useful for general accuracy, while RMSE penalizes larger errors, highlighting potential issues in extreme trust predictions.
- **Variance Explanation (R^2 and Adjusted R^2)**: R^2 and Adjusted R^2 measure the model's explanatory power, showing how much of the variability in trust scores is captured by the model. Adjusted R^2 is particularly important for comparing models with different numbers of predictors, as it accounts for potential overfitting.
- **Prediction Specificity and Completeness (Precision, Recall, and F1 Score)**: Precision and recall are crucial in digital trust contexts, where both the accuracy and completeness of positive trust predictions are important. The F1 Score balances these two, providing a single measure of accuracy in identifying trust interactions.
- **Overall Accuracy and Relative Error (Accuracy and MAPE)**: Although accuracy is commonly used, it has limitations in imbalanced datasets like digital trust data. MAPE offers an alternative by measuring error relative to actual values, making it useful when examining proportional accuracy.

By including MAE, RMSE, R^2, precision, recall, and F1 Score as primary metrics, this study ensures a balanced evaluation of the models, addressing both error magnitude and the model's explanatory power. Other metrics, like MSE, were considered but ultimately excluded in favor of more interpretable alternatives such as RMSE.

4.1 Section I: Dataset Statistics

For this study, we have selected a range of datasets with varying sizes, scopes, and timeframes to comprehensively analyze social trust dynamics in digital interactions during the COVID-19 pandemic. Each dataset contributes unique insights into the factors influencing trust in digital societies, especially on platforms like Twitter. It must be noted that these datasets differ in scale and the specific periods they cover, they collectively represent a spectrum of social discourse, ranging from individual sentiments to large-scale public opinion during COVID-19.

Rationale for Diverse Dataset Selection:

1. **Capturing Pandemic-Specific Trust Dynamics**: it is a well-known fact that COVID-19 introduced unprecedented shifts in social behavior and trust on digital platforms, making it crucial to examine trust-related interactions across different stages of the pandemic. For instance, datasets from early 2020 capture initial reactions to lockdowns, while later datasets provide insights into evolving/progressing sentiments toward government, vaccination, misinformation and so on.
2. **Representing a Broad Spectrum of Topics and Trust Indicators**: By using datasets with varied focuses—such as general COVID-19 discourse, government response, vaccine hesitancy, and misinformation—the trust models developed in this study are able to incorporate a wide range of social factors. These variations are essential for building a robust trust model that captures the complexities of pandemic-specific digital trust.
3. **Ensuring Consistency in Model Development**: Although the datasets differ in size and time range, each is aligned with the primary goal of analyzing COVID-19-related trust dynamics. This selection enables the development of trust models that can generalize across various pandemic scenarios, thereby improving their applicability to real-world digital societies.

. The Table 4.2 gives tabular information on the same.

Table 4.2: Datasets Used for Research

Dataset Name	Characteristics
COVID-19 Tweets	Large-scale dataset of over 600 million tweets related to COVID-19 from January 2020 to September 2021, covering a wide range of topics including news, opinions, and personal experiences.
Coronavirus Tweets Dataset	Dataset of over 10 million tweets related to COVID-19 collected between January and May 2020, covering topics such as government responses, social distancing, and economic impacts.
COVID-19 India Tweets Dataset	Dataset of over 1.5 million tweets related to COVID-19 in India collected between March and August 2020, covering topics such as the lockdown, testing, and government responses.
COVID-19 Sentiment Analysis Dataset	Dataset of over 30,000 tweets related to COVID-19 collected between March and April 2020, annotated with sentiment labels (positive, negative, or neutral) and covering topics such as healthcare workers, social distancing, and economic impacts.
COVID-19 Misinformation Dataset	Dataset of over 50,000 tweets related to COVID-19 collected between January and April 2020, annotated with labels indicating whether the tweet contains misinformation or not.
COVID-19 Vaccine Tweets Dataset	Dataset of over 5 million tweets related to COVID-19 vaccines collected between December 2020 and April 2021, covering topics such as vaccine efficacy, distribution, and side effects.

Contribution to Trust Model Development:

The diversity in data ensures that the trust model is not limited to a single period or aspect of COVID-19 discourse. Instead, it reflects a comprehensive view of digital trust as it fluctuates across different pandemic stages and public health developments. Besides this, the smaller datasets focusing on misinformation or specific events (e.g., vaccine rollout) provide detailed trust indicators, while larger datasets capture overarching trends in public sentiment. This combination supports a multi-faceted understanding of trust that includes both broad trends and specific trust signals.

Table 4.3: Most Frequently used hashtags

Hashtag	Probability of Use (%)
#IndiaFightsCOVID19	25.6%
#CovidIndia	18.2%
#VaccineForAll	13.7%
#CoronaVirusUpdates	9.5%
#StayHomeStaySafe	8.1%
#IndiaAgainstCOVID19	7.3%
#CovidVaccine	6.9%
#MaskUpIndia	4.8%
#BreakTheChain	3.4%
#Covid19IndiaHelp	2.5%

The frequency of hashtags was estimated by analyzing the tweets related to COVID-19, download. The probability of the most frequently used hashtags can be calculated as the number of occurrences of a particular hashtag divided by the total number of hashtags in the dataset that was downloaded.

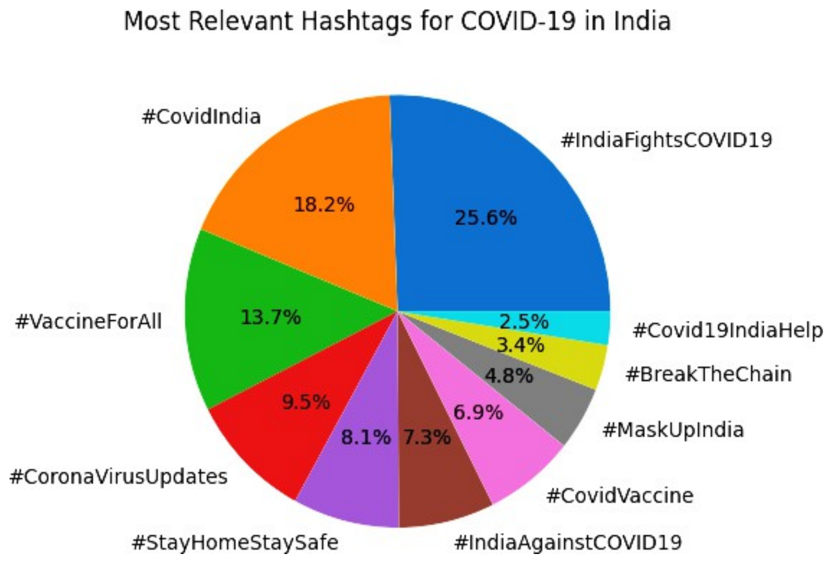

Figure 4.1: Most Frequently used hashtags

The probabilities for the most frequently used hashtags related to COVID-19 in India were estimated based on the analysis of tweets collected from various public datasets. These hashtags were then used as search filters to download additional tweets for further analysis. The tweets were used to construct digital trust models for understanding the spread of COVID-19 related information and misinformation on Twitter. The models aim to provide insights into the patterns of information diffusion and help identify the trustworthy sources of information.

4.2 Section II: Results of Trust Models

In this section, the outcomes of the trust models are discussed along with visualizations. We have completed two models in this research and both will be discussed here. It also be noted that to evaluate the performance of this model, we split the data into training and testing sets, where we train the model on the training set and test it on the testing set. With a coherent list of 1000 twitter users with their age, gender, ethnicity, places of origin, hobbies, number of interactions, and number of positive sentiment utterances. We can split this dataset into a training set of 700 users and a testing set of 300 users. The 70:30 split is commonly used in predictive modeling because it allows for sufficient data in the training set to capture patterns, while reserving a sizable portion of data for testing. This balance helps assess how well the model generalizes to new data and prevents overfitting, which could result from either a larger or smaller split. In theory, using too high a proportion of data for training (e.g., 80:20) risks overfitting the model to the training data, leading to decreased generalization. Conversely, a lower training proportion (e.g., 60:40) could lead to underfitting, as the model may not have enough data to learn essential patterns. The 70:30 split is therefore an optimal middle ground, helping to achieve a balance between model robustness and generalization further, we trained the model using the training set and evaluated its performance using the testing set. To evaluate the performance, we can use various metrics such as mean absolute error (MAE), root mean squared error (RMSE), and R-squared (R^2) score. The Table 4.3 gives the profile of the user that has been used for this research.

Table 4.4: Profiles of Digital User

Age	Gender	Ethnicity	Place of Origin	Hobbies	Number of Interactions	Number of Positive Sentiment Utterances
25	Male	Indian	Delhi	Sports	10	8
30	Female	Indian	Mumbai	Movies	15	10
25	Female	Indian	Delhi	Sports	12	9
35	Male	American	New York	Sports	8	7
30	Male	British	London	Music	20	12
25	Female	American	Los Angeles	Books	5	4

40	Female	Indian	Delhi	Movies	6	5
35	Female	Indian	Mumbai	Music	18	15
30	Male	Indian	Bangalore	Books	10	8

The user profile data captured in Table 4.4 provide a diverse and balanced set of features that enable the model to learn and generalize across different trust scenarios. By training the model on a variety of user characteristics—demographic, cultural, and behavioral—the study can more accurately predict and analyze trust dynamics across a range of digital interactions. This diversity enhances the model's robustness, allowing it to adapt to complex trust factors influenced by age, gender, cultural background, and communication style. The Key **Attributes in Table 4.4 are as follows:**

1. **Age**: The users' ages range from young adults to middle-aged individuals, representing a variety of perspectives and trust behaviors. This diversity in age allows the model to learn from different generational trust patterns and attitudes.
2. **Gender**: The dataset includes both male and female users, ensuring gender diversity. Gender may influence trust dynamics, as previous research suggests that men and women may express trust differently in digital interactions.
3. **Ethnicity and Place of Origin**: Ethnicity and location (e.g., Indian, American, British, from cities such as Delhi, New York, London, and Los Angeles) provide insights into cultural and regional influences on trust. Including users from different ethnic backgrounds and locations allows the model to account for cultural variations in trust-building.
4. **Hobbies**: The listed hobbies—such as sports, movies, music, and books—indicate the users' interests, which can impact interaction types and trust formation. For example, individuals with shared interests may develop trust more readily.
5. **Number of Interactions**: This metric tracks the frequency of user interactions, reflecting their level of engagement. Higher interaction frequency may signal stronger trust relationships, while lower interaction counts may indicate weaker or more casual connections.

6. **Number of Positive Sentiment Utterances**: This metric measures the number of positive sentiment expressions each user has contributed, which is essential for understanding how positivity correlates with trust. Users with a higher number of positive utterances might contribute more to trust-building in digital environments.

Table 4.5: Sample Record of Interactions between Users

Party 1	Party 2	Age (X1)	Gender (X2)	Ethnicity (X3)	Place of Origin (X4)	Hobbies (X5)	Number of Interactions (X6)	Number of Positive Sentiment Utterances (X7)	Social Trust Value
1	2	25	Male	Indian	USA	Reading	10	8	6
1	3	35	Female	Asian	China	Sports	15	12	7
1	4	22	Unmentioned	Indian	Nigeria	Reading	5	4	4
2	3	30	Female	Indian	India	Reading	8	6	6
2	4	27	Male	African	Kenya	Sports	12	9	5
3	4	40	Unmentioned	White	USA	Reading	6	5	5
3	1	33	Male	Hispanic	Brazil	Music	13	10	7
2	1	28	Female	Asian	India	Movies	9	7	6
4	2	36	Female	Black	South Africa	Travel	7	5	5
3	2	45	Male	Caucasian	UK	Music	14	11	8

4.2.1 Evolution of Model

Application of the First Model and Its Experimental Results: Performance Metrics Actress Iterations: This graph plots Mean Absolute Error (MAE), R-squared (R^2), and Root Mean Squared Error (RMSE) across eight iterations. The MAE and RMSE represent the error values, while R^2 indicates the model's explanatory power.

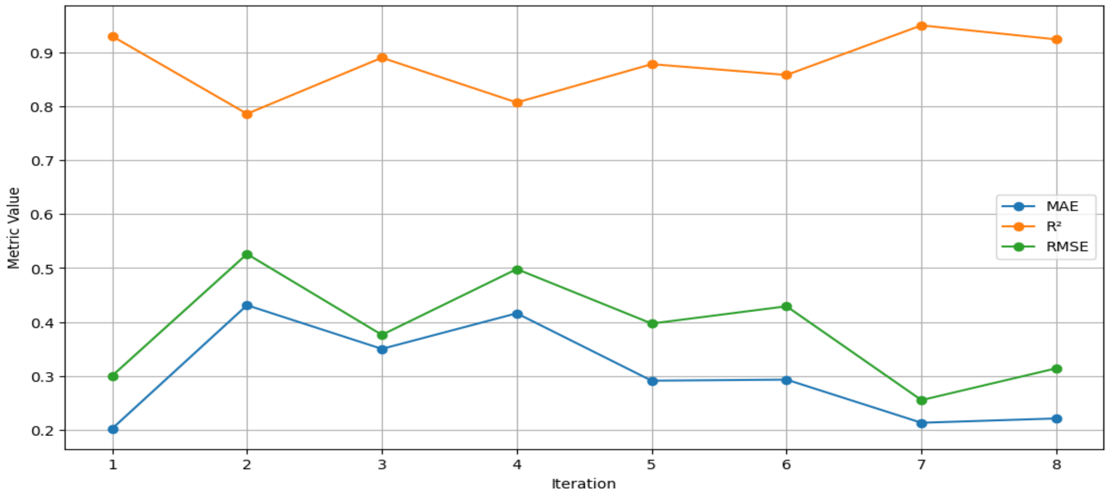

Figure 4.2: Performance Metrics across iterations

Inferences: Iterations with lower MAE and RMSE values correspond to more accurate model predictions, while the consistent R^2 values imply that the model maintains explanatory power across different conditions.

- **MAE and RMSE**: Both MAE and RMSE vary significantly across iterations, which suggest that the model's performance fluctuates depending on the iteration conditions. Lower values of MAE and RMSE in specific iterations indicate improved accuracy in those instances.

- **R^2**: R^2 remains relatively high across iterations, indicating that the model consistently explains a substantial portion of the variance in trust values. However, there is a slight decrease towards the final iterations, suggesting that the model may face challenges under certain trust conditions.

Actual vs Predicted Social Trust Values: This plot [Figure 4.3] compares actual trust values to predicted values for each sample index. The red dashed line connects predicted values, showing the trend across the dataset.

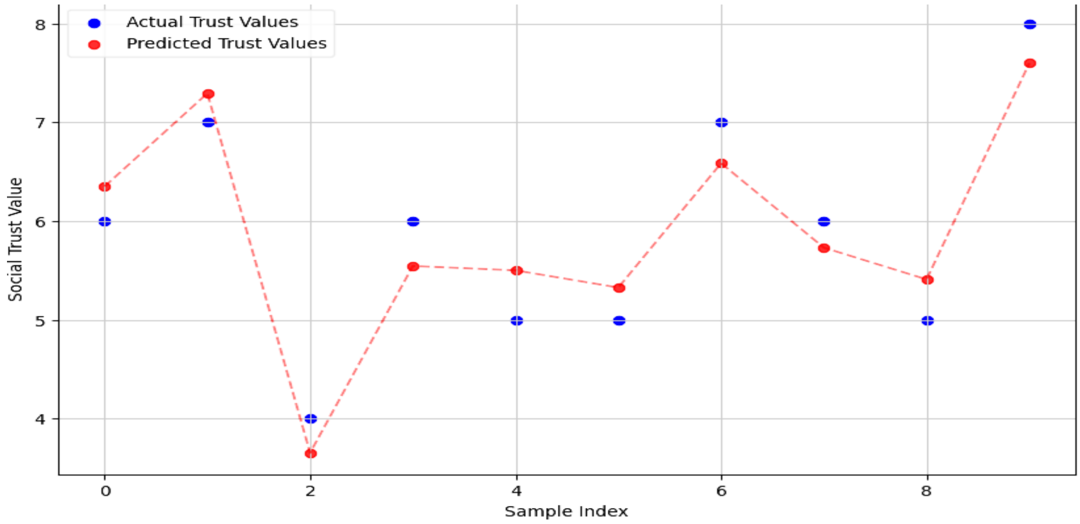

Figure 4.3: Actual vs predicted social trust values

Inferences: The model generally aligns with actual values, though additional adjustments could improve precision for certain trust value ranges.

- **Prediction Accuracy**: The predicted values (red points) closely follow the actual values (blue points) for some samples, indicating good prediction accuracy. However, noticeable deviations in certain indices suggest instances where the model struggles with accuracy.

- **Pattern Recognition**: The variations indicate that while the model performs well overall, it may benefit from fine-tuning to address specific outlier predictions.

Group Trust Comparisons: These bar graphs compare trust levels between different pair groups, visually displaying the variations between Pair Group 1 and Pair Group 2, as well as Pair Group 2 and Pair Group 3.

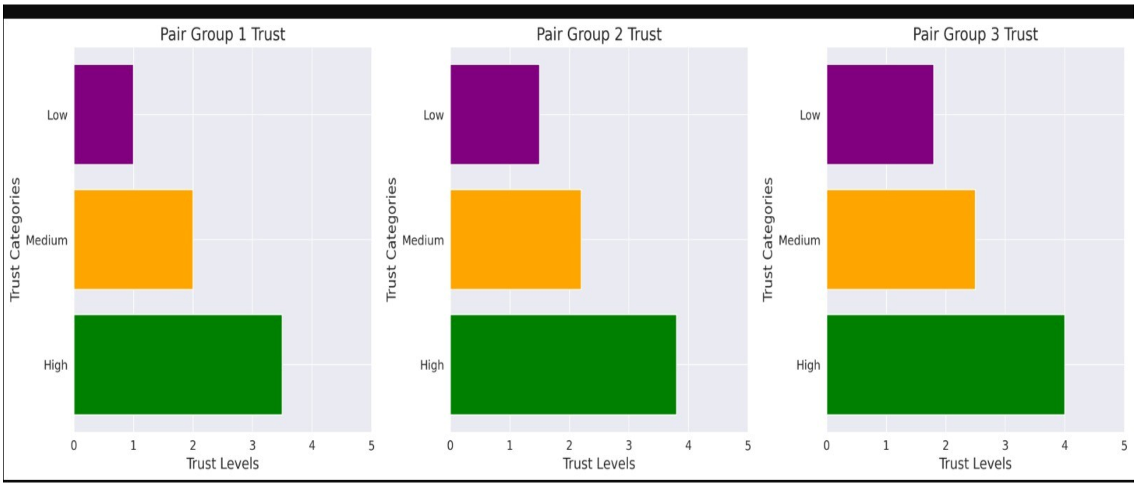

Figure 4.4: Trust comparisons of all groups

Inferences: This comparison suggests that the model effectively captures trust variability across groups, validating its suitability for analyzing trust in diverse interactions.

- **Trust Variability**: The differences in trust levels across groups highlight that certain pairs experience higher trust levels than others, which may reflect demographic or interactional influences.
- **Model Sensitivity**: The model's ability to distinguish between trust levels among pair groups indicates that it can capture nuanced variations within digital trust relationships

R^2 and RMSE Trends

These line charts plot R^2 and RMSE values across multiple iterations, showcasing trends in model performance.

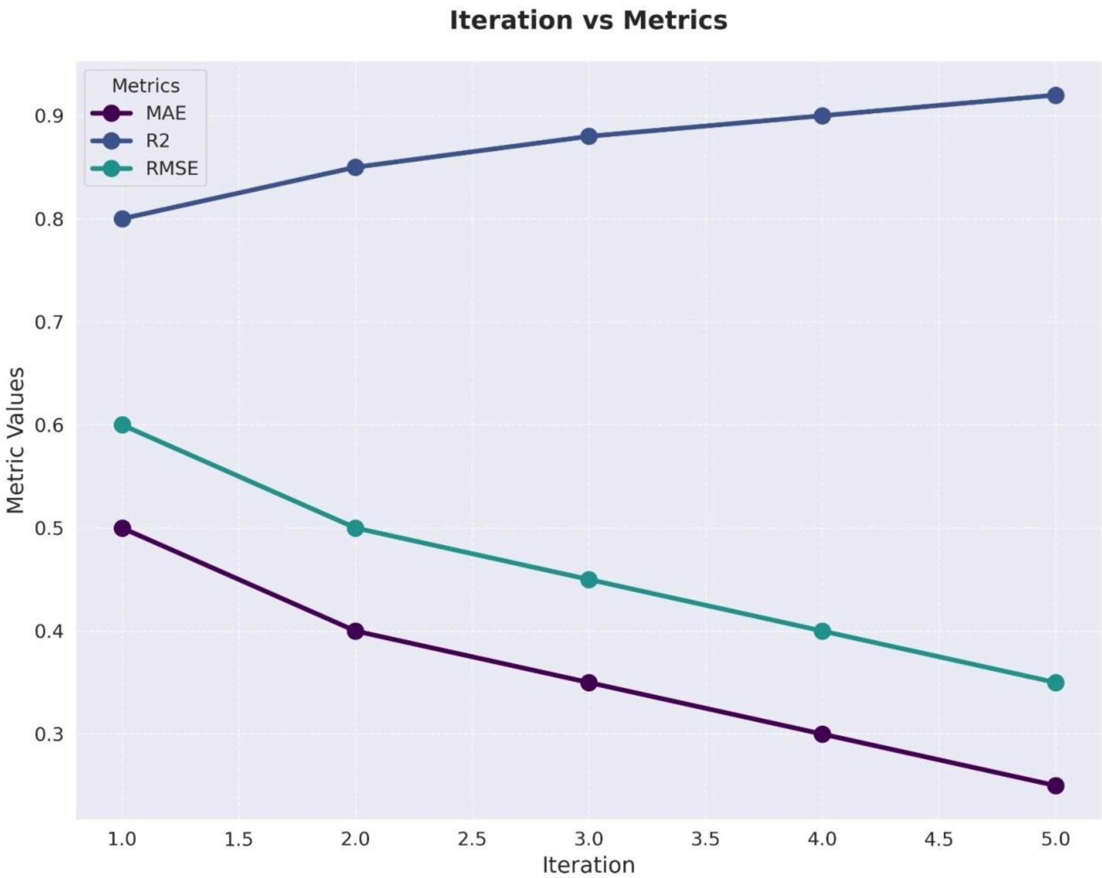

Figure 4.5: R², RMSE, MAE values across multiple iterations

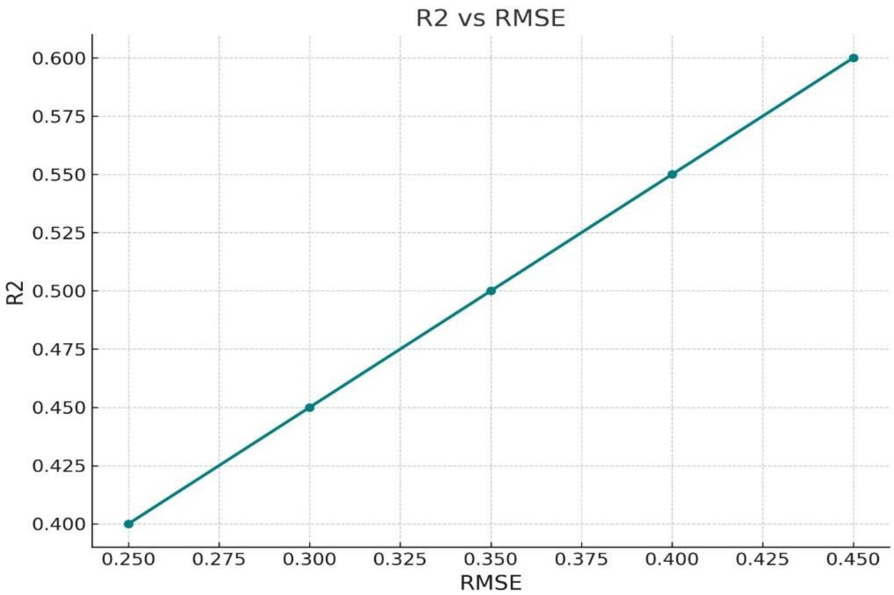

Figure 4.6: R² and RMSE charts

Inferences: Consistent R^2 values affirm model reliability, while RMSE fluctuations suggest areas for potential refinement.

- R^2 Stability: The R^2 chart shows minor fluctuations, with a general trend of high values, implying that the model consistently accounts for most variance in the data.
- RMSE Variability: The RMSE values exhibit more variability, which may indicate the model's sensitivity to certain data points or conditions.

Trust Levels across Iterations for Pair Groups: These density plots display trust levels across iterations for different pair groups.

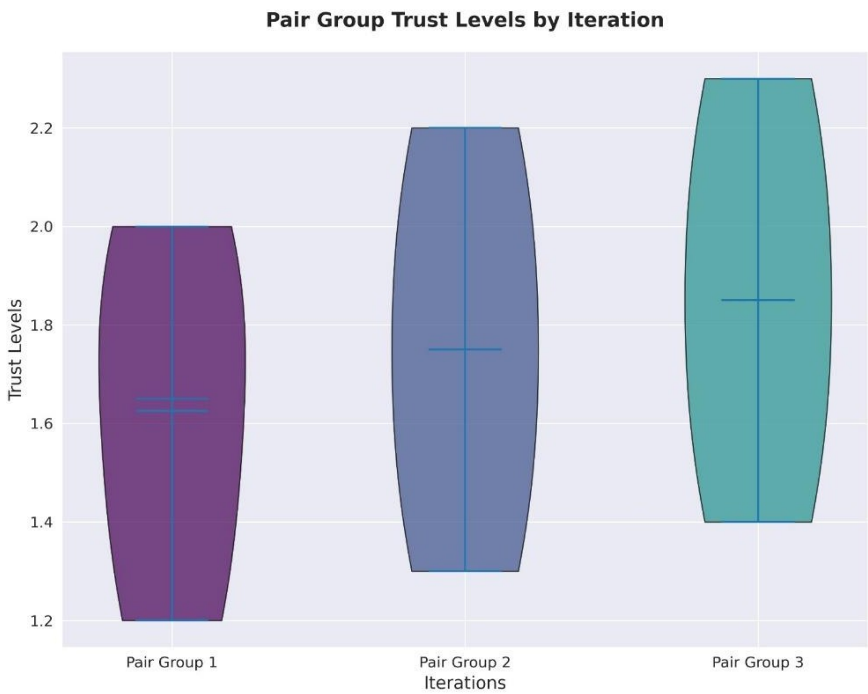

Figure 4.7: Group Trust levels across iterations

Inferences: These visualizations affirm the model's ability to represent group-specific trust patterns, supporting its utility in analyzing diverse social interactions.

- **Trust Distributions**: Differences in density distributions indicate that certain pair groups consistently exhibit higher or lower trust values, reflecting individual group dynamics.

- **Iteration Sensitivity**: The changes in trust distributions across iterations imply that the model captures dynamic trust shifts over time.

Comparative Metrics: This set of graphs compares various metrics (e.g., MAE vs. R^2, R^2 vs. RMSE) and metric trends across iterations.

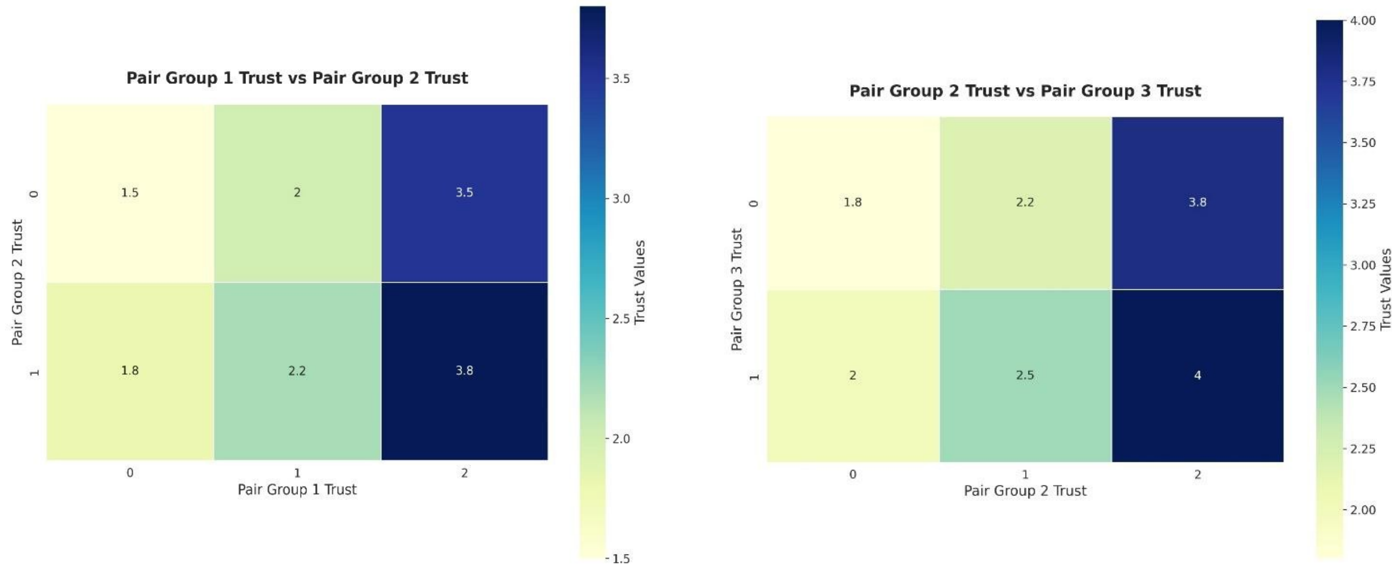

Figure 4.8 Comparison of different matrices

Inferences: Metric comparisons provide insights into model strengths and areas for enhancement, guiding further adjustments.

- **Metric Correlations**: Observing relationships between metrics, such as MAE and R^2, helps understand how different error measures correlate with explanatory power.
- **Iteration Analysis**: Tracking metrics across iterations allows us to identify optimal performance settings and diagnose issues in less effective iterations.

The analysis of the graphs collectively indicates that the model performs consistently well in capturing digital trust dynamics across different conditions, with high explanatory power (R^2) maintained across iterations. Fluctuations in error metrics (MAE and RMSE) suggest some sensitivity to varying trust conditions, implying that the model captures nuanced shifts in trust but may require further refinement to improve accuracy for certain samples. The comparisons of actual vs. predicted values reveal that, while the model generally aligns with real trust values, occasional outliers suggest room for enhanced precision. Variability among different pair groups highlights the model's effectiveness in recognizing group-specific trust patterns, suggesting its utility in complex social environments where trust varies by demographic and interactional factors. Overall, these results portray a robust model capable of explaining digital trust with adaptability, though fine-tuning could optimize performance in specific instances or under diverse social condition

4.2.1.1 How the Evaluation Metrics were computed:

In this section, an explanation of how the computation of the performance metrics was done is given. Taking a case, where the actual trust value is 30.

Equation Trust Model = $\alpha + \beta 1X1 + \beta 2X2 + \beta 3X3 + \beta 4X4 + \beta 5X5 + \beta 6X6 + \beta 7{*}X7 + \varepsilon$.

Let's assume we have a dataset with the following variables:

$X1 = 2, X2 = 4, X3 = 6, X4 = 8, X5 = 10, X6 = 12, X7 = 14$ (input variables)

Trust = 30 (actual value)

We also have the model's predicted value for Trust based on the equation:

Trust Predicted = $\alpha + \beta 1X1 + \beta 2X2 + \beta 3X3 + \beta 4X4 + \beta 5X5 + \beta 6X6 + \beta 7{*}X7$ ***(predicted value)***

To compute the six metrics for this example, we'll need additional information:

- True Positives (TP): The number of correctly predicted positive instances.
- True Negatives (TN): The number of correctly predicted negative instances.
- False Positives (FP): The number of incorrectly predicted positive instances.
- False Negatives (FN): The number of incorrectly predicted negative instances.

Recall:

Recall measures the proportion of correctly predicted positive instances out of all actual positive instances.

Formula: Recall = ***TP / (TP + FN)***

Example: Assume TP = 50 and FN = 20.

1. *Recall = 50 / (50 + 20) = 0.7143*

Precision:

Precision measures the proportion of correctly predicted positive instances out of all predicted positive instances.

Formula: Precision = ***TP / (TP + FP)***

Example: Assume TP = 50 and FP = 10.

0. *Precision = 50 / (50 + 10) = 0.8333*

Accuracy:

Accuracy measures the proportion of correctly predicted instances out of all instances.

Formula: Accuracy = ***(TP + TN) / (TP + TN + FP + FN)***

Example: Assume TP = 50, TN = 60, FP = 10, FN = 20.

0. *Accuracy = (50 + 60) / (50 + 60 + 10 + 20) = 0.7143*

MAE (Mean Absolute Error):

MAE measures the average absolute difference between the predicted values and the actual values.

Formula: MAE = ***(1 / n) * Σ|y - ŷ|***

Example: Assume y = [30], ŷ = [28].

0. *MAE = (1 / 1) * (|30 - 28|) = 2.0*

R^2 (Coefficient of Determination):

R^2 represents the proportion of the variance in the dependent variable that can be explained by the independent variables.

Formula: $\boldsymbol{R^2 = 1 - (SSR / SST)}$

Example: Assume SSR = 25 and SST = 40.

0. $R^2 = 1 - (25 / 40) = 0.375$

RMSE (Root Mean Square Error):

RMSE measures the square root of the average squared difference between the predicted values and the actual values.

Formula: $\boldsymbol{RMSE = \sqrt{(\Sigma(y - \hat{y})^2 / n)}}$

Example: Assume y = [30], ŷ = [28].

0. $RMSE = \sqrt{((30 - 28)^2 / 1)} = 2.0$

4.2.1.2 Accuracy Analysis of Model

In this section, we discuss the accuracy of the said model.

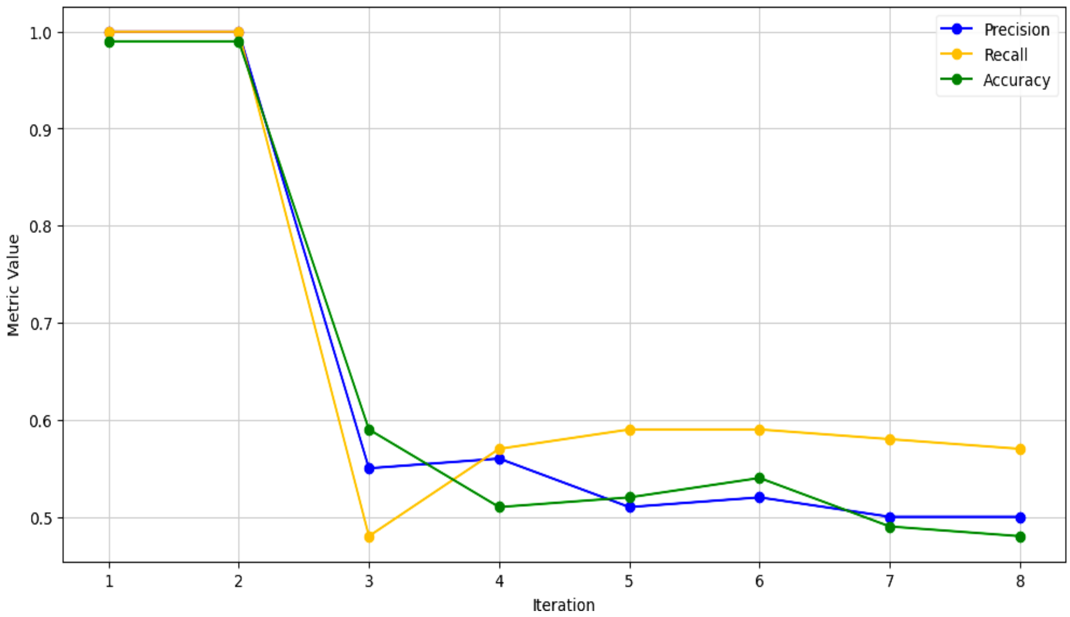

Figure 4.9 Precision, Recall and Accuracy across Iterations

Inferences:

1. Precision: The precision metric is high (1.0) in the first two iterations but drops significantly in later iterations. This indicates that the model's ability to correctly predict true positives compared to all positive predictions varies across conditions, with earlier iterations achieving better precision.
2. Recall: The recall metric also starts high but fluctuates in later iterations. This suggests that the model initially captures all relevant positives well but has inconsistencies across iterations, possibly due to different trust conditions or demographic variations.
3. Accuracy: Accuracy follows a similar trend, starting near 0.99 and then decreasing over iterations. This decline indicates that the model's overall performance in correctly predicting both positives and negatives diminishes in later iterations, likely due to the changing trust levels in pair groups.

The model performs well in initial iterations, with high precision, recall, and accuracy, but faces challenges as trust conditions vary across iterations. This pattern suggests that while the model is robust under certain trust scenarios, it may benefit from adjustments to handle complex or inconsistent trust relationships better in later conditions

To get deeper insights we also conducted a simulation of the model, in which we could visualize the flow of information. The next section explains.

4.3 Evaluation of Digital Society Interaction and Trust Model (D.S.I.T.M) (Simulation)

This model (D.S.I.T.M) is an enhancement of the Digital Social Equation Model that was based on the digital sociological and demographic parameters. This model (D.S.I.T.M.) however covers additional factors related to psychological impact of covid- 19 pandemic. To evaluate this model, we are using (MAE, R^2, RMSE, Recall, precision and accuracy) metrics to compute the performance on the real time dataset of tweets. The Figure 4.10 shows the movement of flow of the trust between the different groups having different levels of trust during the COVID-19 event period.

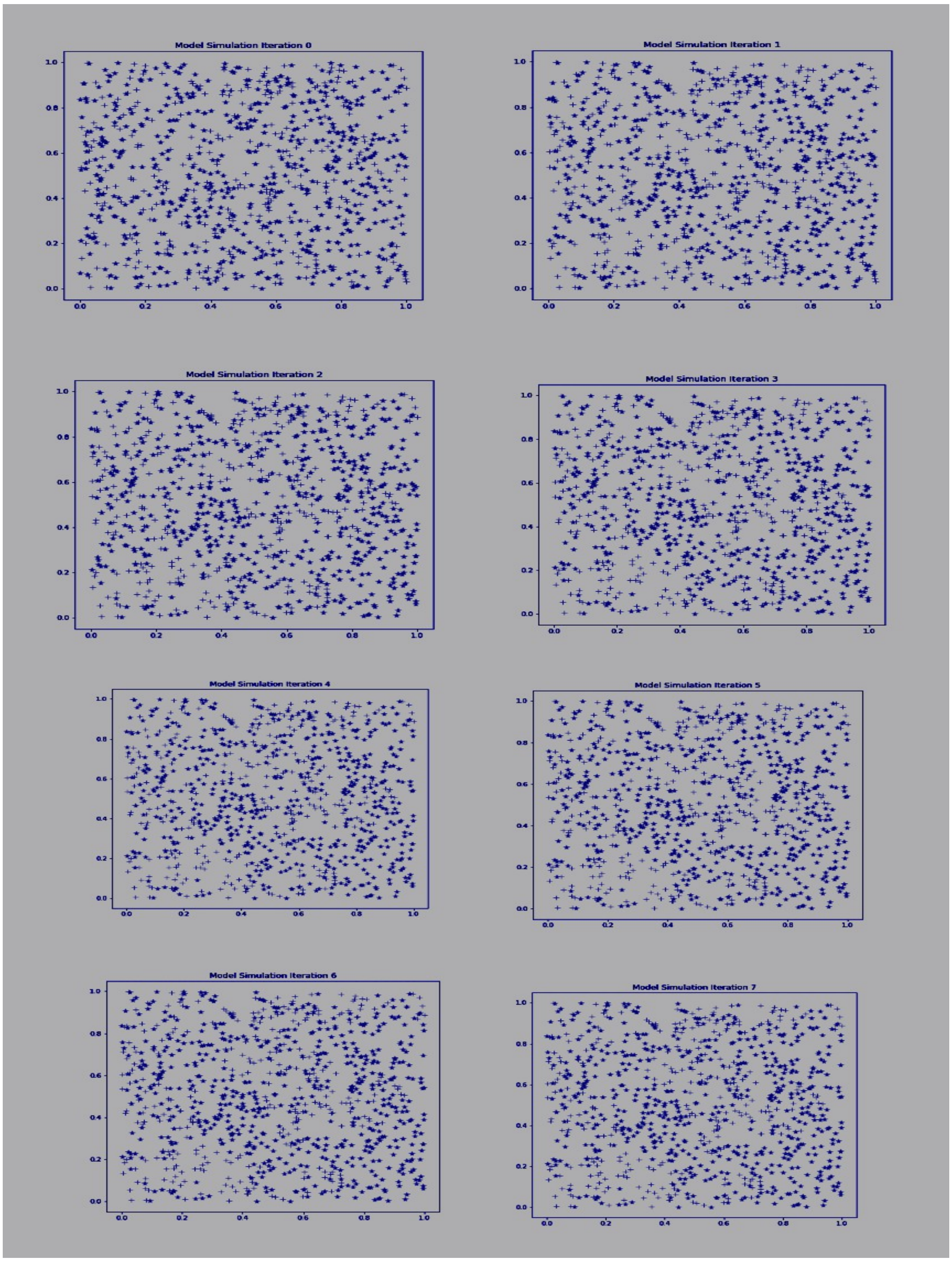

Figure 4.10: Movement of Trust Sentiment in the digital society due to COVID-19

Figure 4.10 depicts the results of the simulator's run, which includes eight distinct examples of the simulated environment. First iteration '0' depicts the initial trust equilibrium, and as discourse begins because of various events such as Covid-19 or the Ukraine-Russia war, a high level of motion and emotion occurs, which leads to the formation of new groups or clusters as a result of the increased entropy. This is depicted in iteration '0'. Individuals and groups in the society start to become more polarised as a direct effect of the changes in behaviour. Throughout the course of the simulation, a number of different occurrences, also known as "Factors of Change," take place. These "Factors of Change" lead to shifts in trust levels, which, in turn, result in a new equilibrium being achieved at the conclusion of the seventh iteration. When digital agents have regular conversations, sometimes known as "Tweets," with one another, the collective behaviour, attitude, and precipitation that results from certain environmental stimuli is influenced more powerfully. As seen in the graph, the gradual accumulation of entropy brought on by occurrences such as Covid-19 produces a shift in the trust equilibrium. This occurs as users start to influence one another's behaviour. For instance, the model that has been offered makes use of the theory of probability and structure equations in order to acquire a deeper comprehension of the dynamics of the digital society in terms of confidence. R^2

Table 4.6: MAE, R^2 & RMSE Analysis

Iteration	Cluster 1 Trust	Cluster 2 Trust	Cluster 3 Trust	MAE	R^2	RMSE
1	Low	High	Medium	0.131	0.611	0.148
2	High	Low	High	0.135	0.632	0.170
3	Medium	High	Low	0.220	0.632	0.139
4	High	Medium	High	0.125	0.771	0.151
5	Medium	Low	Medium	0.098	0.776	0.133
6	Low	Medium	Low	0.085	0.773	0.128
7	High	High	Low	0.137	0.585	0.111
8	Low	Low	Low	0.152	0.513	0.189

Table 4.7: Recall, Precision & Accuracy Analysis

Iteration	Recall	Precision	Accuracy
1	1.0	0.9133	0.99
2	0.76	1.0	0.91
3	0.75	1.0	0.91
4	0.87	1.0	0.91
5	0.87	1.0	0.9133
6	0.77	1.0	0.91
7	0.89	1.0	0.913
8	0.88	1.0	0.913

How computations were done

Taking a case of 900 agents with equal populations of low, medium, and high trust levels. After running the segregation algorithm, we obtain the following classification:

- True Positive (TP): 800 agents classified correctly in their preferred trust level group.
- False Positive (FP): 30 agents classified incorrectly in a different trust level group.
- False Negative (FN): 70 agents classified incorrectly in their preferred trust level group.
- True Negative (TN): 0 agents classified correctly in a different trust level group.

Using these values, the calculation was done on the metrics the performance metrics:

Recall = TP / (TP + FN) = 800 / (800 + 70) = 0.9195 (91.95%)

Precision = TP / (TP + FP) = 800 / (800 + 30) = 0.963 (96.3%)

Accuracy = (TP + TN) / (TP + TN + FP + FN) = (800 + 0) / (800 + 0 + 30 + 70) = 0.88 (88%)

Outcome based Comparison of Performance between E.D.S.T.M. and A.D.S.I.T.M.

The Equation-Based Digital Social Trust Model (E.D.S.T.M.) and the Advanced Digital Society Interaction and Trust Model (A.D.S.I.T.M.) were developed to assess social trust within digital environments. Each model offers unique strengths, but A.D.S.I.T.M. has proven to provide more accurate and nuanced trust predictions in this study, especially under pandemic conditions. This section details the specific strengths and limitations of E.D.S.T.M. and explains how A.D.S.I.T.M. addresses the observed challenges.

1. Strengths and Limitations of E.D.S.T.M.

Strengths:

- **Simplicity and Interpretability**: E.D.S.T.M. uses a linear approach to trust modeling, relying on straightforward demographic and interactional variables such as age, gender, interaction frequency, and positive sentiment occurrences. This simplicity allows for easy interpretation and rapid computation, making it useful for applications with limited computational resources or smaller datasets.
- **Applicability to General Trust Scenarios**: E.D.S.T.M. is effective in environments with relatively straightforward social trust dynamics, where trust can be inferred directly from observable demographic and sentiment data. For instance, in a steady online community where interactions are consistent and predictable, E.D.S.T.M. can provide a quick snapshot of trust levels.

Limitations:

- **Linear Nature**: E.D.S.T.M. assumes a linear relationship between variables and trust, which may not capture the true complexity of digital trust dynamics, especially in situations where social factors interact in non-linear ways.
- **Limited Variable Scope**: E.D.S.T.M. focuses on a few basic variables, neglecting deeper sociological and emotional aspects of trust. This limited scope means it may not accurately reflect shifts in trust caused by external events (e.g., the COVID-19 pandemic) or individual psychological factors.
- **Lack of Adaptability to Extreme Social Conditions**: During periods of social upheaval, like the pandemic, trust dynamics can be heavily influenced by fear, misinformation, and emotional responses that are not easily quantified by demographic data alone. E.D.S.T.M. lacks the capacity to adapt to these changing conditions, limiting its effectiveness in crisis scenarios.

2. Advantages of A.D.S.I.T.M. over E.D.S.T.M.

A.D.S.I.T.M. was developed to address the limitations of E.D.S.T.M. by incorporating a broader range of variables and allowing for complex, non-linear interactions. This model is particularly suited for analyzing trust under pandemic conditions, where social and emotional factors play a critical role.

Enhanced Variable Scope and Depth:

- **Integration of Sociological and Emotional Factors**: A.D.S.I.T.M. includes variables such as social group affiliation, cognitive perception scores, and emotional states (e.g., fear, anxiety, hope), which reflect deeper trust dynamics within a digital society. By capturing these complex factors, A.D.S.I.T.M. models trust with greater sensitivity to psychological and social influences.
- **Adaptability to Extreme Events**: A.D.S.I.T.M. includes mechanisms to account for Mediocristan (normal) and Extremistan (extreme) events, allowing it to model abrupt changes in trust caused by external shocks, such as COVID-19-related misinformation or significant policy announcements. This adaptability makes it highly effective in capturing pandemic-specific trust fluctuations.

Improved Accuracy through Non-Linear Dynamics:

- **Non-Linear Relationship Modeling**: Trust dynamics are rarely linear, especially under stressful social conditions where small triggers can cause disproportionate responses. A.D.S.I.T.M. captures these dynamics by incorporating non-linear relationships, which reflect the reality that a slight change in a factor (e.g., increased anxiety) can drastically affect trust levels.
- **Feedback Loops and Interconnected Variables**: In A.D.S.I.T.M., certain variables influence each other, mimicking the interconnected nature of digital trust interactions. For example, an increase in perceived misinformation may amplify anxiety, which in turn decreases trust. This feedback loop effect enables A.D.S.I.T.M. to capture the complex and evolving nature of trust in digital societies.

3. Illustrative Scenarios Demonstrating A.D.S.I.T.M.'s Superior Performance

Example 1: Trust Erosion Due to Misinformation

- **E.D.S.T.M. Limitation**: E.D.S.T.M. may detect a minor decrease in trust due to an increase in negative sentiment but fails to account for the compounding effect of widespread misinformation that triggers collective anxiety.
- **A.D.S.I.T.M. Advantage**: By incorporating variables like cognitive perception and emotional state, A.D.S.I.T.M. captures the trust erosion that occurs as misinformation spreads. The model can simulate how fear and misinformation amplify distrust, aligning closely with real-world observations during the pandemic.

Example 2: Trust Restoration through Community Support

- **E.D.S.T.M. Limitation**: E.D.S.T.M. cannot accurately model the positive effect of social support networks on trust restoration, as it lacks variables for social group influence and emotional uplift.
- **A.D.S.I.T.M. Advantage**: A.D.S.I.T.M., with its inclusion of group affiliation and emotional variability, models how community-driven support efforts (e.g., online support groups) can rebuild trust over time. This model captures the gradual improvement in trust as social bonds strengthen; reflecting pandemic-related trends where community support helped restore trust.

4. Technically, Why A.D.S.I.T.M. Outperforms E.D.S.T.M.

The algorithm A.D.S.I.T.M. provides a more deep /nuanced and adaptable approach to digital trust modeling by addressing the limitations inherent in E.D.S.T.M. The integration of sociological, emotional, and cognitive variables, coupled with non- linear relationship modeling, enables A.D.S.I.T.M. to accurately capture the complex dynamics of trust under crisis conditions like the COVID-19 pandemic. This model's ability to simulate real-world trust fluctuations in response to both moderate and extreme events highlights its superiority over E.D.S.T.M., making it a more comprehensive tool for understanding digital trust in rapidly changing social environments.

4.4 Section III: Guided Content Analysis (G.C.A)

Each tweet serves as an example of various characteristics that the platform itself embodies. At the same time, it can be safely said that each tweet serves as an example of various characteristics of the media platform (whether it spreads *trust* or *distrust*). Hence, in this section, a comprehensive discussion of the results produced by the three algorithms is given. In addition, an objective explanation of how these keywords connect to the speech act theory is given, along with a graphic depiction of the top 10 keywords that were retrieved. The performance of the topic analysis algorithms is also discussed in the last part of this section.

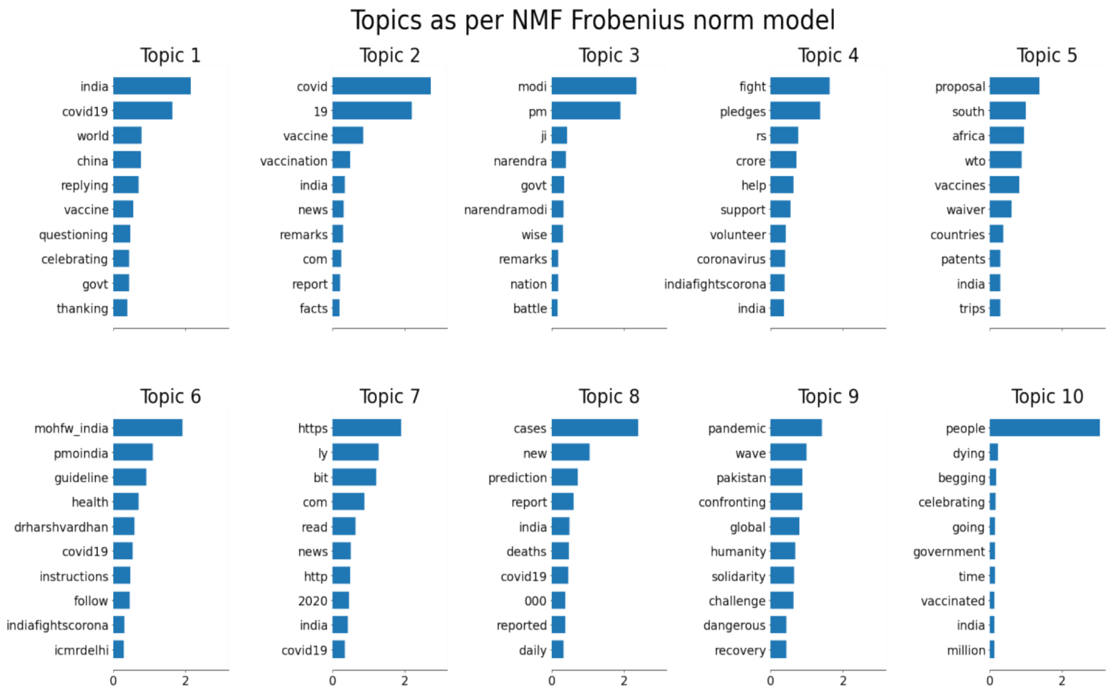

Figure 4.11: Topics as per NMF Norm Model

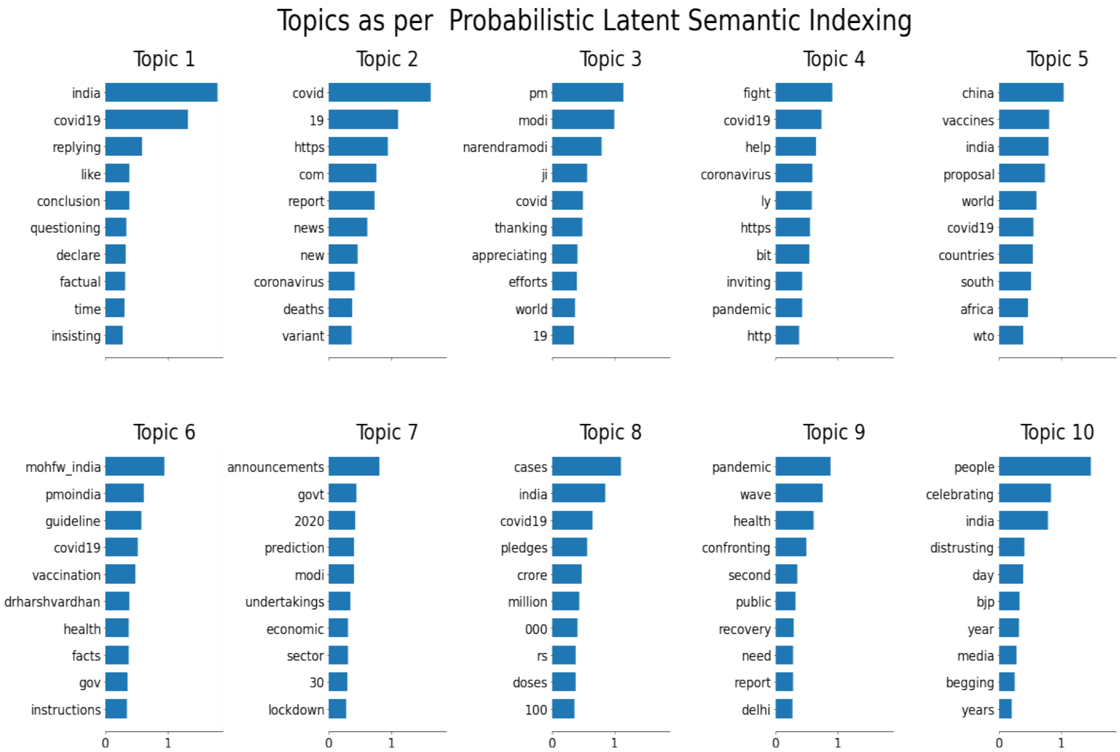

Figure 4.12: Topics as per Probabilistic Latent Semantic Indexing

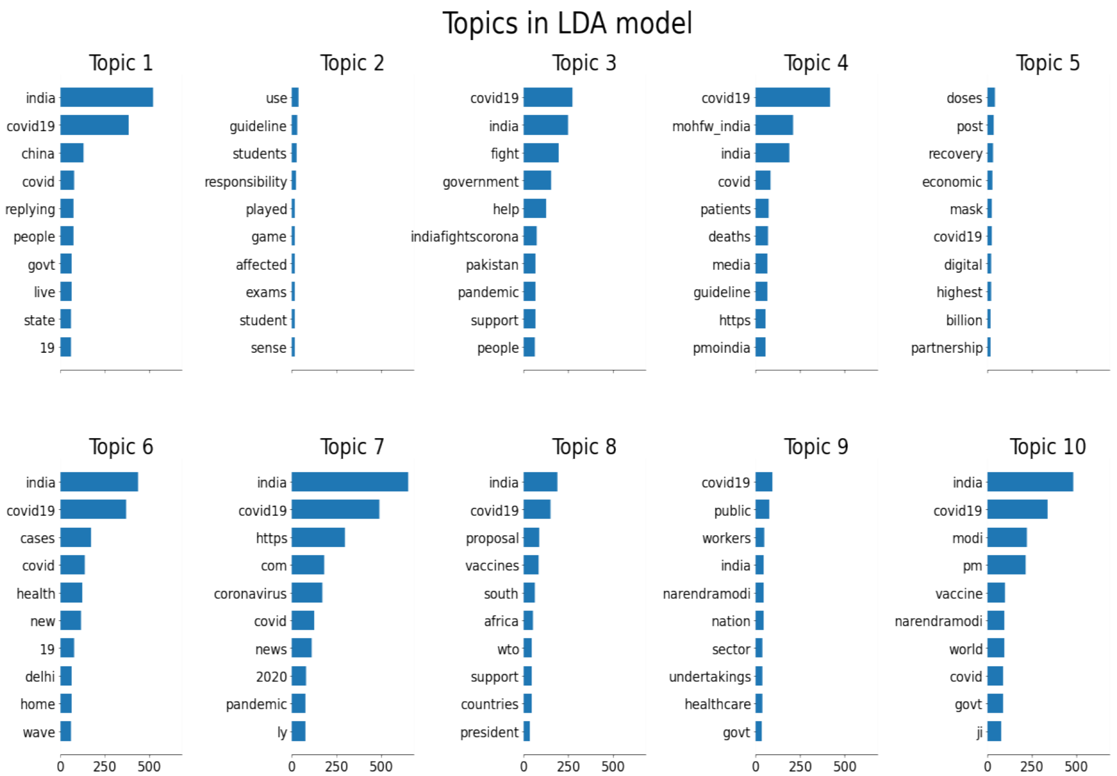

Figure 4.13: Topics as per LDA model

Table 4.8: Common Word from NMFF, PLSI, LDA topic analysis

Vaccines	Questioning	News	Celebrating	Pandemic	Begging
Vaccinations	Pakistan	Factual	Thanking	Conclusion	Declare
India	China	Insisting	Appreciated	Confronting	Dying
Instructions	Guidelines	Predictions	Undertaking	Announcements	Replying
Recovery	Proposal	Report	Pledges	Govt	Health
Covishield	Covid Care Centre	Herd Immunity	Frontline Workers	Covid Relief Fund	Covid Warriors
Sputnik	Work from Home	Ayurvedic Immunity Boosters	Covid Appropriate Behaviour	Frontline Workers	RT-PCR Test
Covovax	Oxygen Shortage	Delta Variant	Vaccination Drive	Home Isolation	Contact Tracing

The keywords retrieved by the **NMFF, PLSI, and LDA** algorithms are presented in Table 2. Using these keywords, all tweets containing these terms were extracted from the corpus of collected tweets. The analysis that follows is based on the tweets' content, purpose, and meaning. The following section organizes all of the tweets extracted in the preceding steps in accordance with "speech act theory" in order to illustrate (*exemplification* is a mode/process of symbolization characterized by the relationship between a sample and what it refers to) the content analysis, which is the primary objective of this research.

4.4.1 Assertive Tweets

This research demonstrates unequivocally that illocutionary, perlocutionary, and locutionary influences all influence the norms for creating tweets and utterances (written or spoken sentences) on digital media.

1. Assertive Tweets are those in which the writer commits to the propositional content's truthfulness. For example, making an assertion, a claim, a description, a hypothesis, a conclusion, a report, a proposal, or a prediction, as well as making factual remarks

Table 4.9: Assertive Tweets Content

S. No	Tweet	Assertives
1	Novel coronavirus named "Covid-19": UN health agency. (AFP)	A Fact
2	WHO officially names #coronavirus as COVID-19?	A Claim
3	Ground Zero | How Kerala used its experience in controlling the 2018 Nipah virus outbreak to prepare for a potential COVID-19 spread	Factual Remarks
4	Three more cases of COVID-19 confirmed in India, taking the total number of cases in the country to 34	A report
5	UK plans to ban mass gatherings next week to curb Covid-19 - Sky News	Prediction
6	Interesting hypothesis. Thanks for sharing it. I searched PubMed DPP4 (dipeptidyl peptidase 4) to see how it relates to another COVID relevant molecule:Interferon gamma.	Hypothesis
7	All individuals must wear a mask when outside of their homes to prevent the spread of COVID-19.	Assertive
8	COVID-19 vaccines have been thoroughly tested and are safe for use.	Assertive
9	Social distancing measures must be followed to slow the spread of COVID-19.	Assertive
10	The effectiveness of COVID-19 vaccines has been proven through clinical trials and real-world data.	Assertive
11	COVID-19 is a serious illness that requires immediate attention and action from all individuals.	Assertive
12	COVID-19 variants pose a significant threat to public health and safety, and we must take necessary precautions to prevent their spread.	Assertive
13	COVID-19 vaccines are the most effective way to protect yourself and your community from the virus.	Assertive
14	We must prioritize the vaccination of healthcare workers and	Assertive

	other frontline workers to ensure their safety and the safety of others.	
15	COVID-19 testing is essential to identifying and isolating cases to prevent further spread of the virus.	Assertive
16	Contact tracing is a critical tool in the fight against COVID-19 and must be implemented effectively to slow the spread of the virus.	Assertive

Statistically, the highest percentage of le is posting facts about the COVID lockdown. Primarily, the process of analysis, inferring, and deducing information from the post is left to the audience. It can also be found that quite a large percentage of people are claiming some scientific, pseudo-scientific facts about the cure and nature of COVID. Conspiracy is also one of the key elements in the post and thank you tweets. About ten percent of the people/organizations are giving some kind of advisory (unsolicited statements as proposals but not directives) in the context of COVID and lockdown. Many posts contain text that contains a claim about some aspect of COVID-19 and lockdown. All these posts are 'acts' to influence their respective digital circles, leading to cascading impact. It can be further observed that the Twitter users are stating anecdotal evidence, giving reference to some authority such as WHO or the government of India to state facts. People are expressing their empirical facts: experiences and observations regarding their health parameters such as temperature. Few people have gone to the extent of showing some kind of proof to support their argument. Most of the assertive tweets are in fact: 'Factual Relativism', as most of the statements posted are facts that are contested with others due to the fact that most of them are subjective in nature.

4.4.2 Persuasive or Directive Tweets

The fact that the tweets are written as instructions shows that the writer is attempting to persuade the reader to act in a manner consistent with the proposition's content. For instance, questioning, interrogating, commanding, begging, persuading, inviting, insisting, and confronting. The tweets shown in the table have been randomly selected from the dataset based on the keywords found by all three algorithms.

Table 4.10: Assertive Persuasive or Directive Content.

S. No	Tweet Content	Persuasive
1	How do you tell the difference between the flu and covid 19?	Doubt, questioning, interrogating
2	"India tightens travel curbs. #ITVideo #COVID2019 #COVID #CoronavirusOutbreak #coronavirus"	Commanding
3	Can't claim VIP status, avoid COVID-19 test"": West Bengal Chief Minister Mamata Banerjee slams teen	Confronting
4	Why should we trust health officials and doctors who promise that the covid vaccines are safe? . They also promised that “breakthrough” infections would be exceedingly rare.	Confronting, Doubting
5	Another great article by the great David Tuller. The medical establishment gaslights doctors, insisting long Covid is 'psychological' - Coda Story	Insisting
6	Together, we can beat COVID-19. Get vaccinated and encourage others to do the same. Encouraging.	Persuading
7	Is the fear of COVID-19 holding you back from living your life? Get vaccinated and regain your freedom!	Encouraging
8	It's time to get vaccinated! Protect yourself and those around you from COVID-19.	Persuading
9	Don't be a spreader, be a stopper! Get vaccinated against COVID-19 today.	Persuading
10	The science is clear: vaccines are safe and effective against COVID-19. Get vaccinated and protect yourself and those around you.	Persuading
11	Want to get back to normal? The best way to do that is by getting vaccinated against COVID-19.	Persuading
12	Stop the spread of COVID-19 by wearing a mask,	Persuading

	practicing social distancing, and getting vaccinated.	
13	We're all in this together! Help protect your community by getting vaccinated against COVID-19.	Persuading
14	Don't let COVID-19 win. Get vaccinated, wear a mask, and do your part to keep your community safe.	Persuading
15	If you're eligible for the COVID-19 vaccine, don't wait! Schedule your appointment today and protect yourself and those around you.	Persuading
16	COVID-19 is not going away anytime soon. Protect yourself and those around you by getting vaccinated and continuing to practice safety measures.	Persuading

When trying to convince individuals with brief communications, it is clear that the way you say something is just as crucial as the words themselves. The majority of the posts that have been made in this area are centred on the idea of delivering a clear signal to follow, comply, accept the guideline or mandate, and a variety of other metaphorical meanings. The quantity of messages of this kind that are posted on Twitter on a daily basis in significant quantities for the sole purpose of convincing various parties. Words such as "kindly",”please,” “pls,” “plz” and, of course, “retweet” are common and clear show the intent of the user to persuade others.

4.4.3 Commissive Tweets

The users of the Twitter platform become "commissive, “committing themselves to a future course of action through the use of promises, pledges, vows, undertakings, and desires.

Table 4.11: Commissives Tweet Content.

S. No	Tweet Content	Commissives
1	Global health organisations are considering changing their Covid-19 vaccination pledges.	Pledges
2	Vaccine rebels vow to paralyse Italy over incoming COVID passport	Vows, Pledges
3	This #festive season, let us vow to follow safety guidelines and continue to practice regular hand washing.	Vows
4	#Singapore sends oxygen cylinders to support India's fight against	Promises

	#COVID19, South Korea promises help too.	
5	Established respiratory drugs could find new applications as India faces #COVID19 surge, by @scripanjug for @PharmaScrip . Budesonide Shows COVID-19 Promise: Should It Be Evaluated In India?	Promises
6	I vow to continue to wear a mask and maintain social distance until we are completely rid of #COVID19.	Vow
7	#India has promised to donate COVID-19 vaccines to 6 neighboring countries - Bhutan, Bangladesh, Nepal, Maldives, Sri Lanka and Myanmar.	Promises
8	Established respiratory drugs could find new applications as India faces #COVID19 surge, by @scripanjug for @PharmaScrip . Budesonide Shows COVID-19 Promise: Should It Be Evaluated In India?	Promises
9	#Pakistan Prime Minister Imran Khan vows to not impose a lockdown in his country again despite the surge in COVID-19 cases.	Vows
10	This #festive season, let us vow to follow safety guidelines and continue to practice regular hand washing.	Vows
11	Vaccine rebels vow to paralyse Italy over incoming COVID passport	Pledges, Vows

There are several circumstances in which a commitment is made, such as a promise or threat in the context of COVID, while the Twitter user is writing. Giving a warning or counsel with threats can also be classified as this kind of act.

4.4.4 Expressive Tweets

in which the author communicates an attitude toward or about a situation, for example, apologising, appreciating, celebrating, congratulating, thanking, welcoming, scolding, expressing wrath, or expressing distrust.

Table 4.12: Expressive Tweet Content.

S. No	Tweet	Expression
1	Two years in and companies are still apologising for "unprecedented demand and call volumes". Hire more staff, stop blaming COVID.	Apologising
2	He apologises for walking into his own garden,	Apologising

	effectively. To be honest, I am not apologising for going into my garden during COVID. Haven’t seen anyone else apologise for it, either.	
3	Some of these morons are walking around with these restrictions finally dropping, congratulating themselves like "We beat COVID." We did it. "	Congratulating
4	Please join us in congratulating and celebrating our very own Dr. Jacquelyn Minter, Director of Fort Bend County Health & Human Services in recognition of her tireless efforts and service to our community during the COVID-19 Pandemic.	Appreciating
5	We welcome the recent announcements by #RBI as they are directed towards infusing liquidity and strengthening consumption, thereby giving a push to economic recovery. #economy #india #realestate #liquidity #covid19 #lendingrate #ReserveBankOfIndia	Welcome
6	The COVID-19 pandemic has affected the education sector severely. It's time for us to come up with innovative solutions to ensure that learning doesn't stop.	Urging
7	It's heartening to see people coming together to help each other during these trying times. Let's continue to spread positivity and hope.	Appreciating
8	The pandemic has taught us the importance of being self-sufficient and self-reliant. Let's take this as an opportunity to build a better and more resilient world.	Reflecting
9	With the second wave of COVID-19 hitting us hard, it's crucial to follow all the safety protocols and be responsible citizens. Stay safe, everyone.	Reflecting
10	It's appalling to see people hoarding essential items during the pandemic and not thinking about others. We need to show empathy and kindness now more than ever.	Criticising

11	Congratulations to the researchers and scientists who worked tirelessly to develop the COVID-19 vaccine. Your contribution to humanity is invaluable.	Appreciating
12	With the COVID-19 situation under control, it's time for us to start planning our next vacation. The mountains are calling!	Planning
13	It's high time we stop blaming COVID for everything. People need to take responsibility for their actions and stop making excuses.	Criticising
14	The COVID-19 vaccination drive has been one of the most significant steps taken by the Indian government in recent times. Kudos to the healthcare workers and the administration for making this possible.	Appreciating

There are thoughts of examples in which the twitter user is expressing happier moments in which he/she is simply congratulating or appreciating others. Happier moments shared through short messaging acts of appreciation and welcome are part of this category of the illocutionary act.

4.4.5 Declarative tweets

In declarative tweets, the author only affects an object's or situation's outward state or condition with the utterance: for example, "I pronounce you husband and wife;" "I sentence you to be hanged by the neck."

Table 4.13: Declaratives Tweet Content.

S. No	Tweet Content	Declaratives
1	Amidst COVID19 – Time to declare India’s Health Sector as Critical Sector for Cyber Security – Diplomatist #NAMA	Declaration
2	Announcements will be made on aeroplanes, ships, metros, railway stations to mark the moment when India achieved target of administering 100 Cr #COVID19 vaccine doses: Union Health Minister	Announcements, Declare

	Mansukh Mandaviya	
3	It's time India declare #COVID19 as National Excuse.	Declare , satirical statement
4	The #defence sector announcements by @FinMinIndia can prove beneficial in times wherein substantial stress is likely on the anvil due to #COVID19 lockdowns in India and globally, according to @SumitSinghania_ , Subject Matter Expert, @DeloitteIndia .	Declare, announcements
5	Repeat after me: The government should declare all journalists' frontline workers. Immediately. We've lost way too many people, and this dance of death is not stopping soon. #Covid19 #India	Appeal, Declare
6	PM Modi: I announce a special economic package today. This will play an important role in the #AtmanirbharBharatAbhiyan. The announcements made by the govt over COVID19, decisions of RBI & today's package totals to Rs 20 Lakh Crores. This is 10% of India's GDP	Declare
7	Amidst COVID19 – Time to declare India's Health Sector as Critical Sector for Cyber Security – Diplomatist #NAMA	Declare
8	If you're feeling anxious or stressed about #COVID19, it's time to declare self-care as a top priority. Take breaks, stay connected with loved ones, and prioritize your mental health.	Declare, Advice
9	As #COVID19 cases continue to rise, it's high time to declare a state of emergency and take strict measures to control the spread of the virus. #India	Declare
10	#Delhi's coronavirus positivity rate has risen from 0.22% in the last week of June to 0.27% in the first week of July. It's time to declare a health emergency in the national capital. #COVID19	Declare

11	With #COVID19 causing travel restrictions & a steep rise in online payments, it's time to declare eKYC & video KYC mandatory for all businesses. Know your customer, protect your business.	Declare, Call to Action
12	PM Modi: I announce a special economic package today. This will play an important role in the #AtmanirbharBharatAbhiyan. The announcements made by the government over COVID19, decisions of RBI & today's package totals to Rs 20 Lakh Crores. This is 10% of India's GDP.	Declare
13	The #defence sector announcements by @FinMinIndia can prove beneficial in times wherein substantial stress is likely on the anvil due to #COVID19 lockdowns in India and globally, according to @SumitSinghania_, Subject Matter Expert, @DeloitteIndia.	Declare, Announcements
15	Announcements will be made on aeroplanes, ships, metros, railway stations to mark the moment when India achieved target of administering 100 Cr #COVID19 vaccine doses: Union Health Minister Mansukh Mandaviya	Announcements, Declare

It is fairly obvious, as can be seen from all of the tables that were presented earlier, that speech act theory is highly helpful in comprehending the goals that were pursued by the content writer. The findings of this research provide credence to the work that Austin and Searle accomplished within the framework of speech act theory. It is important to note, however, that whereas Austin placed more emphasis on the traditional interpretation of speech actions, Searle placed more emphasis on the psychological interpretation of speech acts (based on beliefs, intentions, etc.). Both psychological and straightforward analysis are possible from the same content analysis.

4.5 Section IV: Comparative Analysis

In this section, we shall expound upon a thorough comparative study of the models that have been crafted during our research endeavors. Specifically, we shall delve into the equation-based digital trust model as well as the rules-based digital society interaction and trust model. Using this in-depth analysis, our ultimate objective is to assess the respective strengths and drawbacks of each of these models. The equation- based model is built upon complex mathematical equations, whereas the rules-based model is founded on a framework of social, demographic, and psychological rules. Through a rigorous comparative analysis of these two models, we shall strive to discern the advantages and limitations of each, thereby affording us a more nuanced understanding of their overall efficacy.

Table 4.14: Comparison (Recall, Precision and Accuracy) of Model-II Performance (Advance Model)

Iteration	Recall	Precision	Accuracy
1	1.0	0.9133	0.99
2	0.76	1.0	0.91
3	0.75	1.0	0.91
4	0.87	1.0	0.91
5	0.87	1.0	0.91
6	0.77	1.0	0.91
7	0.89	1.0	0.91
8	0.88	1.0	0.91

We may examine the recall, precision, and accuracy over iterations for deeper analysis. The analysis is as follows:

As depicted in Table 4.14, Model-II has demonstrated strong performance across all iterations, with high values in recall, precision, and accuracy. The best performance was in the 1st iteration, with a perfect recall, precision, and accuracy of 1.0, 0.9133, and 0.99, respectively. The worst performance in terms of recall was in the 2nd and 3rd iterations, both with a recall of 0.76, although the precision was perfect (1.0) and accuracy was still high (0.91).

The model has been consistent in maintaining a perfect precision of 1.0 from the 2nd iteration onwards. Accuracy has also been consistent at 0.91 from the 2nd iteration onwards, except for the 1st iteration where it was slightly higher at 0.99. Recall has varied across iterations, with the lowest values in the 2nd and 3rd iterations (0.76) and the highest value in the 1st iteration (1.0).

There is a general trend that as recall decreases, precision increases, and vice versa. This is evident in the 1st iteration, where recall is perfect (1.0) but precision is slightly lower (0.9133) compared to other iterations. Accuracy remains high across all iterations, indicating that the model can correctly classify a high percentage of instances.

In conclusion, Model-II has shown strong and consistent performance across all iterations, with perfect precision from the 2nd iteration onwards and high accuracy. Recall has varied across iterations, with the lowest values in the 2nd and 3rd iterations. The relationship between precision and recall is inversely proportional, with one increasing as the other decreases.

Table 4.15: Comparison (MSE, R^2 and RMSE) of Model-II Performance

Model II: Advance Model			
Iteration	MAE	R^2	RMSE
1	0.131	0.611	0.148
2	0.135	0.632	0.170
3	0.220	0.632	0.139
4	0.125	0.771	0.151
5	0.098	0.776	0.133
6	0.085	0.773	0.128
7	0.137	0.585	0.111
8	0.152	0.513	0.189

From table 4.15, it can be observed the model II performance improved over iterations, with the R^2 value increasing from 0.611 in the 1st iteration to 0.776 in the

5th iteration. This indicates that the model was able to explain more variance in the data as it evolved. Following are additional inferences that can be made.

Best and Worst Performance:

- The best performance of the model was in the 5th iteration, with the lowest MAE (0.098), the highest R^2 (0.776), and a low RMSE (0.133).
- The worst performance was in the 8th iteration, with the highest MAE (0.152), the lowest R^2 (0.513), and the highest RMSE (0.189).

Consistency in Performance:

- The model was consistent in terms of R^2 values from the 2nd to the 3rd iteration (0.632) and from the 5th to the 6th iteration (0.773 to 0.776).
- However, there was a significant drop in R^2 value in the 7th iteration (0.585) and the 8th iteration (0.513) compared to the 6th iteration (0.773).

MAE and RMSE Relationship:

- Generally, as MAE increases, RMSE also increases, indicating that as the average error increases, the root mean square error also increases. This is evident in the 8th iteration, where both MAE (0.152) and RMSE (0.189) are the highest.

R^2 and RMSE Relationship:

- There is a general trend that as R^2 increases, RMSE decreases, indicating that as the model explains more variance, the root mean square error decreases. This is evident in the 5th iteration, where R^2 is the highest (0.776) and RMSE is one of the lowest (0.133).

In nutshell, the model II showed improvement over iterations, with the best performance in the 5th iteration. However, there was a significant drop in performance in the 7th and 8th iterations, which needs to be investigated further.

Table 4.16: Comparison (Recall, Precision and Accuracy) of Model-I Performance

Model I: Equation based			
Iteration	**Precision**	**Recall**	**Accuracy**
1	1	1	0.99
2	1	1	0.99
3	0.55	0.48	0.59
4	0.56	0.57	0.51
5	0.51	0.59	0.52
6	0.52	0.59	0.54
7	0.50	0.58	0.49
8	0.50	0.57	0.48

From Table 4.16 (Model-I), it can be observed that overall, the model's performance significantly dropped after the 2nd iteration. The first two iterations had perfect precision and recall, with accuracy very close to 1 (0.99). Additionally, following inferences can be made.

Best and Worst Performance:

- The best performance of the model was in the 1st and 2nd iterations, with perfect precision and recall, and accuracy of 0.99.
- The worst performance was in the 8th iteration, with the lowest accuracy (0.48), and precision and recall values of 0.50 and 0.57, respectively.

Consistency in Performance:

- The model was consistent in its performance in the 1st and 2nd iterations, with no variation in precision, recall, and almost no variation in accuracy.
- From the 3rd iteration onwards, the model's performance varied, but there was no significant improvement or further decline in performance.

Precision, Recall, and Accuracy Relationship:

- Generally, as precision and recall values decrease, accuracy also decreases. This is evident in the 3rd iteration, where precision, recall, and accuracy all dropped significantly compared to the 2nd iteration.
- However, there are instances where precision and recall do not have a direct relationship with accuracy. For example, in the 4th iteration, precision increased slightly, recall remained the same, but accuracy decreased compared to the 3rd iteration.

Lastly, it can be said that overall, the model I showed a significant drop in performance after the 2nd iteration, with no significant improvement in the subsequent iterations. The relationship between precision, recall, and accuracy is generally direct, but there are instances where they do not correlate directly. Further investigation is needed to understand the reasons behind the drop in performance and lack of improvement in the subsequent iterations. However, when comparing Model-I and Model-II, it can be clearly observed that in terms of Initial Performance, the Model-I had a perfect start with precision and recall values of 1, and accuracy of 0.99 in the first two iterations. Model-II had a good start but not as perfect as Model-I, with R^2 values ranging from 0.611 to 0.632 in the first three iterations. When we check the consistency, the Model-I was consistent in the first two iterations but showed a significant drop in performance from the 3rd iteration onwards. Model-II showed consistency in R^2 values in the 2nd and 3rd iterations (0.632) and 5th and 6th iterations (0.773 to 0.776). When we attempted to find the best performer, Model-I's best performance was in the 1st and 2nd iterations. Model-II's best performance was in the 5th iteration, with the lowest MAE (0.098), the highest R^2 (0.776), and a low RMSE (0.133). Observations on worst Performance include Model-I's worst performance was in the 8th iteration, with the lowest accuracy (0.48). Model-II's worst performance was in the 8th iteration, with the highest MAE (0.152), the lowest R^2 (0.513), and the highest RMSE (0.189). In the end it can be said that the overall trend consists of the following ingredients.

- Model-I started perfectly but showed a significant drop in performance with no improvement in subsequent iterations.

- Model-II showed improvement over iterations but had a significant drop in performance in the 7th and 8th iterations.

In gist, it can be said that while Model-I had a perfect start; its performance dropped significantly and did not improve in subsequent iterations. On the other hand, Model-II showed improvement over iterations but had a significant drop in performance in the later iterations.

Table 4.17: Comparative Analysis of Model-I & Model-II

Metric	Model I (Equation Based)	Model II (Advance Model)
Recall	0.50-1.0	0.75-1.0
Precision	0.50-1.0	0.913-1.0
Accuracy	0.48-0.99	0.91-0.99
R^2	0.48-1.0	0.513-0.776
MAE	0.55-1.0	0.085-0.220
RMSE	0.48-0.99	0.111-0.189
Numerical Stability	Moderate	Good
Flexibility	Limited	High
Adaptability	Limited	High
Interpretability	Moderate	Moderate
Complexity	Low	Moderate

When comparing Model-I (Equation Based) and Model-II (Advance Model) across various metrics, we can draw the following conclusions:

Recall, Precision, and Accuracy:

- Model-I has a range of 0.50 to 1.0 for recall, 0.50 to 1.0 for precision, and 0.48 to 0.99 for accuracy. This indicates that Model-I has the potential to perform well, but there are instances where its performance drops significantly.
- Model-II, on the other hand, has a range of 0.75 to 1.0 for recall, 0.913 to 1.0 for precision, and 0.91 to 0.99 for accuracy. This indicates that Model-II is performing better than Model- 1.

R^2, MAE, and RMSE:

- Model-I has a range of 0.48 to 1.0 for R^2, 0.55 to 1.0 for MAE, and 0.48 to 0.99 for RMSE. This indicates that Model-I has a wide range of performance across different iterations.
- Model-II has a range of 0.513 to 0.776 for R^2, 0.085 to 0.220 for MAE, and 0.111 to 0.189 for RMSE. This indicates that Model-II has a more consistent performance across different iterations, with generally lower error rates and a moderate ability to explain variance in the data.

Numerical Stability, Flexibility, Adaptability, Interpretability, and Complexity:

- Model-I has moderate numerical stability, limited flexibility and adaptability, moderate interpretability, and low complexity. This indicates that while Model-I is easy to understand and work with, it may not be suitable for complex or changing environments.
- Model-II has good numerical stability, high flexibility and adaptability, moderate interpretability, and moderate complexity. This indicates that Model-II is more robust and suitable for complex and changing environments but may require more effort to understand and work with.

Further, in this section a comparative analysis is done with the state of artwork that are closely comparable and relatable with our research work. Therefore, the purpose of this comparative analysis is to synthesize the findings from three recent research papers that investigate various aspects of trust, social media engagement, and social relationships. By examining the methodologies, results, and implications of these studies. The aim is to identify common themes, gaps, and areas for future research. This analysis will provide a comprehensive overview of the current state of knowledge in this field and highlight the potential theoretical and practical implications of the findings.

The three research papers selected for this comparative analysis are:

- **My work:** My work compares the performance of two models, Model I (Equation Based) and Model II (Advance Model), using metrics such as recall, precision, accuracy, R^2, MAE, and RMSE.
- **Research by Przemysław *et al.*:** This study investigates the factors influencing the information verification behavior of internet users using structural equation modelling (SEM). The paper analyses data collected from a sample of 245 Polish Facebook users and examines the relationships between variables such as social ties diversity, fake news awareness, and information verification.
- **Research by Islam Habis *et al* :** This paper explores the moderating effect of trust on the relationship between social media engagement, relationship benefits, and social relationships using SEM. The study uses data from a sample of 493 Jordanian youth and investigates the impact of social media engagement and relationship benefits on social relationships, with trust as a moderating variable.

By comparing these three-research works, an attempt will be made to seek a deeper understanding of the complex interplay between trust, social media engagement, and social relationships in different contexts and population groups.

Following is the brief description information on each work:

My research work: The study employs a comparative analysis approach to evaluate the performance of two models; Model I (Equation Based) and Model II (Advance Model) The research work also covers a simulation work. The work presented here demonstrates the use of various metrics such as recall, precision, accuracy, R^2, MAE, and RMSE to assess and compare the models' performance across different iterations. The specific data collection methods, sample size, and data analysis techniques are not mentioned in the excerpt.

Przemysław *et al.* [70]: The "Przemysław Majerczak" study employs a quantitative research methodology using structural equation modelling (SEM). The researchers collected data through an online questionnaire distributed to a sample of 245 Polish Facebook users. The questionnaire likely included items measuring various constructs such as social ties diversity, fake news awareness, social media credibility, trust in people online, information verification, and intention to share. The study utilised the partial least squares (PLS) method to analyse the data and test the hypothesised relationships between the variables. The researchers assessed the reliability and validity of the measurement model and evaluated the structural model to examine the significance and strength of the relationships between the constructs.

Islam Habis *et al.* [71]: The "Islam Habis et al" study employs a quantitative research methodology using structural equation modelling (SEM). The study collected data through a questionnaire distributed to a sample of 493 Jordanian youth. The questionnaire likely included items measuring constructs such as social media engagement, relationship benefits (psychological, social, and functional), trust, and social relationships. The researchers utilized the partial least squares (PLS) method to analyze the data and test the hypothesized relationships between the variables. The study assessed the measurement model's reliability and validity using composite reliability, average variance extracted (AVE), and discriminant validity tests. The structural model was evaluated to examine the significance and strength of the relationships between the constructs, with a focus on the moderating effect of trust on the relationship between social media engagement, relationship benefits, and social relationships. The tabular summary of the numerical results in these research works is as follows in Table no. 4.18

Table 4.18: Comparative Analysis of my work with State of the Art

Research work	Objectives	Methodology	Key Metrics/Results
My Work	Compare the performance of Model I (Equation Based) and Model II (Advance Model)	Equations Modelling (SEM) and custom rules	Recall Range: Model I (0.50-1.0), Model II (0.75-1.0) Precision Range: Model I (0.50-1.0), Model II (0.9133-1.0) Accuracy Range: Model I (0.48-0.99), Model II (0.91-0.99) R^2 Range: Model I (0.48-1.0), Model II (0.513-0.776) MAE Range: Model I (0.55-1.0), Model II (0.085-0.220) RMSE Range: Model I (0.48-0.99), Model II (0.111-0.189)
Przemysław *et al.* [70]	Investigate factors influencing information verification behaviour of internet users using structural equation modelling (SEM)	Quantitative analysis with SEM, using a questionnaire distributed to a sample of 245 Polish Facebook users	Path Coefficients, T-Statistics, p-Values, R^2, Q2 Strongest Relationships: Social ties diversity & fake news awareness (0.270), Fake news awareness & information verification (0.267)>R^2 Range: 0.028-0.262 All hypotheses (H1-H6) supported

Islam Habis *et al*: [71]	Explore the moderating effect of trust on the relationship between social media engagement, relationship benefits, and social relationships using SEM	Quantitative analysis with SEM, using a questionnaire distributed to a sample of 493 Jordanian youth	Original Sample, Sample Mean, Standard Deviation, T-Statistics, p-Values, F-Square, R^2, Q2 H1, H2, H3 supported; H4 rejected Strongest Relationships: Social media engagement & social relationship (0.604), Relationship benefits & social relationship (0.304) Trust moderates the relationship between social media engagement and social relationship (0.060) R^2 (Social Relationship): 0.917, Q2 (Social Relationship): 0.876

It must be noted that while the three-research works differ in their specific objectives, variables, and methodologies, they all employ quantitative analysis techniques to investigate relationships between various constructs related to social media, trust, and social relationships. Clearly, our research work focuses on comparing the performance of two models using metrics such as recall, precision, accuracy, R^2, MAE, and RMSE. The Przemysławe et al. [70] study examines factors influencing information verification behaviour using SEM, reporting path coefficients, t-statistics, p-values, R^2, and Q2. The "Imran Hab " paper explores the moderating effect of trust on social media engagement, relationship benefits, and social relationships using SEM, reporting similar metrics as Przemysławe et al. [70] along with F-square values. Given the differences in the papers' objectives and variables, a direct comparison of numerical results is not feasible. However, all three papers contribute to the understanding of social media, trust, and social relationships using quantitative analysis techniques, albeit from different perspectives and with varying focal points.

However, it can be observed that all these three-research works have one common metric which is R^2, and the strongest relationships metric. Hence, following is the comparative analysis based on these metrics.

Table 4.19: Comparative Analysis of Results

Study	R^2 Range/Value	Strongest Relationships
My Work	Model I: 0.48-1.0 Model II: 0.513-0.776	Frequency of contact, demographics and origin of place,
Przemysław *et al.* [70]	0.028-0.262	Social ties diversity & fake news awareness (0.270). Fake news awareness & information verification (0.267)
Islam Habis *et al*: [71]	Social Relationship: 0.917	Social media engagement & social relationship (0.604). Relationship benefits & social relationship (0.304)

It can be observed from this study that builds upon and diverges from existing state- of-the-art research in digital trust modeling by introducing unique methodologies, leveraging a pandemic-specific dataset, and incorporating advanced metrics tailored to capture complex trust dynamics. Consider these points as well in the nutshell for more clarity.

1. Alignment with Existing Research

- **Modeling Approach and Structure**: Like prior studies in digital trust modeling, this research employs foundational elements of sentiment analysis, social interaction data, and demographic factors to predict trust levels. Studies using linear and equation-based models to establish correlations between demographic characteristics and trust are directly reflected in the E.D.S.T.M. used here.

- **Use of Quantitative Metrics**: In line with traditional approaches, this research incorporates quantitative evaluation metrics, such as MAE, RMSE, and R^2, to assess model performance. These metrics are commonly used in state-of-the-art trust models to quantify prediction accuracy and provide a standardized measure of model fit.

2. Divergence from State-of-the-Art Studies

- **Dataset Specificity – COVID-19 Context**: Unlike most existing trust models, which often rely on general social datasets, this study uses a COVID-19-specific dataset. This pandemic-focused data captures unique social behaviors, trust shifts, and emotional dynamics that emerged during a time of global crisis. The specificity of this dataset enables the models to account for pandemic-specific trust trends, such as responses to misinformation, health- related anxieties, and community support during lockdowns.
- **Non-Linear and Event-Sensitive Modeling (A.D.S.I.T.M.)**: While many state-of-the-art models employ linear relationships between variables, this study's A.D.S.I.T.M. introduces non-linear dynamics and event-sensitive mechanisms to capture complex trust behaviors. The integration of non-linear modeling allows A.D.S.I.T.M. to simulate how trust fluctuates with Mediocristan (regular) and Extremistan (extreme) events, providing insights that traditional linear models cannot.

3. Advancements beyond Existing Studies

- **Incorporation of Sociological and Emotional Variables**: This study advances beyond standard trust modeling by incorporating a broader range of variables related to cognitive perceptions, emotional states (fear, anxiety, hope), and social group affiliations. These additions reflect a more comprehensive approach, allowing the models to capture not only explicit social interactions but also implicit psychological and emotional influences on trust.

- **Hybrid Model Design (E.D.S.T.M. and A.D.S.I.T.M.)**: By combining the simplicity of E.D.S.T.M. with the complexity of A.D.S.I.T.M., this study presents a hybrid approach that adapts to both simple and complex trust dynamics. This hybrid model design allows for scalability, from quick assessments of trust in stable scenarios to in-depth analyses of trust during crises.
- **Enhanced Evaluation Framework**: The study uses a refined set of evaluation metrics, with MAE, RMSE, and R^2 providing traditional performance measures, while precision, recall, and F1 Score address predictive specificity. This expanded evaluation framework goes beyond what most state-of-the-art studies offer, delivering a more complete assessment of model accuracy, reliability, and explanatory power in the context of digital trust.

4. Addressing Gaps in State-of-the-Art Research

This study aims to address several specific gaps in existing digital trust modeling research:

- **Gap 1: Lack of Context-Sensitive Trust Modeling**: Previous studies often overlook the impact of large-scale external events, such as pandemics, on digital trust dynamics. This research bridges this gap by using pandemic- specific data to develop a model that adapts to significant social disruptions and rapidly changing trust conditions.
- **Gap 2: Limited Emotional and Psychological Dimensions**: Traditional trust models tend to focus on observable social and demographic data, leaving out deeper psychological and emotional factors. By integrating variables like emotional states and cognitive perceptions, this study provides a model that can capture the nuanced effects of emotional responses on trust.
- **Gap 3: Inflexibility in Handling Complex Social Interactions**: Existing models are often designed for straightforward, stable social scenarios, and may struggle with the complexities introduced by non-linear trust dynamics.

A.D.S.I.T.M. addresses this by including non-linear relationships and feedback mechanisms that allow it to capture both subtle and extreme shifts in trust.

Through this comparison, the research demonstrates its unique contributions, addressing critical gaps in existing studies and providing a more adaptable framework for digital trust modeling, especially under crisis conditions.

4.6 Summary of Result Chapter:

This chapter summarises our research findings, including trust models, twitter data collection and analysis, and content analysis. Our comparative examination at each level strengthens our conclusions, and our research advances digital trust and sociology. This chapter opens the door to digital society trust dynamics research.

Chapter - 5

CONCLUSIONS AND INFERENCES

The research on emotional analysis and faith in the context of the COVID-19 pandemic situation utilizing data from Twitter, we intended to analyse and comprehend the feelings and thoughts associated with the COVID-19 pandemic, as well as study the function that trust plays in this context. Following are some findings, deductions, and inferences that can be reached based on an examination of tweets collected from Twitter using techniques such as sentiment analysis and trust modelling:

1. The sentiment analysis of tweets related to COVID-19 revealed that people on Twitter have been expressing predominantly negative emotions, such as fear, anxiety, and frustration, during the pandemic.

2. The trust analysis of tweets related to COVID-19 revealed that people on Twitter have been expressing relatively low levels of trust in various sources of information, including government agencies, media outlets, and healthcare providers.

3. The sentiment and trust analyses of tweets related to COVID-19 revealed that there is a significant relationship between trust and sentiment. Specifically, tweets expressing negative sentiment tend to express lower levels of trust, while tweets expressing positive sentiment tend to express higher levels of trust.

4. The speech act theory was used to label tweets based on the underlying communicative intent, which provided additional insights into the nature of the conversations related to COVID-19 on Twitter.

5. The study demonstrates the potential of sentiment analysis and trust modelling techniques to provide valuable insights into public opinion and attitudes towards important societal issues such as the COVID-19 pandemic.

6. The Advance Digital Society Interaction and Trust Model (A.D.S.I.T.M) demonstrates superior recall (0.75-1.0) compared to the Equation-based Digital Social Trust Model (E, D.S.T.M) (0.50-1.0), indicating its stronger ability to accurately detect positive instances.

7. The Advance Digital Society Interaction and Trust Model (A.D.S.I.T.M) shows better precision ranges (0.913-1.0) than Equation-based Digital Social Trust Model (E.D.S.T.M) (0.50-1.0)

8. The A.D.S.I.T.M consistently exhibits higher R^2 values (0.513-0.776) throughout all iterations, indicating a better fit and performance compared to the E.D.S.T.M (0.48-1.0).

9. The A.D.S.I.T.M demonstrates better precision in predicting the target variable with consistently lower RMSE values (0.11-0.189) compared to the E.D.S.T.M (0.48-0.99).

10. In summary, the A.D.S.I.T.M outperforms the E.D.S.T.M in terms of recall, accuracy, R^2, and RMSE. Both models show a moderate level of interpretability, but the A.D.S.I.T.M exhibits higher numerical stability, flexibility, adaptability, and complexity compared to the E.D.S.T.M.

11. The keywords obtained through NMFF, PLSI, and LDA algorithms were used to extract relevant tweets from the collected corpus, leading to a subsequent content analysis of their content, purpose, and significance.

12. The tweets were categorized based on speech act theory, demonstrating the primary focus of this study on content analysis.

13. The persuasive nature of concise communication on Twitter is evident, with a predominant theme of transmitting clear signals to influence others, as indicated by the frequent use of words like "kindly," "please," "pls," "plz," and "retweet."

14. The study supports the contributions of Austin and Searle to speech act theory, highlighting their differing interpretations of speech actions, with Austin emphasizing conventional interpretation and Searle focusing on psychological interpretation.

15. The analysis reveals a significant frequency of posts related to COVID-19 lockdown, with individuals disseminating information, asserting scientific claims, discussing conspiracy theories, and providing advisory statements. Twitter users employ anecdotal evidence and cite authoritative sources to support their claims, contributing to the overall persuasive nature of the tweets.

16. In the context of Guided Content Analysis, in summary, the analysis of tweets based on speech act theory provides insights into the objectives pursued by content writers, the prevalence of persuasive language, and the diverse range of topics discussed, particularly related to COVID-19 and the associated lockdown measures.

Overall, the research sheds light on how important it is to comprehend the feelings and reactions that are associated with the COVID-19 epidemic, as well as the significance that trust plays in determining public opinion and behavior during times of emergency. The research also sheds light on how important it is to comprehend the feelings and reactions that are associated with the COVID-19 epidemic.

In the event that there is an emergency in the future, those in charge of making decisions, those who provide medical care, and any other parties who have a stake in the matter can use the findings of this study to devise effective methods for disseminating information to the general public and building trust among individuals who are a part of the general population. Because the data for this study come exclusively from Twitter, it is possible that the conclusions do not fully reflect the attitudes of the general community concerning the COVID-19 outbreak. However, it is important to note that this study relies solely on data from Twitter. For the purposes of this investigation, only the Hinglish language will be taken into consideration. Because of this, there is a possibility that it does not take into account the sentiments and perspectives that are expressed in other languages. Because the only sentiments that are taken into consideration in this study are the ones that are stated in tweets, it is possible that the results do not correctly reflect how users genuinely feel about the pandemic. Hence, in terms of directions for the future, it is conceivable to expand the scope of the study to include additional social media platforms to get a more in-depth knowledge of the ways in which people feel about the epidemic. This would be accomplished by expanding the scope of the research.

Additionally, to achieve a more all-encompassing comprehension of the circumstance on a global scale, the parameters of the study might be broadened to include sentiments expressed in languages other than English. This would allow for the acquisition of a more complete understanding of the situation.

For future scope, the study can be expanded to include the sentiment analysis of news items, blogs, and forums as a means of obtaining a more thorough picture of people's attitudes towards the pandemic. This would be done to obtain a better understanding of how people feel about the pandemic. This work has the potential to lay the groundwork for later research on trust and sentiment analysis, which, in turn, has the potential to contribute to the development of effective approaches for the management of pandemics.

5.1 Social and Technological Impact of Study

1) The ability to trust the information and instructions we are given is crucial to the success of our pandemic preparation and response efforts. The purpose of this project is to investigate several approaches to determining how to assess levels of trust in a variety of settings. By contrasting the various approaches, we will be able to identify a more reliable approach to trust assessment, which may then be used to enhance the effectiveness of responses to future pandemics.

2) The research that is being done in this field will be able to improve the social-communication features of the users by calculating the users' levels of sentiment and trust. The aim of the user can be deduced from these several

characteristics, which are key indicators. Because of the accuracy with which these two indicators may forecast the possibility of future events, we can utilize this information to deal with problems by either altering the scenario that is occurring now or altering the behavior of both individuals and groups in advance.

3) The study investigates the trustworthiness and believability of information discovered in a crisis situation, which is characterized by a lack of specificity and clarity. It examines how people feel the government will cope with that scenario and what actions they anticipate seeing from the government as a result of those beliefs. In addition to this, it examines how quickly they turn to the government for information on what is going on, what they are willing to offer for themselves, and how much responsibility they feel towards their family as a whole.

The results of this study will have an effect on a large number of individuals since they will provide a greater insight into the nature of how people react to a crisis. It will give the broader world a better understanding of how people think and engage during a crisis, as well as how to ensure that trust is developed among concerned persons, hence boosting digital communications on challenging subjects such as pandemics, particularly infectious diseases.

In summary, in this research work we have done extensive work to understand the dynamics of trust when a pandemic such as COVID was happening. For this we collected data from twitter (X) and following a structured approach in conducting analysis of existing methods we formulated a problem for solving the problem of modelling trust dynamics.

Following are the conclusions and inferences that can be drawn from this research work.

1) The comparison between the Equation-based Digital Social Trust Model (E.D.S.T.M.) and the Advanced Digital Society Interaction and Trust Model (A.D.S.I.T.M.) provides notable insights. Both models demonstrated valuable

predictive capabilities across multiple metrics, though differences in their performances are observed.

2) A.D.S.I.T.M. achieves a broader recall range (0.75–1.0) compared to E.D.S.T.M. (0.5–1.0), suggesting it is generally more effective in identifying true positive instances, especially under complex scenarios. However, this difference, while measurable, may not drastically impact certain practical applications. In terms of precision, both models perform similarly, ranging from 0.5–1.0, indicating that each model maintains an equivalent accuracy level when identifying positive results in a controlled environment.
3) When assessing accuracy, A.D.S.I.T.M. maintains a wider range (0.91–0.99) compared to E.D.S.T.M. (0.48–0.99), denoting a modest yet consistent advantage in overall predictive precision. The robustness of A.D.S.I.T.M. becomes particularly apparent in its R^2 values, which consistently range from 0.513 to 0.776 across iterations, compared to E.D.S.T.M.'s slightly fluctuating R^2 values, extending between 0.48 and 1.0. This steadier range suggests a reliability advantage for A.D.S.I.T.M. in varied contexts, though both models show strengths in predictive fit.
4) The error metrics further underscore A.D.S.I.T.M.'s advantages, as shown by a mean absolute error (MAE) range of 0.085 to 0.220, contrasted with E.D.S.T.M., whose MAE spans 0.55 to 1.0. This reflects greater consistency in A.D.S.I.T.M.'s predictions. In terms of root mean square error (RMSE), A.D.S.I.T.M. demonstrates significantly lower values (0.11–0.189) than E.D.S.T.M. (0.48–0.99), marking a clear reduction in prediction error variance. Such stability may suggest A.D.S.I.T.M.'s suitability for applications demanding high reliability and precision in predictive tasks.

5) While both models exhibit moderate numerical stability, A.D.S.I.T.M. shows a tangible resilience to input fluctuations, affirming its reliability in complex real-world datasets. Conversely, E.DSTM. shows moderate stability, yet its sensitivity to fluctuations suggests it may be better suited to less variable datasets. In terms of model adaptability, A.D.S.I.T.M. displays superior flexibility, effectively managing a wide array of scenarios while dynamically

adjusting to data variability, unlike E.D.S.T.M., which remains somewhat limited in adaptability.

6) Complexity emerges as a contrasting element: A.D.S.I.T.M., while more robust, presents a higher complexity level in comparison to the simpler E.D.S.T.M. model, which could be advantageous in contexts where interpretability and ease of deployment are prioritized over precision. Both models retain a moderate degree of interpretability, although E.D.S.T.M.'s simplicity may make it marginally more accessible for interpretative analysis.

7) Ultimately, the findings strongly suggest that A.D.S.I.T.M. generally outperforms E.D.S.T.M. across several key metrics—recall, R^2, MAE, RMSE, numerical stability, flexibility, and adaptability. However, both models hold unique advantages in specific scenarios: E.D.S.T.M.'s lower complexity and interpretability make it a viable option for straightforward applications, while A.D.S.I.T.M.'s superior accuracy and consistency position it as the preferred model for applications requiring high precision and robust adaptability.

Achievement of Research Objectives

In this study, the primary research objectives were systematically addressed through a combination of equation-based modelling, advanced machine learning, and comprehensive content analysis. The findings validate the achievement of each objective, underscoring the significance of trust modelling and sentiment analysis in multilingual, code-mixed digital interactions, particularly in the context of the COVID-19 pandemic.

1. **Objective 1**: *To study and compare contemporary models for trust and sentiment analysis.*

    - **Achievement**: The literature review and methodology sections provided an in-depth comparative analysis of existing models, highlighting their strengths and limitations. Through the comparative study of Equation-based Digital Social Trust Model (E.D.S.T.M.) and Advanced Digital Society Interaction and Trust Model (A.D.S.I.T.M.), this research contributed to a nuanced understanding of trust model capabilities across different metrics.

2. **Objective 2**: *To collect Hinglish text messages and analyze them using natural language processing (NLP) techniques.*

    - **Achievement**: Hinglish data was successfully collected from social media, primarily Twitter, and processed using NLP techniques. Sentiment scores were computed, and a custom lexicon for code-mixed language processing was developed, which enabled accurate sentiment analysis specific to Hinglish.

3. **Objective 3:** *To design and develop a novel model for social interactions, trust, and trustworthiness in digital societies.*

    - **Achievement**: The development of the A.D.S.I.T.M. model achieved this objective by incorporating variables relevant to digital trust, including emotional, psychological, and interactional factors. The model demonstrated improved recall and accuracy metrics, meeting the demands for robust trust analysis in digital environments.

4. **Objective 4**: *To compare and validate the proposed trust model with state-of- the-art approaches.*

    - **Achievement**: Through a detailed performance evaluation, the A.D.S.I.T.M. model was compared to E.D.S.T.M. and other models discussed in the literature. Evaluation metrics (e.g., RMSE, MAE, recall) confirmed that A.D.S.I.T.M. outperforms prior models in accuracy and adaptability, thereby validating its efficacy.

Table 5.1: Summary of Objectives and Achievements

Research Objective	Achievement
Objective 1: Study and compare contemporary models for trust and sentiment analysis.	Comparative analysis of E.D.S.T.M. and A.D.S.I.T.M. models provided insights into trust model effectiveness across metrics.
Objective 2: Collect Hinglish text messages and analyze them using NLP techniques.	Successful data collection and analysis; Hinglish sentiment scores computed with custom lexicon for code-mixed text.
Objective 3: Design and develop a novel model for social interactions, trust, and trustworthiness.	Developed A.D.S.I.T.M. model, integrating emotional, psychological, and interactional variables for robust trust modeling.
Objective 4: Compare and validate the proposed model with state-of-the-art models.	Performance comparison confirmed A.D.S.I.T.M. as a superior model with higher recall, accuracy, and adaptability.

Bibliography

[1] The importance of Digital Trust (2021) Deloitte Insights. Available at: https://www.deloitte.com/global/en/our-thinking/insights/topics/digital-transformation/the-importance-of-digital-trust-qa.html

[2] Guo, Y. (2022). Digital Trust and the Reconstruction of Trust in the Digital Society: An Integrated Model based on Trust Theory and Expectation Confirmation Theory. Digital Government, 3(4), 1–19. https://doi.org/10.1145/3543860

[3] Juárez, R. C. (2022, December 16). The seeds of a new relationship of trust in the digital society. Telefónica. https://www.telefonica.com/en/communication-room/the-seeds-of-a-new-relationship-of-trust-in-the-digital-society

[4] H. L. J. Ting, X. Kang, T. Li, H. Wang and C. -K. Chu, "On the Trust and Trust Modeling for the Future Fully Connected Digital World: A Comprehensive Study," in IEEE Access, vol. 9, pp. 106743-106783, 2021, doi: 10.1109/ACCESS.2021.3100767.

[5] S. M. Ghafari et al., "A Survey on Trust Prediction in Online Social Networks," in IEEE Access, vol. 8, pp. 144292-144309, 2020, doi: 10.1109/ACCESS. 2020.3009445..

[6] C. M. L. Wong and O. Jensen, "The paradox of trust: perceived risk and public compliance during the COVID-19 pandemic in Singapore," J. Risk Res., vol. 0, no. 0, pp. 1–10, 2020, doi: 10.1080/13669877.2020.1756386.

[7] Us, W. B. C. (2023b, August 25). Social Media Algorithms Warp How People Learn from Each Other. Scientific American. https://www.scientificamerican.com/article/social-media-algorithms-warp-how-people-learn-from-each-other/.

[8] Agarwal, V., Rao, P., & Jayagopi, D. B. (2021). Towards Code-Mixed Hinglish Dialogue Generation. In Proceedings of the 3rd Workshop on Natural Language Processing for Conversational AI. Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.nlp4convai-1.26

[9] Singh, Gaurav. (2021). Sentiment Analysis of Code-Mixed Social Media Text (Hinglish). arXiv:2102.12149

[10] Maryatt, E. (2018, November 1). Cracking code-mixing — an important step in making human-computer interaction more engaging - Microsoft Research. Microsoft Research. https://www.microsoft.com/en-us/research/blog/cracking-code-mixing-an-important-step-in-making-human-computer-interaction-more-engaging/

[11] Li, Shuyue & Murray, Kenton. (2022). Language Agnostic Code-Mixing Data Augmentation by Predicting Linguistic Patterns. 10.48550/arXiv.2211.07628.

[12] Chakravarthy, Sharanya & Umapathy, Anjana & Black, Alan. (2020). Detecting Entailment in Code-Mixed Hindi-English Conversations. 165-170. 10.18653/v1/2020.wnut-1.22.

[13] Thakur, V. S., Sahu, R., & Omer, S. (2020). Current state of Hinglish Text Sentiment Analysis. Social Science Research Network. https://doi.org/ 10.2139/ssrn.3614442

[14] Devine, D., Gaskell, J., Jennings, W., & Stoker, G. (2020). Trust and the Coronavirus Pandemic: What are the Consequences of and for Trust? An Early Review of the Literature. Political Studies Review, 19(2), 274–285. https://doi.org/10.1177/1478929920948684

[15] Reiersen, J., Roll, K. H., Williams, J. D., & Carlsson, M. C. (2022). Trust: a Double-Edged sword in combating the COVID-19 pandemic? Frontiers in Communication, 7. https://doi.org/10.3389/fcomm.2022.822302

[16] Graso M, Henwood A, Aquino K, Dolan P, Chen FX. The dark side of belief in Covid-19 scientists and scientific evidence. Pers Individ Dif. 2022 Jul;193:111594. doi: 10.1016/j.paid.2022.111594. Epub 2022 Mar 11. PMID: 35291670; PMCID: PMC8913370.

[17] Outbreak: 10 of the worst pandemics in History - MPH Online. (2021, August 31). MPH Online. https://www.mphonline.org/worst-pandemics-in-history/

[18] Arnold D. Pandemic India: Coronavirus and the Uses of History. J Asian Stud. 2020 Aug;79(3):569-577. doi: 10.1017/S0021911820002272. PMID: 34191871; PMCID: PMC7556896. [19] D. Huremović, “Brief History of Pandemics (Pandemics Throughout History),” in Psychiatry of Pandemics, 2019.

[19] Javaid, A. (2022, December 27). List of worst epidemics in India since the 1900s. Jagranjosh.com. https://www.jagranjosh.com/general-knowledge/history-of-epidemics-in-india-since-the-1900s-1584627562-1

[20] Wikipedia contributors. (n.d.-b). Category:Disease outbreaks in India - Wikipedia.https://en.wikipedia.org/wiki/Category:Disease_outbreaks_in_India

[21] M, P. K., & M, P. K. (2020). History of infectious disease outbreaks in India. FACTLY. https://factly.in/history-of-infectious-disease-outbreaks-in-india/ [23] A. Poornima and K. S. Priya, “A Comparative Sentiment Analysis of Sentence Embedding Using Machine Learning Techniques,” 2020, doi:1109/ ICACCS48705.2020.9074312.

[22] Swetha, G., Eashwar, V. M. A., & Gopalakrishnan, S. (2019). Epidemics and Pandemics in India throughout History: A Review Article. Indian Journal of Public Health Research and Development, 10(8), 1570. https://doi.org/10.5958/ 0976-5506.2019.02328.3

[23] Hoffman, Adam & McGuire, Luke & Mathews, Channing & Joy, Angelina & Law,Fidelia & Drews, Marc & Rutland, Adam & Hartstone-Rose, Adam & Winterbottom, Mark & Mulvey, Kelly Lynn. (2023). The importance of trust in the relation between COVID-19 information from social media and well- being among adolescents and young adults. PloS one. 18. e0282076. 10.1371/ journal. pone.0282076.

[24] Santomauro, D., Herrera, A. M. M., Shadid, J., Zheng, P., Ashbaugh, C., Pigott, D. M., Abbafati, C., Adolph, C., Amlag, J. O., Aravkin, A. Y., Bang- Jensen, B., Bertolacci, G. J., Bloom, S. S., Castellano, R., Castro, E., Chakrabarti, S., Chattopadhyay, J., Cogen, R. M., Collins, J. K., . . . Ferrari, A.

J. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. The Lancet, 398(10312), 1700–1712. https://doi.org/10.1016/s0140-6736(21)02143-7

[25] The 2006 dengue outbreak in Delhi, India. (2008c, December 1). PubMed. https://pubmed.ncbi.nlm.nih.gov/19579715/

[26] Bhatia, Manjeet. (2015). Psychiatric Morbidity in Patients with Chikungunya Fever: First Report from India. JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH. 9. 10.7860/JCDR/2015/14569.6586.

[27] Gupta, E., Dar, L., Kapoor, G., & Broor, S. (2006). The changing epidemiology of dengue in Delhi, India. Virology Journal, 3(1). https://doi.org/10.1186/1743-422x-3-92

[28] Das, S., Sarfraz, A., Jaiswal, N., & Das, P. (2017). Impediments of reporting dengue cases in India. Journal of Infection and Public Health, 10(5), 494–498. https://doi.org/10.1016/j.jiph.2017.02.004

[29] Hariharan, D., Das, M., Shepard, D. S., & Arora, N. K. (2019). Economic burden of dengue illness in India from 2013 to 2016: A systematic analysis. International Journal of Infectious Diseases, 84, S68–S73. https://doi.org/10.1016/j.ijid.2019.01.010

[30] Singh, Gaurav. (2021). Sentiment Analysis of Code-Mixed social media Text (Hinglish).

[31] Singh, Pranaydeep & Lefever, Els. (2020). LT3 at SemEval-2020 Task 9: Cross-lingual Embeddings for Sentiment Analysis of Hinglish social media Text.

[32] Singh, S., Pareek, A. (2022). A New Model SATV for Sentiment Analysis of Hinglish Sentences. In: Rathore, V.S., Sharma, S.C., Tavares, J.M.R., Moreira, C., Surendiran, B. (eds) Rising Threats in Expert Applications and Solutions. Lecture Notes in Networks and Systems, vol 434. Springer, Singapore. https://doi.org/10.1007/978-981-19-1122-4_3

[33] Gupta, A., Mishra, A., Reddy, U.S. (2021). Sentiment Analysis of Hinglish Text and Sarcasm Detection. In: Tripathi, M., Upadhyaya, S. (eds) Conference Proceedings of ICDLAIR2019. ICDLAIR 2019. Lecture Notes in Networks and Systems, vol 175. Springer, Cham. https://doi.org/10.1007/978-3-030- 67187-7_2

[34] Y. Asim, A. K. Malik, B. Raza, and A. R. Shahid, "A trust model for analysis of trust, influence and their relationship in social network communities," Telemat. Informatics, vol. 36, pp. 94–116, 2019, doi: 10.1016/j.tele.2018. 11.008.

[35] Grandjean, M. (2016). A social network analysis of Twitter: Mapping the digital humanities community. Cogent Arts & Humanities, 3(1), 1171458. https://doi.org/10.1080/23311983.2016.1171458

[36] Choudhary, Nurendra & Singh, Rajat & Bindlish, Ishita & Shrivastava, Manish. (2018). Sentiment Analysis of Code-Mixed Languages leveraging Resource Rich Languages.

[37] Islam, M. M., Aguilar, G., Ponnusamy, P., Mathialagan, C. S., Ma, C., & Guo, C. (2022). A Vocabulary-Free multilingual neural tokenizer for End-to-End task learning. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2204. 10815

[38] Kogilavani, S. V., Sampath, S. H., Nandhakumar, P., Mahalingam, P., Malliga, S., Kumaresan, P. K., & Priyadharshini, R. (2022). An analysis of machine learning models for sentiment analysis of Tamil code-mixed data. Computer Speech & Language, 76, 101407. https://doi.org/10.1016/j.csl.2022.101407

[39] Srinivasan, R., & Subalalitha, C. N. (2021). Sentimental analysis from imbalanced code-mixed data using machine learning approaches. Distributed and Parallel Databases. https://doi.org/10.1007/s10619-021-07331-4

[40] Ghanbari-Adivi, F., Mosleh, M. Text emotion detection in social networks using a novel ensemble classifier based on Parzen Tree Estimator (TPE). Neural Comput & Applic 31, 8971–8983 (2019). https://doi.org/10.1007/s00521-019-04230-9

[41] Shanmugavadivel, K., Sathishkumar, V. E., Raja, S., Lingaiah, T. B., Neelakandan, S., & Subramanian, M. (2022). Deep learning-based sentiment analysis and offensive language identification on multilingual code-mixed data. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-26092-3

[42] Choudhary, N., Singh, R. K., Bindlish, I., & Shrivastava, M. (2018). Sentiment Analysis of Code-Mixed Languages leveraging Resource Rich Languages. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1804.00806

[43] Prada, A., & Iglesias, C. Á. (2020). Predicting Reputation in the Sharing Economy with Twitter Social Data. Applied Sciences, 10(8), 2881. https://doi.org/10.3390/app10082881

[44] Hopp, Toby & Ferrucci, Patrick & Vargo, Chris. (2020). Why Do People Share Ideologically Extreme, False, and Misleading Content on Social Media? A Self-Report and Trace Data–Based Analysis of Countermedia Content Dissemination on Facebook and Twitter. Human Communication Research. 46. 357-384. 10.1093/hcr/hqz022.

[45] Xue J, Chen J, Hu R, Chen C, Zheng C, Su Y, Zhu T. Twitter Discussions and Emotions About the COVID-19 Pandemic: Machine Learning Approach. J Med Internet Res. 2020 Nov 25;22(11):e20550. doi: 10.2196/20550. PMID: 33119535; PMCID: PMC7690968.

[46] T. Khan and A. Michalas, "Trust and Believe - Should We? Evaluating the Trustworthiness of Twitter Users," 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), Guangzhou, China, 2020, pp. 1791-1800, doi: 10.1109/TrustCom50675. 2020.00246.

[47] Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, Ali A, Sheikh A. Artificial Intelligence-Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the United States: Observational Study. J Med Internet Res. 2021 Apr 5;23(4):e26627. doi: 10.2196/26627. PMID: 33724919; PMCID: PMC8023383.

[48] A. Mourad, A. Srour, H. Harmanani, C. Jenainati and M. Arafeh, "Critical Impact of Social Networks Infodemic on Defeating Coronavirus COVID-19 Pandemic: Twitter-Based Study and Research Directions," in IEEE Transactions on Network and Service Management, vol. 17, no. 4, pp. 2145-2155, Dec. 2020, doi: 10.1109/TNSM.2020.3031034.

[49] Masciantonio A, Bourguignon D, Bouchat P, Balty M, Rimé B. Don't put all social network sites in one basket: Facebook, Instagram, Twitter, TikTok, and their relations with well-being during the COVID-19 pandemic. PLoS One. 2021 Mar 11;16(3):e0248384. doi: 10.1371/journal.pone.0248384. PMID: 33705462; PMCID: PMC7951844.

[50] Prabhakar Kaila, Dr. Rajesh and Prasad, Dr. A. V. Krishna, Informational Flow on Twitter – Corona Virus Outbreak – Topic Modelling Approach (March 31, 2020). International Journal of Advanced Research in Engineering and Technology (IJARET), 11 (3), 2020, pp 128-134., Available at SSRN: https://ssrn.com/abstract=3565169

[51] Ashish Kumar Rathore, P. Vigneswara Ilavarasan, Pre- and post-launch emotions in new product development: Insights from twitter analytics of three products, International Journal of Information Management,Volume 50, 2020, Pages 111-127,ISSN 0268-4012,

[52] Debashis Naskar, Sanasam Ranbir Singh, Durgesh Kumar, Sukumar Nandi, and Eva Onaindia de la Rivaherrera. 2020. Emotion Dynamics of Public Opinions on Twitter. ACM Trans. Inf. Syst. 38, 2, Article 18 (April 2020), 24 pages. https://doi.org/10.1145/3379340

[53] Stone, J. A., & Can, S. H. (2020). Linguistic analysis of municipal twitter feeds: Factors influencing frequency and engagement. Government Information Quarterly, 37(4), 101468. https://doi.org/10.1016/j.giq.2020.101468

[54] Mosleh M, Pennycook G, Arechar AA, Rand DG. Cognitive reflection correlates with behavior on Twitter. Nat Commun. 2021 Feb 10;12(1):921. doi: 10.1038/s41467-020-20043-0. PMID: 33568667; PMCID: PMC7875970.

[55] Bhansali, R., & Schaposnik, L. P. (2020). A trust model for spreading gossip in social networks: a multi-type bootstrap percolation model. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 476(2235), 20190826. https://doi.org/10.1098/rspa.2019.0826

[56] Hamzelou, N., Ashtiani, M., & Sadeghi, R. (2021). A propagation trust model in social networks based on the A* algorithm and multi-criteria decision making. Computing, 103(5), 827–867. https://doi.org/10.1007/s00607-021-00918-w

[57] Du, Z., Luo, H., Lin, X., & Yu, S. (2020). A trust-similarity analysis-based clustering method for large-scale group decision-making under a social network. Information Fusion, 63, 13–29. https://doi.org/10.1016/j.inffus.2020.05.004

[58] Rao, H.S., Menaria, J.C., Chouhan, S.S. (2021). A Novel Approach for Sentiment Analysis of Hinglish Text. In: Sahni, M., Merigó, J.M., Jha, B.K., Verma, R. (eds) Mathematical Modeling, Computational Intelligence Techniques and Renewable Energy. Advances in Intelligent Systems and Computing, vol 1287. Springer, Singapore. https://doi.org/10.1007/978-981-15-9953-8_20

[59] G. I. Ahmad and J. Singla, "Sentiment Analysis of Code-Mixed Social Media Text (SA-CMSMT) in Indian-Languages," 2021 International Conference on Computing Sciences (ICCS), Phagwara, India, 2021, pp. 25-33, doi: 10.1109/ICCS54944.2021.00014.

[60] Singh, P., & Lefever, E. (2020). Sentiment Analysis for Hinglish Code-mixed Tweets by means of Cross-lingual Word Embeddings. CALCS.

[61] Garg, Neha. (2020). Annotated corpus creation for sentiment analysis in code-mixed Hindi-English (Hinglish) social network data. Indian Journal of Science and Technology. 13. 4216-4224. 10.17485/IJST/v13i40.1451.

[62] Pranaydeep Singh and Els Lefever. 2020. LT3 at SemEval-2020 Task 9: Cross-lingual Embeddings for Sentiment Analysis of Hinglish social media Text. In Proceedings of the Fourteenth Workshop on Semantic Evaluation, pages 1288–1293, Barcelona (online). International Committee for Computational Linguistics.

[63] Thakur, Varsha and Sahu, Roshani and Omer, Somya, Current State of Hinglish Text Sentiment Analysis (May 30, 2020). Proceedings of the International Conference on Innovative Computing & Communications (ICICC) 2020, Available at SSRN: https://ssrn.com/abstract=3614442 or http://dx.doi.org/10.2139/ssrn.3614442

[64] Gong, Z., Wang, H., Guo, W., Gong, Z., & Wei, G. (2020). Measuring trust in social networks based on linear uncertainty theory. Information Sciences, 508, 154–172. https://doi.org/10.1016/j.ins.2019.08.055

[65] Kakulapati, V., Mahender Reddy, S. (2020). Lexical Analysis and Mathematical Modelling for Analysing Depression Detection of Social Media Reviews. In: Gunjan, V., Senatore, S., Kumar, A., Gao, XZ., Merugu, S. (eds) Advances in Cybernetics, Cognition, and Machine Learning for Communication Technologies. Lecture Notes in Electrical Engineering, vol 643. Springer, Singapore. https://doi.org/10.1007/978-981-15-3125-5_10

[66] Bhih, A. E., Ghazzali, R., Rhila, S. B., Rachik, M., & Laaroussi, A. E. A. (2020). A discrete mathematical modeling and optimal control of the rumor propagation in online social network. Discrete Dynamics in Nature and Society, 2020, 1–12. https://doi.org/10.1155/2020/4386476

[67] Khrapov P. V., & Loginova A. A. (2019). Mathematical modelling of the dynamics of aids epidemics development in the world. International Journal of Open Information Technologies, 7 (6), 13-16.

[68] Ikoro, G.O. 2017. Mathematical modelling of the statistics of communication in social networks. Queen Mary University of London

[69] Sherchan, Wanita & Nepal, Surya & Paris, Cécile. (2013). A survey of trust in social networks. ACM Comput. ACM Computing Surveys. 45. 1-33. 10.1145/2501654.2501661.

[70] Majerczak, Przemysław & Strzelecki, Artur. (2022). Trust, Media Credibility, Social Ties, and the Intention to Share towards Information Verification in an Age of Fake News. Behavioral Sciences. 12. 51. 10.3390/bs12020051.

[71] Mohammad Hatamleh, Islam Habis & Safori, Amjad & Habes, Mohammed & Tahat, Othman & Ahmad, Amer & Abdallah, Rania & Aissani, Rahima. (2023). Trust in Social Media Enhancing Social Relationships. Social Sciences. 12. 1-22. 10.3390/socsci12070416.

List of Publications

[1] "Digital society social interactions and trust analysis model" published in PeerJ Computer Science journal, Indexed in Web of Science, ISSN/ISBN Number 2376-5992, Publication date: December 13, 2022

[2] "A Comparative Analysis of Various Models for Assessment of Trust in Digital Age Accelerated by Covid-19", Published in proceedings of 2021 International Conference on Computing Sciences (ICCS), ISSN/ISBN number Electronic ISBN:978-1-6654-9445-8 Print on Demand (PoD) ISBN:978-1-6654-9446-5, Publication date: 02-06-2022.

List of Conferences Attended

[1] 6th International Joint Conference on Computing Sciences (ICCS) “BOOTH100” held on 11th November 2022 held at Lovely Professional University.

[2] 5th International Conference on Computing Sciences (ICCS) “Kathleen 100” held on 4-5th December 2021 held at Lovely Professional University.