International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 11 Issue 10 ǁ October 2023 ǁ PP. 242-254

An overview of cultural shifts in society using theNLP algorithm, a constructive study for sentimentanalysis on the Twitter platform

1st Dr. Mohammed Nazeh Alimam
dept. Tech and Software) University of Europe for applied science)
Potsdam, Germany mohammednazeh.alimam@UE-germany.de

2nd Zeynab Piryaee
dept. Data Science)
University of Europe for applied science
Potsdam, Germany Zeynab.piryaee@UE-germany.de

Abstract—This thesis explores cultural shifts within society through sentiment analysis. The main objective is to create an au- tomated framework for extracting and analyzing sentiments from tweets, categorizing them as positive, negative, or neutral using machine learning techniques. The approach combines lexicon- based sentiment analysis with supervised machine learning algo- rithms for classification. Performance evaluation of the hybrid framework includes precision, accuracy, recall, and F1 score metrics. Among the four algorithms (SVM, Naive Bayes, Logistic Regression, and Random Forest) tested, Logistic Regression performs best, achieving the highest accuracy (0.7697) and F1- score (0.77) in sentiment analysis. However, other algorithms also show respectable results, with SVM and Random Forest achieving competitive accuracy scores (0.7623 and 0.7445), and Naive Bayes achieving a fair score (0.6900). A line chart comparing the models visually highlights Logistic Regression's consistent superiority across all metrics. SVM and Random Forest perform similarly, with SVM slightly ahead in accuracy and F1-score, and Random Forest excelling in precision and recall. Naive Bayes lags in accuracy and F1-score, suggesting challenges in handling sentiment complexities in this domain. Given its high accuracy and balanced performance, Logistic Regression is recommended as the preferred model for sentiment analysis of cultural shift reviews on Twitter. Nonetheless, model selection should consider factors like interpretability, training time, and specific application needs. In conclusion, this thesis leverages sentiment analysis on Twitter to provide insights into cultural shifts, showcasing the power of NLP algorithms in understanding public sentiments during critical events. These findings enhance understanding of collective sentiments and cultural transformations, aiding informed decision-making and social analysis.

Index Terms—*movement, machine learning, NLP, cultural shift, algorithms,lexicon-based, sentiment analysis, tweets.*

Date of Submission: 09-10-2023 Date of acceptance: 23-10-2023

I. INTRODUCTION

The rapid growth of social media platforms has ushered in a new era of information sharing, redefining the way people communicate and express their opinions in society. Among these platforms, Twitter has gained immense popularity as a microblogging site, especially during significant events like the COVID-19 pandemic, movements, and political issue. With millions of users actively engaging on the platform, Twitter serves as a valuable source of public sentiment and cultural shifts.[11][1] Sentiment analysis, a subfield of natural language processing (NLP), plays a crucial role in understanding the sentiments and emotions expressed by users in their tweets. Analysing sentiments on Twitter during critical events can offer valuable insights into societal perceptions, public sentiments, and cultural transformations. This study aims to address the need for a constructive approach to sentiment analysis on the Twitter platform, specifically focused on understanding cultural shifts in society. To achieve this, the research proposes the development of an automated framework that extracts and analyses sentiments from tweets. The framework employs a combination of lexicon-based techniques for tweet sentiment analysis and labeling, along with supervised machine learning methods for tweet classification.[8] By leveraging machine learning algorithms, the framework enables the categorization of tweets into positive, negative, and neutral sentiments, pro- viding a comprehensive overview of the sentiments expressed by users. Performance evaluation of the hybrid framework is carried out using precision, accuracy, recall, and F1 score as key measures, offering a quantitative assessment of the model's effectiveness.

Among the algorithms tested, Logistic Regression emerges as the top performer, demonstrating the highest accuracy and F1-score in the sentiment analysis task. However, the study also acknowledges the competitive perfor- mance of SVM, Random Forest, and Naive Bayes algorithms, despite being slightly outperformed by Logistic Regression. The findings from this constructive study hold significant implications for understanding public perceptions and culturalshifts during critical events.
Through sentiment analysis, the research contributes to a deeper comprehension of collective sentiments in society, aiding. [24] decision-makers, researchers, and social analysts in making informed choices based on the sentiments expressedon Twitter.

II. LITERATURE REVIEW

Overview of cultural shifts in society: The text discusses the impact of the digital age on cultural shifts and outlines key points related to this transformation. The digital age has ushered in a profound cultural transformation. Here, we delve

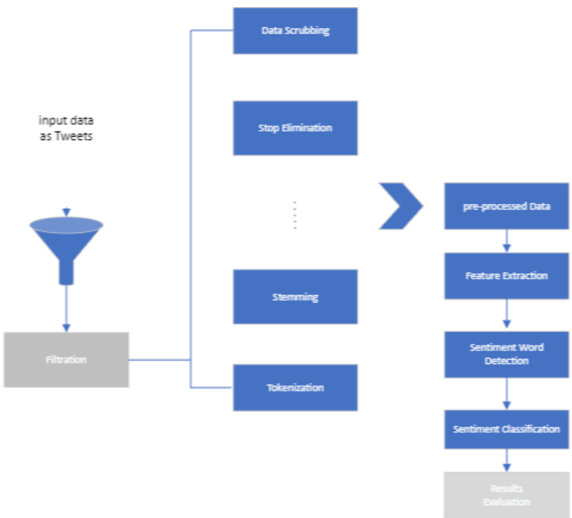

Fig. 1. Figure 1.1 Sentiment analysis approach on Twitter.

into key aspects that directly shape this cultural shift: User Empowerment: In the digital era, user experience has shifted from passivity to active engagement. People can access a di- verse array of content and even create their own. Social media platforms serve as hubs for content sharing, enabling users to disseminate information, collaborate on content creation, and tap into collective knowledge and creativity.
Communication Evolution: Digital technologies and the internet have revolutionized communication. Traditional face- to-face interactions have been supplemented, if not replaced, by digital channels such as social media, messaging apps, and video conferencing. This shift has accelerated the pace of communication and transcended geographical boundaries.
Cultural Adaptation: Cultural transformation is pivotal in the digital age. Cultural institutions, organizations, and individuals have embraced digital platforms to engage wider audiences, share narratives, and preserve heritage. Online education op- portunities enable global access to cultural knowledge, foster- ing cross-cultural understanding and collaboration.
Cultural Diversity: The digital age has given rise to a vast array of cultural products, ideas, and behaviors online, from viral trends to social media subcultures. This rich tapestry of cultural variation provides fertile ground for the study of cultural evolution.
Perception Redefined: Our perception of communication, culture, and business management skills has been reshaped by the digital age. While it offers new avenues for interaction, it also poses challenges such as information overload and digitaletiquette nuances.
Business Agility: Businesses must adapt swiftly in the digital age. Technological advancements and evolving con- sumer expectations demand continuous innovation and agility,essential skills for managers.
The production and transmission of culture have been equally revolutionized by digital media, with several pivotal shifts:
Accessibility and Democratization: Digital media has de- mocratized cultural production, allowing individuals to create and share content without traditional gatekeepers. This has diversified cultural voices and perspectives.
User-Generated Content: Online platforms empower users to actively contribute to cultural content, from videos

to music,fostering a participatory culture.
Digital Distribution and Consumption: Digital media has transformed how we access and consume cultural content,transcending physical formats and enabling personalized rec- ommendations.
Remix Culture and Mashups: Easy access to existing con- tent has fueled a culture of remixing and mashups, giving rise to new forms of art and expression.
Virality and Global Reach: Digital media amplifies the potential for cultural content to go viral, spreading across the internet and gaining global attention.
Cultural Preservation and Archiving: Digital media allows for the preservation and digitization of cultural artifacts, mak- ing them more accessible for future generations.
Twitter, in particular, serves as a gateway to culture, influ- encing and reflecting cultural shifts:
Influence of Social Media: Twitter plays a pivotal role in shaping cultural phenomena and societal attitudes, thanks toits vast user base and global reach.
Echo Chambers: However, it also raises concerns about echo chambers, where users are exposed primarily to like-minded perspectives, potentially reinforcing existing beliefs.
Fear of Missing Out (FOMO): The constant stream of content on social media can induce a fear of missing out, prompting impulsive behavior to keep up with trends.

Twitter, as a platform, has a significant impact on social attitudes and movements:
Social Attitude Shaping: Twitter allows users to share ideas and perspectives, contributing to the shaping of social attitudeson various issues.
Hashtags and Movements: Hashtags like MeToo and Black- LivesMatter have amplified social movements, leading to legislative changes.
In summary, the digital age has triggered a profound cultural shift, fundamentally altering how we create, consume, and interact with culture and information, with Twitter standing as a significant influencer in this ever-evolving landscape.

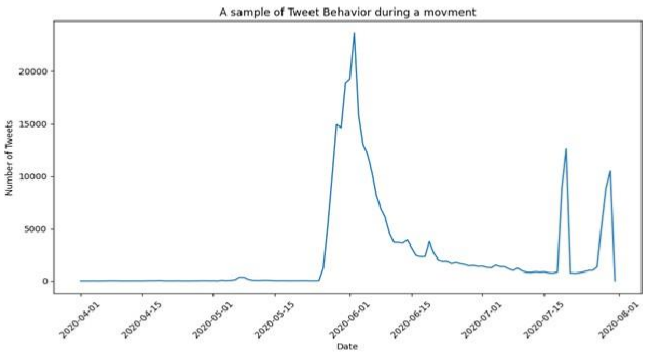

Fig. 2. a sample of tweet behavior during MeToo

A. Using natural language processing (NLP) to investigate cultural change and developments in the Twitter platform:

Language serves as a means of communication through which we express ourselves, read, and write. Our thoughts, decisions, and plans are often formulated using natural lan- guage, specifically through words. However, in the era of artificial intelligence (AI), a fundamental question arises: Can we effectively communicate with computers using our natural language? Essentially, we face the challenge of developing applications in natural language processing (NLP) as comput- ers typically require structured data, while human speech is unstructured and often contains ambiguity. In this context, Nat- ural Language Processing (NLP) can be described as a branch of Computer Science, particularly Artificial Intelligence (AI), that focuses on equipping computers with the ability to com- prehend and process human language. In essence, the primary objective of NLP is to program computers to analyse andprocess vast amounts of data written or spoken in naturallanguage.

B. NLP Phases

The following diagram shows the phases or logical steps in natural language processing.

C. Sentiment analysis on Twitter

Advancements in NLP, sentiment lexicons, and machine learning led to rapid growth in Twitter sentiment analysis. Researchers addressed challenges posed by informal tweets, including emoticons and slang.

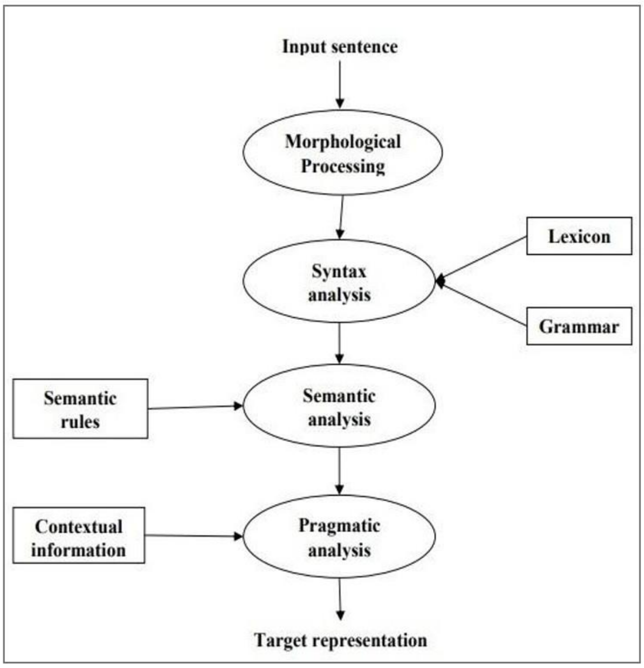

Fig. 3. logical steps in natural language processing

Twitter introduced the Twitter API, enabling large-scale data collection for sentiment analysis and sparking further research. Sentiment analysis on Twitter is used in marketing, politics, and more, with continuous improvements like deep learning models and tailored sentiment lexicons.

This field remains active and dynamic, with ongoing re- search enhancing the accuracy of sentiment analysis on Twitterdata, which helps understand public sentiment and monitor brands.

1) section There are several approaches to performing sentiment analysis on Twitter:: Lexicon-based approach: This method uses predefined sentiment lexicons or dictionaries that contain words and their associated sentiment scores. The senti- ment scores can be positive, negative, or neutral. The approach involves matching words in tweets with the lexicon and ag- gregating the scores to determine overall sentiment. Lexicons are built through manual annotation or automated methods using large collections of text data. Examples of widely used sentiment lexicons include SentiWordNet, AFINN, and VADER (Valence Aware Dictionary and Sentiment Reasoner). Word-level Sentiment Analysis: In the Lexicon-based ap- proach, individual words in a tweet are matched against the sentiment lexicon. Each word is assigned a sentiment score based on its entry in the lexicon. These scores can beaggregated to compute an overall sentiment score for the tweet.

For example, positive and negative scores can be summed up to determine whether the sentiment is positive or negative.

Machine learning-based approach: This approach involves training a machine learning model using a labelled dataset. The dataset consists of tweets that are manually labelled as positive, negative, or neutral. The model learns patterns and features from the labelled data and applies them to classify new, unseen tweets. Challenges: Lexicon-based approaches face certain challenges. They may struggle with sarcasm, irony, context-dependent sentiment, and the evolving nature of language on social media platforms. Lexicons may not cover all possible words or phrases, leading to potential gaps in sentiment analysis. Additionally, sentiment lexicons are often language-dependent, and their effectiveness may vary across different languages and cultures.[26]

Hybrid Approaches: To enhance the performance of the Lexicon-based approach, researchers often combine it with other techniques. Hybrid approaches may incorporate ma- chine learning algorithms to fine-tune sentiment analysis using lexicon-based methods as a starting point. This combination leverages the strengths of both approaches and helps address the limitations of individual methods.

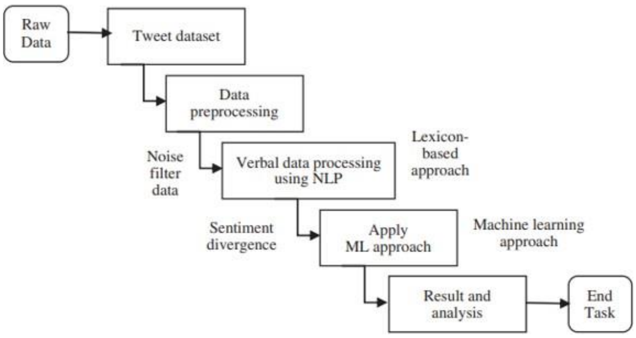

Fig. 4. framework of sentiment analyst

D. This research used machine learning algorithms for senti- ment analysis on Twitter include:

Naive Bayes: These classifiers are efficient and assume feature independence, making them suitable for large Twitter datasets.

Support Vector Machines (SVM): SVM finds optimal hyper- planes to separate sentiment classes, handling complex Twittersentiment patterns.

Logistic Regression: This statistical technique predicts senti- ment labels based on input features, offering a straightforwardapproach.

Random Forest: It combines multiple decision trees, re- ducing overfitting and capturing nuanced Twitter sentiment patterns.

1) Algorithms comparison table:: This research tries an ex- tensive exploration of sentiment analysis methods and machine learning algorithms, positioning them as formidable tools for deciphering sentiments amidst a diverse range of contexts. This chapter lays a solid groundwork for the subsequent research journey, with a specific focus on the intricate dynamics of sentiment fluctuations within the sphere of cultural shifts on the Twitter platform.

As societal dynamics continue to evolve, catalysed by events of global significance, sentiment analysis emerges as a potent lens through which to unravel the multifaceted layers of human response. This chapter delves into the intricacies of sentiment analysis, considering a spectrum of parameters that encapsulate an array of impactful events. From mon- umental movements that have galvanized masses, such as socio-political uprisings and environmental advocacy, to the unprecedented challenges posed by the COVID-19 pandemic, this research sheds light on the sentiments that reverberate across Twitter’s virtual landscape.

Table 3.1:The algorithms chosen for this study

Model	**Type**	**Tuned Parameters**	**Description**
Naïve Bayes	Supervised Learning	Smoothing (alpha)	Naïve Bayes is a probabilistic classifier based on Bayes' theorem, commonly used for text classification tasks.
RF	Ensemble Learning	Number of Trees, Max Depth	Random Forest is an ensemble of decision trees that aggregates their outputs, offering improved accuracy and robustness.
LR	Supervised Learning	Regularization Strength (C)	Despite the name, Logistic Regression is a linear classifier that predicts the probability of a binary outcome using a sigmoid function. It's commonly used for binary classification tasks.
SVM	Supervised Learning	Regularization Parameter (C), Kernel	SVM (Support Vector Machine) is a powerful classifier that separates data into different classes using hyperplanes.

TABLE I
ALGORITHMS COMPARISON TABLE

III. METHODOLOGY

This research involved gathering tweets from Twitter to predict their value and identify their unique characteristics. The initial dataset consisted of tweets, which underwent pre- processing to extract relevant information. This extracted data was then used for estimating and analysing the classification of the collected tweets.

A. Data Preprocessing

Fig. 5. Framework of tweet analysis mechanism

B. Data collection and preprocessing

The process of preparing the data involves crucial tasks such as eliminating short words and white spaces, tokenizing the text, assigning part-of-speech tags, and performing lemmati- zation. Eventually, I recombined the lemmatized text to its original form.[22][23]

1) Tokenization:: Split tweets into individual words or tokens.

2) Lowercasing:: Convert all tokens to lowercase to ensureconsistent processing.

3) Stop word Removal:: Remove common words (e.g., ”and,” ”the,” ”is”) that don’t contribute much to sentiment analysis.

4) Removing Special Characters:: Remove punctuation, hashtags, URLs, and other non-essential symbols.

C. Feature Selection

1) Unigrams and Bigrams:: Consider individual words (un- igrams) and pairs of consecutive words (bigrams) as features.

2) TF-IDF (Term Frequency-Inverse Document Frequency):: Calculate the importance of words in each tweet relative to the entire dataset. Words that are common in a particular tweet but rare in the entire dataset are given higher weights. The process of feature selection in text classification involves extracting both semantic and syntactic features. The semantic feature encompasses the sentiment associated with each word, while the syntactic feature includes unigrams, bigrams, and n-grams, which are groups of associated words.

[6] Unigrams refer to single words like ”Global” ”Corona ” and ”life” whereas bigrams are combinations of two words like ”Me Too”, ”Global Warming”, and ”Corona Virus” All Unigrams, Bigrams under the category of n-grams, which are commonly used for text classification.

3) Unigrams: $\text{extract}_{u}nigrams(text)$

4) Bigrams: $\text{extract}_{b}igrams(text)$

5) Trigrams: $\text{extract}_{t}rigrams(text)$

6) $Syntactic_{f}eatures$: unigrams + bigrams + trigrams

D. Feature Selection

1) Bag-of-Words (BoW): Create a matrix where each row corresponds to a tweet and each column corresponds to aunique word. The matrix values can represent word countsor TF-IDF weights.

2) Sentiment Features: Introduce features like sentiment scores (e.g., using the VADER lexicon) and emotion distri- butions for each tweet.

3) Define a function to perform sentiment analysis using VADER:

E. perform sentiment analysis

VADER, a potent text sentiment extraction technique, ex- hibits remarkable proficiency in categorizing tweet contents into positive, negative, and neutral sentiments. As evidenced in the article, VADER achieved an impressive accuracy of 0.96 in sentiment extraction from microblogging sites, making it a reliable and effective choice for tweet sentiment analysis. The

```
compound_score = sentiment_scores['compound']
if compound_score >= 0.05
sentiment = 'positive'
elif compound_score <= -0.05
sentiment = 'negative'

else
        sentiment = 'neutral'

   return sentiment, sentiment_score
```

technique's accuracy is attributed to its reliance on a lexicon- based dictionary, which associates sentiments with words, ensuring precise sentiment extraction from tweets.

An outstanding aspect of VADER is its contextual un- derstanding of phrases, enabling it to appropriately label sentiments in phrases like "diminishing the COVID cases" as positive. Given that the article discusses data consisting of short tweet snippets, employing a limited-class sentiment extraction approach with two or three main classes (positive, negative, and neutral) proves to be an appropriate and well-supported strategy, as validated in prior studies [17], [23], [20]. To effectively distribute tweets among the sentiment classes, the article utilizes a 0.05 threshold value for the compound score. Employing conditional computation, tweets are assigned to their respective categories: those with a compound score greater than or equal to the threshold value are classified aspositive, those with a compound score less than or equal to -0.05 are labeled as negative, while tweets with compound scores between +0.05 and -0.05 are categorized as neutral. This meticulous approach ensures accurate and comprehensive sentiment classification, as demonstrated in the findings of thearticle.

Sentiment analysis helps to understand the sentiment or emotion behind each piece of text, which can be valuable for understanding user opinions, trends, or reactions. On the other hand, TF-IDF feature extraction is a common technique for representing text data numerically, which is crucial for many NLP tasks like text classification, clustering, and topic modelling.[12][10] approach of this thesis is analysing the sentiment of the text while also using the same text data for other analyses. using the sentiment scores as features alongside the TF-IDF vectors for a more comprehensive analysis.

F. Classifiers:

"In the pursuit of gaining deeper insights into the cultural shifts within our society, this study harnesses the power of advanced Natural Language Processing (NLP) algorithms for sentiment analysis. The aim is to unveil the evolving sentiments expressed in tweets on the Twitter platform.[15] To achieve this objective, a selection of robust supervised machine learning techniques has been meticulously chosen as the classifiers for sentiment classification. These classifiers are not only adept at processing text-based data but also offer distinct advantages in understanding and interpreting sentimenttrends.[22]

The chosen algorithms for this study are:

1) Na¨ıve Bayes: Known for its efficiency in text classifi- cation, it's used for categorizing tweets into sentiment classes due to its simplicity and computational efficiency.

2) Logistic Regression (LR): Despite its name, it's a power- ful statistical technique for understanding sentiment dynamics. It calculates outcome probabilities and interprets feature coef- ficients to uncover linguistic patterns.

3) Random Forest (RF): This ensemble learning method combines decision trees to capture complex sentiment nuancesand mitigate overfitting. It also reveals influential terms in sentiment changes.

4) Support Vector Machine (SVM): SVM identifies op- timal hyperplanes to separate sentiment classes, crucial for discerning nonlinear sentiment patterns, especially in evolvingcultural narratives.

G. Evaluation Metrics

One paragraph or more explaining the evaluation metrics with equations for each metric, where possible.

H. Experimental settings

One paragraph for experimental settings of your and com- peting methods (if any). (Optional) One paragraph for hyper- parameter settings and network architecture

I. Dataset and Computational Infrastructure

The analysis is conducted using a publicly available tweet dataset obtained from Kaggle.com [56]. This dataset contains tweets related to social events and includes additional informa- tion such as location, retweets, followers, friends, tweet count, and hashtags. Due to data anonymization, our study did not require ethics review and approval. The dataset comprises a total of 179,108 tweets collected from 2019 to 2022.

Applying machine learning techniques to textual data de- mands significant computational resources, substantial time, and cost investments. To address these challenges, this thesis implemented classification framework pipeline on the Google Cloud Platform (GCP)

J. Performance Metrics

In evaluating the performance of the machine learning techniques employed, we utilized several classification per- formance metrics, including accuracy, precision, recall, and F1-score. Each metric provides valuable insights into the effectiveness and reliability of the classifiers applied to assess sentiments during diverse social events.

1) Precision : TP/(TP + FP)

2) Recall : TP/(TP + FN)

3) F1 score : (2(precision x recall))/(precision + recall)

4) Accuracy : (TP + TN)/(TP + FP + FN + TN) The utilized machine learning techniques were thoroughly evaluated, and the analysis involved a range of classification performance metrics, including accuracy, precision, recall, and F1-score. Formally, TP (True Positives) represented the correctly classi- fied positive samples, while TN (True Negatives) indicated the accurately classified negative samples. Conversely, FP (False Positives) represented the number of negative samples mistak- enly classified as positive, and FN (False Negatives) denoted the positive samples erroneously classified as negatives. The overall performance was assessed using the accuracy metric. Precision is a metric that gauges the accuracy and relevance of the generated results. It quantifies the ratio of true positive instances to the combined total of true positive and false positive instances. On the other hand, recall is a metric that evaluates the completeness and correctness of the classified tweets. It calculates the ratio of true positive instances to the combined total of true positive and false negative instances. The F1-Score is a composite metric that combines both recall and precision scores, providing a balanced measure of the model's overall performance [20]. Performance measures are computed for all the algorithms. Performance measures such as precision, recall, and f1-scores. Figure 4 to Figure 6 is used to represent the sentiment classification with pre-processes andMLTSA algorithm.

IV. RESULTS, DISCUSSION

The sentiment analysis framework employed in this study to explore cultural shifts on Twitter has yielded insightful out- comes, illuminating the evolving sentiments embedded within the platform's vast textual landscape. This chapter delves into the findings, interpretations, and implications arising from the application of four distinct algorithms—Support Vector Ma- chine (SVM), Naive Bayes, Logistic Regression, and Random Forest—to decipher the sentiments expressed in tweets.[20] Algorithm Performance Analysis

A. section The performance of the sentiment

analysis algorithms has been thoroughly evaluated using various metrics. Each algorithm's accuracy, precision, recall, and F1-score have been computed to gauge their effectiveness in classifying sentiments. Below is a summary of the perfor- mance metrics for each algorithm:

B. Algorithmic Comparison and Interpretation

[htbp] The results reveal distinct insights into the perfor- mance of each algorithm and their suitability for sentiment analysis within the context of cultural shifts. Noteworthy observations can be made from the metrics:

SVM demonstrates a respectable accuracy of 0.7623. It excels in predicting ”negative” sentiments, maintaining an impressive precision and recall, while its performance on ”neutral” and ”positive” sentiments is also noteworthy.

Naive Bayes exhibits an accuracy of 0.6900. While it demonstrates strong recall for ”negative” sentiments, its preci-sion for ”neutral” sentiments is particularly high. However, the

Metric	Value
Accuracy	0.7623
Precision (Negative)	0.81
Precision (Neutral)	0.67
Precision (Positive)	0.78
Recall (Negative)	0.85
Recall (Neutral)	0.65
Recall (Positive)	0.74
F1-Score (Negative)	0.83
F1-Score (Neutral)	0.66
F1-Score (Positive)	0.76

Fig. 6. SVM Performance Metrics

Metric	Value
Accuracy	0.6900
Precision (Negative)	0.64
Precision (Neutral)	0.80
Precision (Positive)	0.75
Recall (Negative)	0.96
Recall (Neutral)	0.32
Recall (Positive)	0.67
F1-Score (Negative)	0.77
F1-Score (Neutral)	0.46
F1-Score (Positive)	0.71

Fig. 7. Naive Bayes Performance Metrics

Metric	Value
Accuracy	0.7697
Precision (Negative)	0.82
Precision (Neutral)	0.69
Precision (Positive)	0.78
Recall (Negative)	0.87
Recall (Neutral)	0.65
Recall (Positive)	0.75
F1-Score (Negative)	0.84
F1-Score (Neutral)	0.67
F1-Score (Positive)	0.77

Fig. 8. Logistic Regression Performance Metrics

Metric	Value
Accuracy	0.7503
Precision (Negative)	0.76
Precision (Neutral)	0.69
Precision (Positive)	0.79
Recall (Negative)	0.87
Recall (Neutral)	0.60
Recall (Positive)	0.72
F1-Score (Negative)	0.82
F1-Score (Neutral)	0.64
F1-Score (Positive)	0.75

Fig. 9. Random Forest Performance Metrics

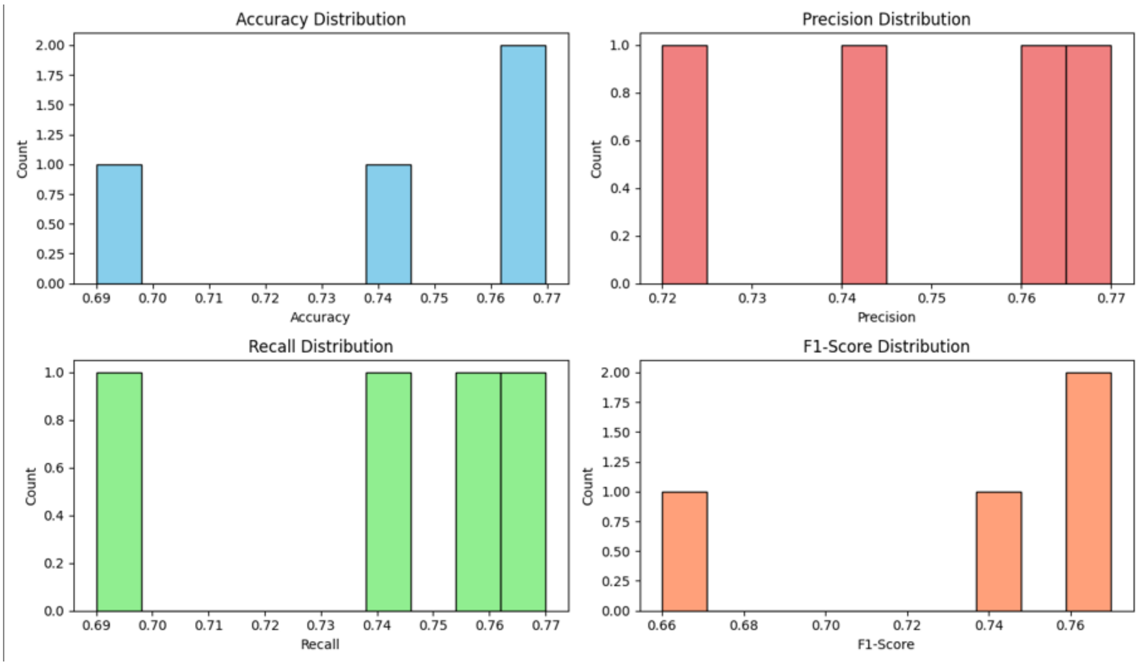

Fig. 10. matrices distribution of sentiment analysis model.

lower F1-scores across the board suggest challenges in finding the right balance between precision and recall. Logistic Regression emerges as a standout performer with an accuracy of 0.7697. It achieves the highest F1-scores across all sentiment categories, indicating a robust balance between precision and recall. Its capability to handle nuanced sentiment shifts is evident, making it a reliable choice for analysing cultural sentiment evolution. Random Forest maintains a commendable accuracy of 0.7503. It performs well in identifying "negative" and "positive" sentiments, albeit with a relatively lower recall for "neutral" sentiments. The overall F1-scores reflect its ability to capture sentiment trends despite the intricate landscape of cultural shifts. In sum, the "Logistic Regression" algorithm emerges as the most versatile and balanced model for predicting sentiment in the context of cultural shift reviews on Twitter. Its accuracy, precision, recall, and F1-scores collectively position it as a robust choice, reflecting its capacity to capture the intricate nuances of sentiments tied to evolving cultural narratives.

V. DISCUSSION

Five to six paragraphs discussing the results (at least one paragraph for each research question). Your opinion on how good/bad the results are. Draw inferences from the results here. Explain novelty of your contributions and what was missing that you have explored here. Discuss how results from your proposed method compare with other existing contemporary methods in this section. Any other point you would like to

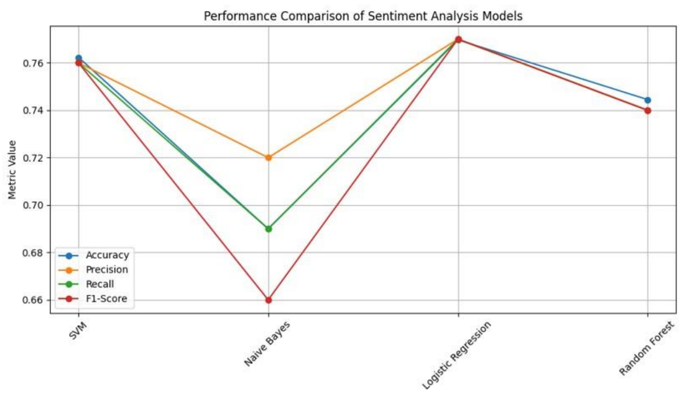

Fig. 11. performance comparison of sentiment analysis model

discuss related to this study. One paragraph for what are the future directions in your opinion for continuing this study. Limitations of your work, if any? Any assumptions that effect your analysis?

VI. CONCLUSION

In the contemporary global landscape, the profound im- pact of cultural and social transformations on human societies is undeniable. The ability to discern and forecast these transformative shifts holds immense scientific and practical significance, not only informing sociological planning but also shaping effective policy-making. Within this context, the analysis of sentiments and perspectives emerges as a potent instrument, serving as a representative gauge of cultural and social orientations, thereby facilitating the study of these intricate changes.

This research has embarked on a comprehensive exploration of cultural metamorphoses within the dynamic ecosystem of Twitter, employing the formidable synergy of sentiment analysis and Natural Language Processing (NLP). Through the strategic deployment of diverse algorithms including Logistic Regression (LR), Naive Bayes, Random Forest, and Support Vector Machine (SVM), a spectrum of models for sentiment and perspective analysis has been meticulously crafted. Rig- orous investigations have underscored the nuances of each algorithm, culminating in a comparative assessment of their efficacy.

The culmination of sentiment analysis outcomes, coupled with insights derived from meticulous data analysis and the synthesis of diverse machine learning models, unequivocally underscores Twitter's role as a wellspring of invaluable data. This data reservoir holds the potential to detect and anticipate cultural revolutions and social movements. Notably, this analytical prowess transcends mere trend monitoring, extending its impact on the spheres of sociology, cultural analysis, and political strategy, thus fortifying society's ability to navigate and manage these transformative currents.

The synthesis of this research highlights the success achieved in deploying varied sentiment analysis algorithms and conducting insightful model comparisons. This diversity in approach enhances the granularity of change trend depiction, empowering analysts to forge judicious decisions aimed at the advancement of societal, cultural, and political dimensions.

In summation, the findings of this study affirm that sen- timent and perspective analysis, gleaned from the Twitter milieu, emerge as influential tools for comprehending and forecasting cultural and social transformations. By harnessing an array of algorithms, this research dissects individual in- clinations and attitudes during moments of cultural evolution with enhanced precision. This newfound insight empowers researchers and analysts not only to foresee cultural and social orientations but also to bolster societal strengths and rectify shortcomings, thereby furnishing solutions for adroitly managing these paradigm shifts.

Moreover, the meticulous comparison of disparate machine learning algorithm models uncovers a nuanced landscape of strengths and limitations, offering a glimpse into their optimal alignment with specific topics. This discernment equips analysts to judiciously select an algorithm tailored to sentiment analysis, thoughtfully considering the intricacies of specific issues and research goals.

In conclusive reflection, this research eloquently under- scores the formidable synergy between sentiment analysis, natural language processing, and machine learning algorithms, constituting a robust apparatus for unraveling the intricate tapestry of cultural and social metamorphoses. This analytical arsenal amplifies our insights into change trends, augments the precision of predictions, and redounds to the advancement of societal well-being, particularly amid swift and substantial transformations. As novel methodologies for sentiment and perspective analysis are introduced, this research propels us toward a more nuanced understanding of the symbiotic inter- play between culture, society, and their relentless evolution.

REFERENCES

[1]. Rodogno, R. Personal identity online. Philos. Tech- nol. 2011, 25, 309–328.

[2]. Langlois, G. Social media and the care of the self. In Digital Existence: Ontology, Ethics, and Transcendence in Digital Culture; Lagerkvist, A., Ed.; Routledge: New York, NY, USA, 2019; pp.156–170

[3]. Prinz, J. Culture and Cognitive Science; In The Stanford Encyclopedia of Philosophy (Summer 2020 Edition). Available online: https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=culture-cogsci (accessedon 14 January 2021).

[4]. Levin, I. Cultural Trends in a Digital Society. In Proceedings of the 10th International Symposium on Tools and Methods of Competitive Engineering TMCE, Budapest, Hungary, 19–23 May 2014; ISBN 978-94-6186- 177-1.

[5]. K. Lagutina, V. Larionov, V. Petryakov, N. Lagutina, Paramonov, and I. Shchitov, "Sentiment Classification into Three Classes Applying Multinomial Bayes Algorithm, N- grams, and Thesaurus," in 2019 24th Conference of Open Innovations Association (FRUCT), 2019, pp. 214–219.

[6]. S. Arlot and A. Celisse, "A survey of cross-validation proce- dures for model selection," Stat. Surv., vol. 4, no. none, pp. 40–79, 2010, doi: 10.1214/09-SS054.

[7]. Ilya Levin, Dan Mamlok, "Culture and Society in the Digital Age",12(2), 68; https://doi.org/10.3390/info12020068, 5 February 2021

[8]. Arun, Ayyagari Srinagesh, Multi-lingual Twitter sentiment analysis using machine learning, December 2020, pp. $21_3$4.

[9]. Masoom Raza, Rashmi Phalnikar, 'COURSE SYL- LABUS COMPARISON USING ONTOLOGY DESIGN AND NLP', April. -2023, pp. $21_2$8.

[10]. B. B. Elov1, Sh. M. Khamroeva1and Z. Y. Xusainova1, "The pipeline processing of NLP",

https://doi.org/10.1051/e3sconf/202341303011,August2023, pp. 5₁1.
[11]. R. Khan, R. Khan, P. Shrivastava, A. Kapoor, A. Tiwari, and A. Mittal, "Social media analysis with AI: Sen- timent Analysis techniques for analyzing Twitter COVID-19 Data," J. Crit. Rev., vol. 7, no. 9, pp. 2761- -2774, 2020.
[12]. G. Becquin, End-to-end NLP Pipelines in Rust (2020)https://doi.org/10.18653/v1/2020.nlposs-1.4
[13]. Norah Fahad Alshammari, Amal Abdullah,' State-of-the-art review on Twit- ter Sentiment Analysis', DOI:10.1109/CAIS.2019.8769465, May 2019, pp.20₂6.
[14]. H. Kaur, S. U. Ahsaan, B. Alankar, and V. Chang, "A Proposed Sentiment Analysis Deep Learning Algorithm for Analysing COVID-19 Tweets," Inf. Syst. Front., 2021, doi: 10.1007/s10796-021-10135-7.
[15]. P. Gupta, S. Kumar, R. Suman, and V. Kumar, "Sen- timent Analysis of Lockdown in India During COVID-19: A Case Study on Twitter," IEEE Transactions on ComputationalSocial Systems, 2020.
[16]. S. Narr, et al., "Language-Independent Twitter Senti- ment Analysis," Workshop on Knowledge Discovery, Data Mining and Machine Learning (KDML 2012), 2012
[17]. S. Strassel, et al., "Handbook of Natural Language Processing and Machine Translation: DARPA Global Autonomous Lan- guage Exploitation," Springer-Verlag, 2011.
[18]. A. Balahur and M. Turchi, "Improving sentiment analysis in twitter using multilingual machine translated data," Proceedings of the International Conference Recent Advances in Natural Lan-guage Processing, pp. 49-55, 2013.
[19]. E. Kouloumpis, etal., "Twitter Sentiment Analysis: The Good the Bad and the OMG!" in Proceedings of Fifth International AAAI Confer- ence on Weblogs and Social Media, pp. 538-541, 2011. 20.
[20]. H. Saif, et al., "Semantic sentiment analysis of twitter," In- ternational Semantic Web Conference, pp. 508-524, 2012. 21.
[21]. Brönnimann, "Multilanguage sentiment-analysis of Twitter data on the example of Swiss politicians," M.Sc. Thesis, University of Applied Sciences Northwestern Switzerland, 2014.
[22]. Z. Y. Xusainova, NLP: tokenizatsiya, stemming, lemmatizatsiya va nutq qismlarini teglash. O'zbek amaliy filologiyasi istiqbollari, Respublika ilmiy-amaliy konferensiyato'plami. Elektron nashr, Toshkent: ToshDOTAU, 159-163(2022)
[23]. R. M. Merchant and N. Lurie, "Social Media and Emergency Preparedness in Response to Novel Coronavirus," JAMA, vol. 323, no. 20, pp. 2011–2012, May 2020, doi:10.1001/jama.2020.4469
[24]. F. Rustam, M. Khalid, W. Aslam, V. Rupapara, A. Mehmood, and G. S. Choi, "A performance comparison of supervised machine learning models for Covid-19 tweets sen- timent analysis," PLoS One, vol. 16, no. 2, p. e0245909, 2021.
[25]. I. Mozetič, et al., "Multilingual Twitter Sentiment Clas- sification: The Role of Human Annotators," PLOSONE, vol. 11, no. 5, 2016.
[26]. Piyush Vyas, Martin Reisslein Fellow, Bhaskar P. Rimal, Senior Member, Gitika Vyas, Ganga Prasad Basyal, Student Member, and Prathamesh Muzumdar, "Automated Classifica- tion of Societal Sentiments on Twitter with Machine Learning"August 21,PP(99):1-1
[27]. C. J. Hutto and E. Gilbert, "VADER: A Parsimo- nious Rule-based Model for Sentiment Analysis of Social Media Text," Proc. Int. AAAI Conf. Web Soc. Media, vol.8, no. 1, 2014, Accessed: Dec. 18, 2020. [Online]. Available: http://sentic.net/
[28]. Amanpreet Singh, Narina Thakur, Aakanksha Sharma., 'A review of supervised machine learning algorithms' 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), 16 March 2016.
[29]. A. Singh, N. Thakur, and A. Sharma, "A review of supervised machine learning algorithms," in 2016 3rd Inter- national Conference on Computing for Sustainable Global Development (INDIACom), Mar. 2016, pp. 1310–1315.