PROJECT TOPIC: 'WHAT TIME OF THE DAY ARE PEOPLE MORE POSITIVE OR NEGATIVE'

"Diurnal Patterns of Sentiments: A Time-based Sentiment Analysis of Tweets Across the Globe"

Abstract

This study delves into the intricate connection between time of day and human sentiment expression on the global Twitter platform. In an era where millions use social media to share thoughts and emotions, understanding the temporal dynamics of sentiment becomes increasingly important. Sentiment analysis, a crucial tool for detecting and understanding emotions, is central to this research.

The study employs a blend of supervised and unsupervised machine learning techniques, along with deep learning methods featuring Long Short-Term Memory (LSTM) models and attention mechanism, to uncover temporal sentiment trends across diverse global locations.

The findings reveal a consistent diurnal pattern in Twitter sentiment. Mornings are marked by heightened positivity, peaking at 1 a.m., while negativity tends to increase in the evening, reaching its peak at 4 p.m. The results also indicate a gradual rise in pessimism over time, caused by external factors. These insights have significant implications across various domains, including business, governance, mental health, and technology.

This study contributes to sentiment analysis, temporal dynamics, and social media research, demonstrating the effectiveness of advanced machine learning and deep learning algorithms. The LSTM hybrid model achieves an impressive 98% accuracy in sentiment classification. Additionally, the research provides valuable insights into the evolving nature of human emotions and technology's potential in uncovering hidden sentiment patterns.

The implications of this research are wide-reaching, enabling businesses to strategically time content delivery, policymakers to align decisions with public sentiment, and aiding mental health practitioners in mood disorder detection.

CHAPTER 1 INTRODUCTION

I.1 Introduction

Millions of people across the world daily use social media to express their personal views, opinions, interests, and feelings (Mostafa et al., 2022). They tend to communicate by speaking and writing with their families, social groups, work teams, the general public, and vice versa through social networking platforms in their natural languages such as English, thus revealing their emotions within the identified texts (Saha et al., 2021).

There is a common saying that *the pen is mightier than the sword*. This underscores the value of written text and its objectivity and subjectivity. Recognizing and understanding these aspects is vital in daily life. Human subjectivity encompasses feeling, emotion, sentiment, and opinions, as noted by Mostafa et al. (2022), making its detection a concern for researchers and positioning sentiment analysis as a crucial tool. Emotions significantly impact decision-making, social interactions, and well-being.

The recent focus lies on the temporal dynamics and pattern classification of emotions, especially within the realm of social media platforms (Jasim et al., 2022). With platforms like Twitter offering a wealth of user-generated content, there is an opportunity to explore the link between emotions and time of day. This study aims to employ sentiment analysis and modelling techniques to examine when individuals tend to display highly positive or negative emotions.

1.2 Significance and Potential Impact of the Study

The relationship between time of day and mood has been explored in studies such as Kahneman et al. (2004), but limited research has considered specific populations. Quercia et al. (2012) found that global tweets reflect morning and early afternoon happiness shifting to evening pessimism. This research project's impact spans various domains, offering insights into temporal emotional patterns. It aids decision support systems for real-time monitoring (Kaklauskas A., 2023) and provides insights for global decision-makers (Hogenboom et al., 2014) to assess public sentiment at specific times. It's crucial for early mood disorder detection (Jeon et al., 2016), enabling targeted health strategies. Identifying periods of extreme sentiments informs critical case management and supports digital communication-based emotional understanding, enhancing decision-making processes. To achieve this, the subsequent inquiries for the research will be addressed:

- What are the diurnal patterns of sentiment in tweets from individuals in the world?
- What are the times of most positive and negative emotions?
- What are the peak times of both positive and negative emotions within the tweets?
- What are the possible locations of most positive and negative sentiments?

The following hypothesis will be tested in this research project:

- Hypothesis: People are more positive in the morning and more negative in the evening.

CHAPTER TWO LITERATURE REVIEW

2.1 Sentiment Analysis (SA)

For the issue of opinion mining in numerous fields and applications, a wide range of tactics and adaptations have been created. While some academics have published several papers to categorize the polarity of written content and visuals, other researchers have focused on certain tasks: identifying sentiments in words (Hatzivassiloglou & McKeown,1997a; Narayanan et al., 2013), recognizing the source of a viewpoint (Choi et al., 2005), choosing a theme phrase (Nasukawa & Yi, 2003). Sentiment analysis as a process of identifying and extracting the sentiment of a piece of text, such as a tweet, news article, or product review, has been used to track public opinion, identify trends, and make predictions. Three techniques—machine learning, lexicon-based, and hybrid or combination approaches—are often used in sentiment analysis. Dictionary-based methods categorize words using sentiment lexicons like SentiWordNet and VADER without the need for training data. These techniques result in a sentiment rating scale from -1 to 1 (i.e., "Negative," "Positive"), however, they are unable to accurately classify opinions that depend on context. Supervised and unsupervised approaches fall under the category of machine learning. The Gaussian and Support Vector Machine (SVM) are the two most efficient supervised text algorithms (Zainuddin & Selamat, 2014; Prabowo & Thelwall, 2009; Pang et al., 2002). A significant obstacle is that supervised procedures depend on extensive data training, which can be time-consuming. Supervised algorithms must select the right attributes for sentiment analysis. Pre-processing approaches include stemming, lemmatizing, and part-of-speech tagging, for instance (Pang et al., 2002; Tripathy et al., 2016). It should be worth noting that the Hybrid technique combines the two methodologies (Appel et al., 2016). Kolchyna et al. (2015) is similar but differs in the datasets and pre-processing techniques used, such as tokenizers, part-of-speech tagging, etc. The latter combined the two approaches to achieve a better performance when compared to the individual approaches.

2.2 Application of Sentiment Analysis

Sentiment Analysis's scope extends beyond product reviews to social media and personal blogs, detecting emotions (Mantyla et al., 2018). It's crucial for individuals, corporations, and governments. Businesses utilize it for customer insights (Jin et al., 2017; Amplayo R. et al., 2018; Dou Z-Y., 2017; Wu Z. et al., 2018), enhancing consumer confidence (Salinca A., 2017). Governments gauge public sentiment on elections, and policies, like the 2016 US presidential election (Alashri et al., 2016). Sentiment analysis improves recommendations.

2.3 Twitter Sentiment

Recent studies in sentiment analysis (SA) often focus on analyzing Twitter data. According to Pat & Paroubek (2010), a method was proposed using emoticons to label microblogging data, training a multinomial NB classifier with n-grams and POS tags. O'Connor et al. (2010) linked sentiment measurements from word frequencies in tweets to political and consumer confidence polls, revealing public sentiments. Lai (2011) employed Twitter sentiment analysis with SentiStrength to assess a president's performance. Agarwal et al. (2011) combined sentiment and non-sentiment features to study manually annotated tweets, shedding light on emotional nuances among Twitter users.

CHAPTER 3 METHODOLOGY

3.1 Data Source and Selection

The dataset used in this research work consists of Twitter data found at Kaggle collected during April 2009 titled sentiment140_with_location (Μαριοσ Μ. Κ, 2017). The data was chosen to ensure a diverse range of tweets from different locations around the world across the 24 hours of the day.

3.2 Data Privacy Considerations

For privacy and ethics, tweets were anonymized, personal details were removed before analysis (Wu et al., 2017). Only publicly available data was used with no user identification or account linkage attempts. Additionally, the research project obtained appropriate ethical approvals and followed institutional guidelines for handling and storing data. The dataset will be securely stored and accessed only by authorized team members, ensuring confidentiality and data protection.

3.3 Data Preparations

Having imported the necessary libraries for sentiment analysis, the data were subjected to cleaning and preprocessing to remove noise and irrelevant information using a natural language tool kit. This was done to remove the Twitter-specific elements, characters, punctuations, and digits; tweet conversion to lowercase and tokenization were done to ensure uniformity in the data.

3.4 Data Labelling and Statistics

For sentiment classification and evaluation, each tweet in the data was labelled using a sentiment intensity analyzer on a tokenized list of clean tweets (Bird et al., 2009). Comparing the positive and negative polarity scores, the respective sentiment labels (positive (2), negative (0) and neutral (1)) were set. A label encoding task was performed to assign numerical values to the sentiment labels. The dataset statistics were analysed to gain insights into the data as shown below in Figure 1,2.

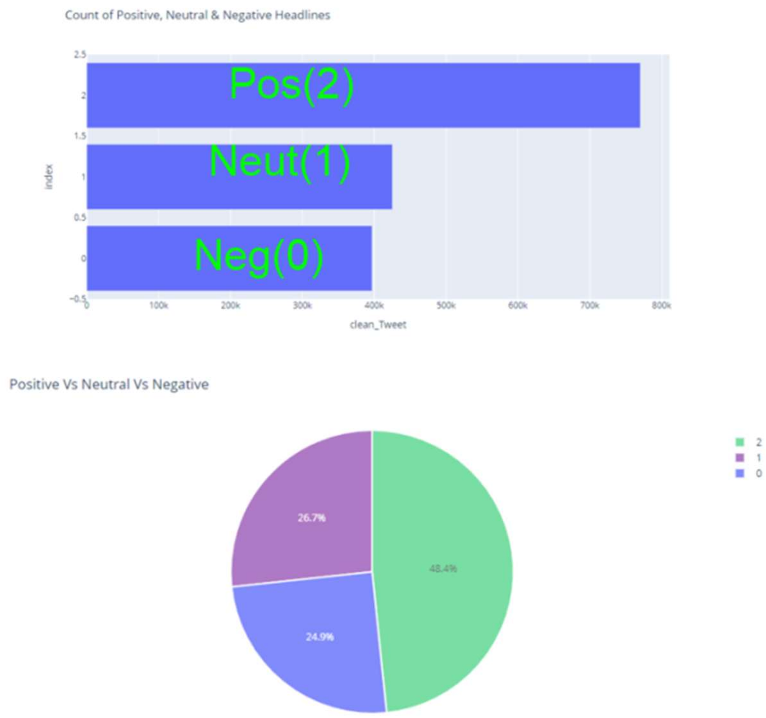


Figure 1: Sentiment Distributions of Cleaned and Labelled Tweets.

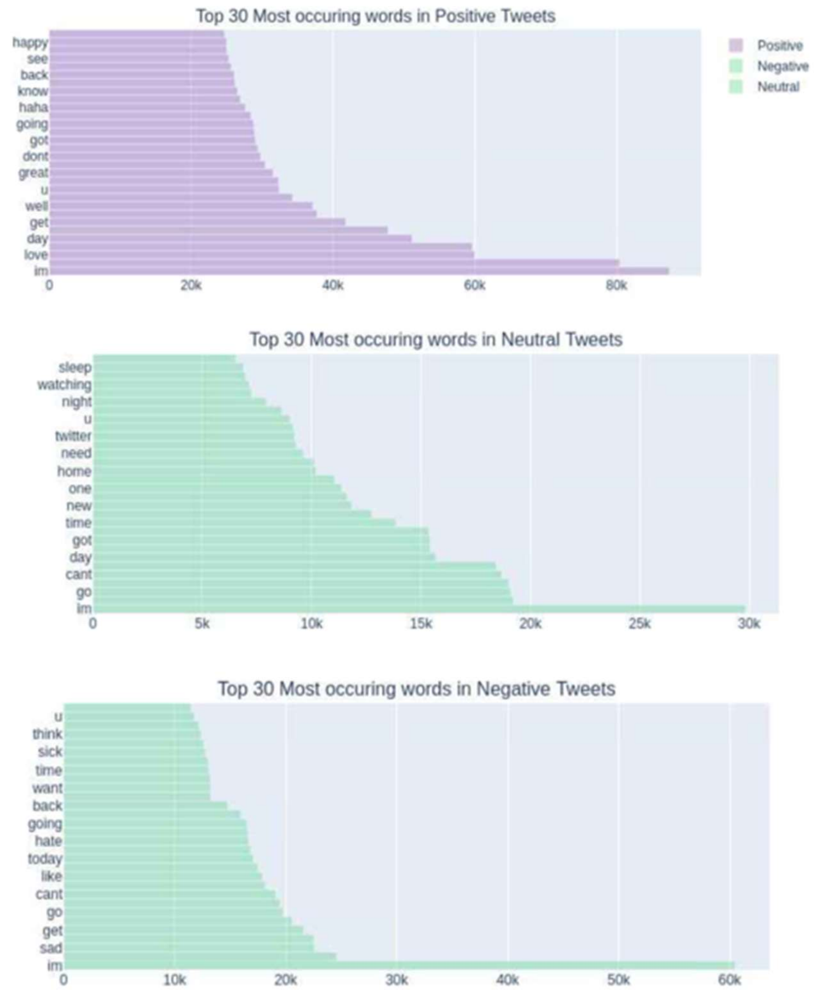


Figure 2: Most Occurring Words Across the Positive-Neutral-Negative Tweets.

Further preprocessing tasks were done to generate a word cloud for the lemmatized words extracted from positive, neutral and negative tweets respectively demonstrated in Figure 3.

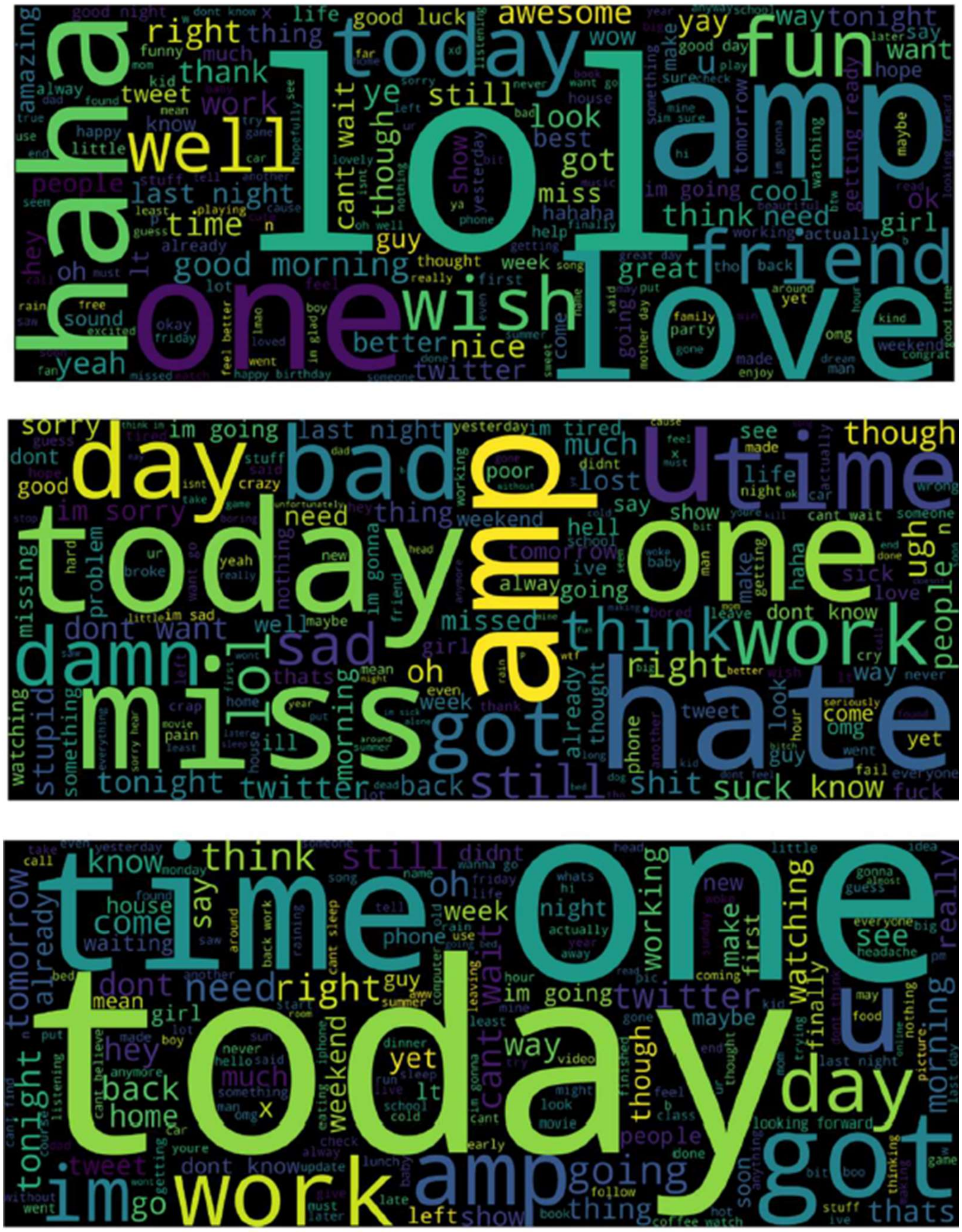


Figure 3: Sentiment-Based Word Clouds for Cleaned Tweets

3.5 Supervised Learning for Sentiment Categorization

In this section, we explore the application of supervised learning techniques for sentiment classification namely logistic regression (LR), Naive Bayes (NB), and an ensemble model combining decision trees, random forests, and k-nearest neighbours (DT+RF+KNN) and Long Short-Term Memory (LSTM) architectures.

3.5.1 Machine Learning Algorithms

Three models were exploited namely Logistic Regression (LR), Multinomial Naïve Bayes (MNB) and Ensemble to evaluate the sentiment classifications. MNB classifier was applied to categorize tweets into sentiment classes based on their word frequencies (Pak & Paroubek, 2010). Also, a logistic regressor was adopted for a similar task. Both models were trained on the split dataset with a test size of 0.2 and a reproducibility random seed of 42. The transform method of the count vectorizer class was initiated for its numerical representation(document-term-matrix). The synthetic minority over-sampling technique (SMOTE) was deployed to balance the class distribution, prevent bias and optimize the performance of the training data.

Using the *make-classification* function, an ensemble model was formed with 1000 data points and 10 input features. This model fuses predictions from three classifiers through majority voting, capitalizing on their complementary strengths (Yang & Ghose, 2010).

Cross-validation was performed successively to evaluate individual model performance on unseen data, revealing mean accuracy and accuracy score standard deviation, showcasing model robustness.

3.5.2 Deep Learning for Sentiment Classification

A Long Short-Term Memory (LSTM) model was further considered for sentiment classification. It is extensively applied in NLP tasks; commonly designed to handle long sequences, address vanishing gradients, relay long-range dependencies, allow gating mechanisms while capturing contextual hints, hence leading to improved precision in predicting sentiment. The model was relatively enhanced with bidirectional LSTM and Gated Recurrent Unit (GRU) forward-backward patterns. An attention mechanism is applied to the model's output to assign different weights to different time steps of the input sequence thus, giving more importance to the most relevant information (Li et al.,2020). The concatenated attention context and final LSTM output were processed through another bidirectional LSTM, dropout, and Dense layer for classification. The Hybrid model was compiled with Adam optimizer, categorical cross-entropy loss, and accuracy metric.

3.6 Evaluation Metrics

The performance of sentiment classification models was examined using standard metrics such as accuracy, precision, recall, and F1-score. The confusion matrix for each model was investigated showing the comparison between the actual labels (ground truth) and predicted labels for the test data.

3.7 Unsupervised Learning Approach

The methodology also implemented basic unsupervised learning techniques to uncover underlying patterns within a sentiment dataset without predefined labels. The techniques included k-means clustering which is a popular clustering algorithm that partitions a dataset into a specified number of clusters to reveal sentiment patterns in tweets as well as offer a structured approach. The elbow method was also implemented to monitor overfitting. This was done by finding the 'elbow point' where improvement slows down thus indicating the best ideal cluster count. Silhouette score was evaluated to measure how well the clusters are separated within the clustering algorithm.

3.8 Time-based Sentiment Analysis

We studied diurnal sentiment patterns in tweets by segmenting them into intervals (weekly, daily, hourly) for sentiment analysis. This reveals fluctuations, peak sentiment counts, extreme sentiments, and recurring patterns. We examined the day's sentiment changes, peak positive/negative periods via 24-hour sentiment visualization. Sentiment partitioning across time intervals and locations was performed to discover periods of extreme sentiments, most sentiment counts, specific times and locations of sentiment-related incidents' occurrence.

For additional sentiment analysis, VADER (Valence Aware Dictionary and Sentiment Reasoner) was initiated to perform some cumulative distribution functions and probability density functions of sentiment values across the tweets. Autocorrelation analyses were done on the tweets to infer some patterns and dependencies present in the time series data.

CHAPTER FOUR RESULTS AND DISCUSSIONS

4.1 Results and Discussions

In this chapter, we delve into the comprehensive results' presentation and analysis of our findings, unravelling the outcomes of our study and engaging in insightful discussions that shed light on the diurnal sentiment patterns exhibited in tweets from around the world.

4.2 Supervised Learning Evaluation

Machine Learning Classifiers: The performance of the models was evaluated, and the results are shown below.

Table 1: Results of the Machine learning models.

S/N	MODEL	ACCURACY (%)	PRECISION (%)	RECALL (%)	F1-SCORE (%)	MEAN ACCURACY (%)	STANDARD DEVIATION
1	MNB	0.86	0.87	0.85	0.86	0.83	0.03
2	LR	0.94	0.93	0.94	0.93	0.86	0.02
3	Ensemble	0.91	0.92	0.89	0.90	0.89	0.17

Considering the accuracy (Proportion of instances classified correctly among the total instances); MNB gave an accuracy of 86%, whereas LR and ensemble models accorded accuracies of 94% and 91% respectively. The precision which represents the ratio of accurately predicted positive instances to all instances predicted as positive had LR as the highest followed by the ensemble and MNB successively. Recall which tells the size of correctly predicted positive instances out of all actual positive instances had 94% from model 2 (LR) proceeded by models 3 (Ensemble) and 1 (MNB) in sequence. For the F1-Score formally known as the harmonic mean of precision and recall; model 2 (LR) provided a good balanced measure across the 3 models with a score of 93%. The models after systematically partitioning the dataset into folds for training and validation. We observed that LR performed best with the lowest standard deviation of 0.02 among other classifiers with a mean accuracy of 86%.

Based on these criteria, **LR** model seems to better than the others. It has the highest accuracy, precision, recall, FI-score and second-highest mean accuracy and the lowest standard deviation among the three models.

DL Model (LSTM): The model initially produced an accuracy of 48% with slight overfitting. There was a tremendous improvement after the introduction of Bidirectional LSTM, GRU layers and attention mechanism respectively hence an accuracy of **98**% was finally obtained as shown below. The model was tested on the text and gave a prediction of **a positive label**, unlike the previous models. Results show a mildly overfitted hybrid model; training accuracy is greater than validation accuracy by a **0.004** difference same applicable to loss. This implies that the model's generalization performance is not severely compromised, thus making it the **most balanced choice**.

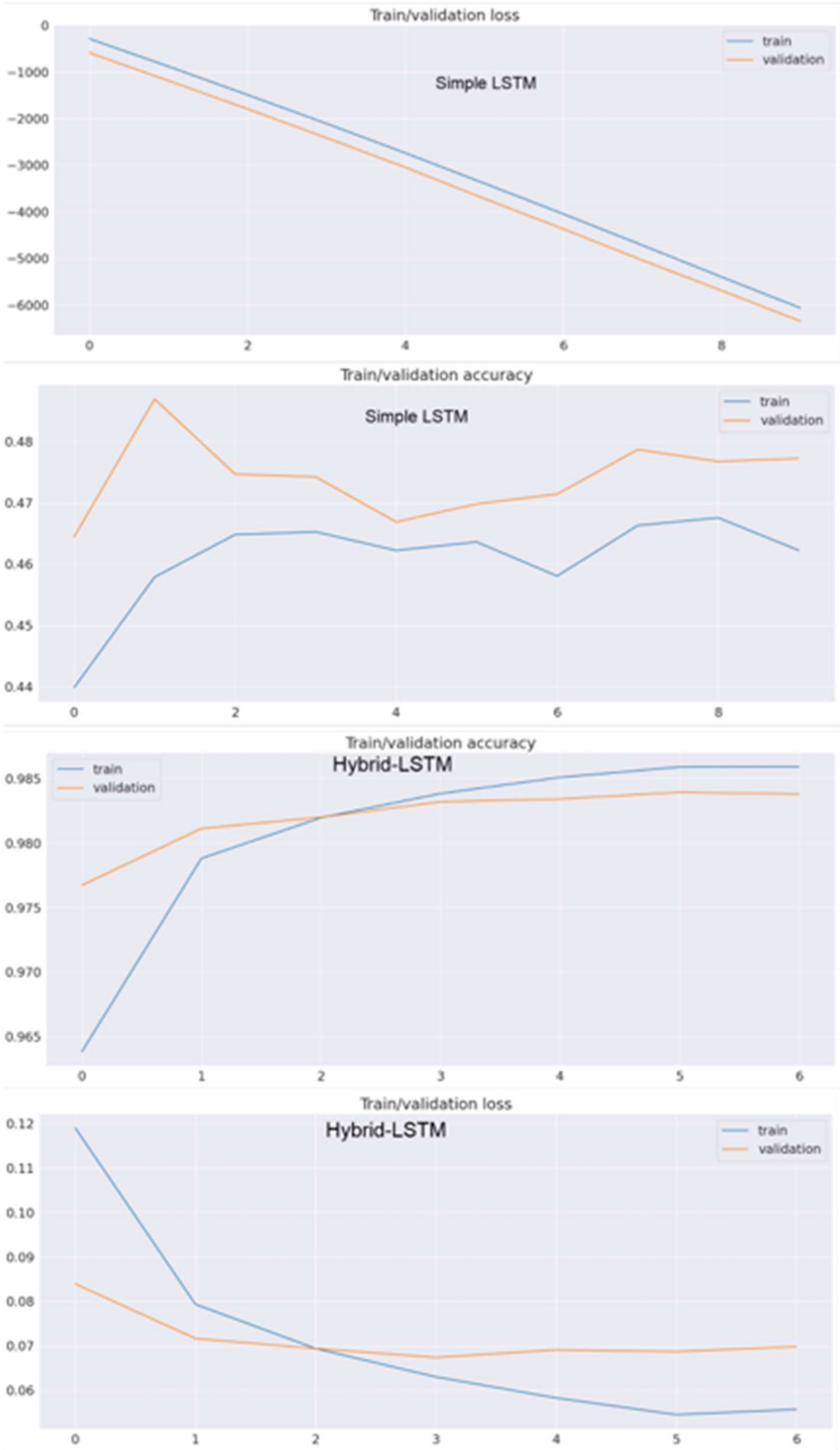


Figure 4: Train-Validation Plots for Simple LSTM and Attention Mechanism-Enhanced LSTM.

4.3 Unsupervised Learning Approach

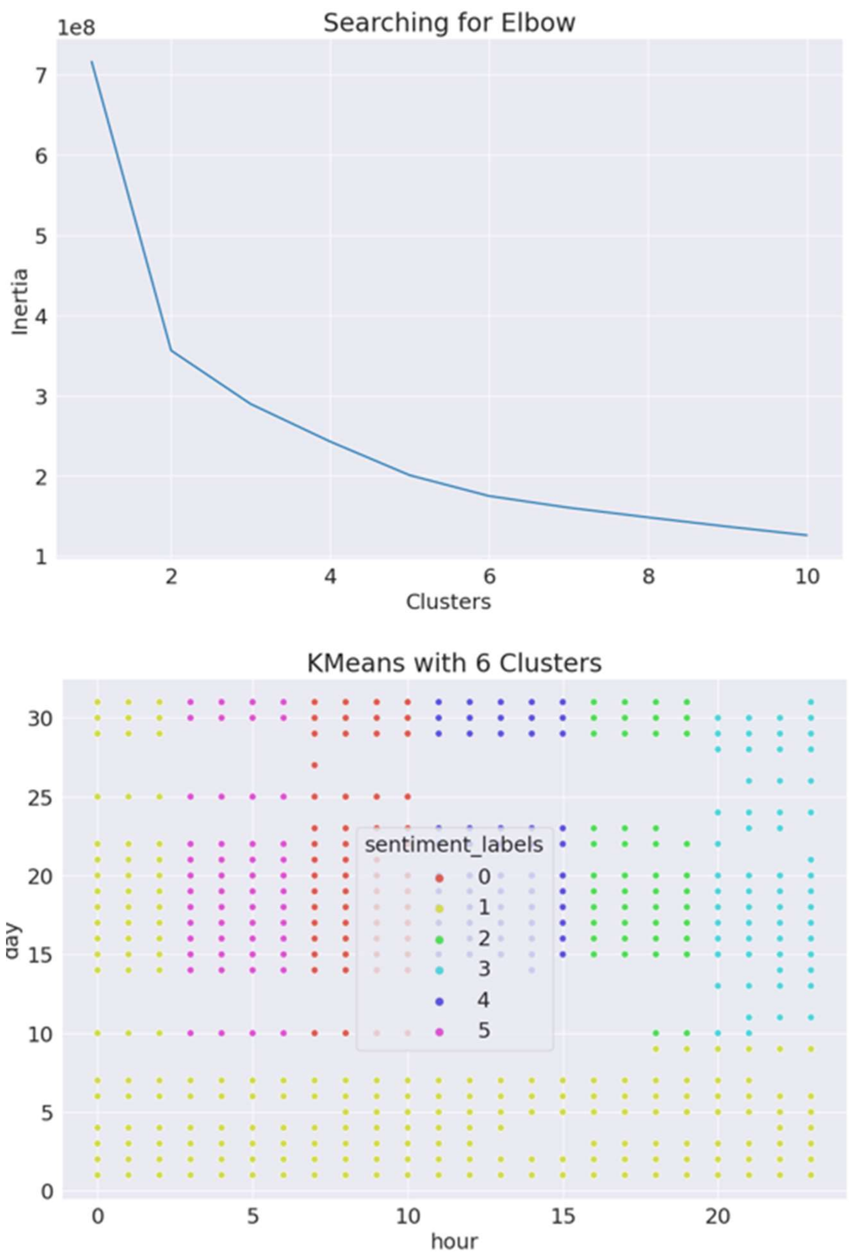


Figure 5: Diagrams showing K-means Clustering and K-methods.

After examining the plots, it seems the model with 6 clusters provided a more appropriate reasonable approach when compared to the one with 2 clusters. The following observations were made:

- Label 0 (Morning) reflects sentiments between 7 a.m. and 10 a.m., contributing insights into early daytime sentiments.
- Label 1 (Continuous) encompasses all 24 hours and 31 days, highlighting a consistent sentiment pattern, invaluable for trend analysis.
- Label 2 (Evening) represents sentiment shifts from 4 p.m. to 7 p.m., offering insights into post-work sentiments.
- Label 3 (Late Night) focuses on sentiments from 8 p.m. to 11 p.m., providing insights into nighttime sentiments.

- Label 4 (Midday) covers sentiments between 11 a.m. and 3 p.m., indicating sentiment dynamics during midday hours.
- Label 5 (early morning) reveals sentiment trends from 3 a.m. to 6 a.m., aiding insights into early risers' sentiments.

These 6 clusters provide a comprehensive view of daily sentiment patterns in the tweets, each corresponding to specific times and days. These highlight varied sentiment patterns across different times and days, aiding sentiment analysis. Identifying time-based clusters reveals trends, prompting investigations into their causes. Different sentiment labels across time and days offer insights into sentiment dynamics. For instance, a 24-hour, full-month coverage cluster (Label 1) suggests a *consistent sentiment pattern*, potentially influenced by *non-time-related factors* like news events or trends.

This understanding is essential in sentiment analysis as it enables the recognition of time-based trends and factors impacting sentiments. It empowers businesses, researchers, and policymakers to tailor strategies, content delivery, and decision-making to align with these patterns for more effective engagement and communication (Syakur et al., 2018).

The silhouette score of **0.629** indicates well-separated clusters thus, validating the chosen approach. The Davies Bouldin score of **0.508** reaffirms distinct clusters and strong cohesion within them.

4.4 Time-Based Sentiment Analysis

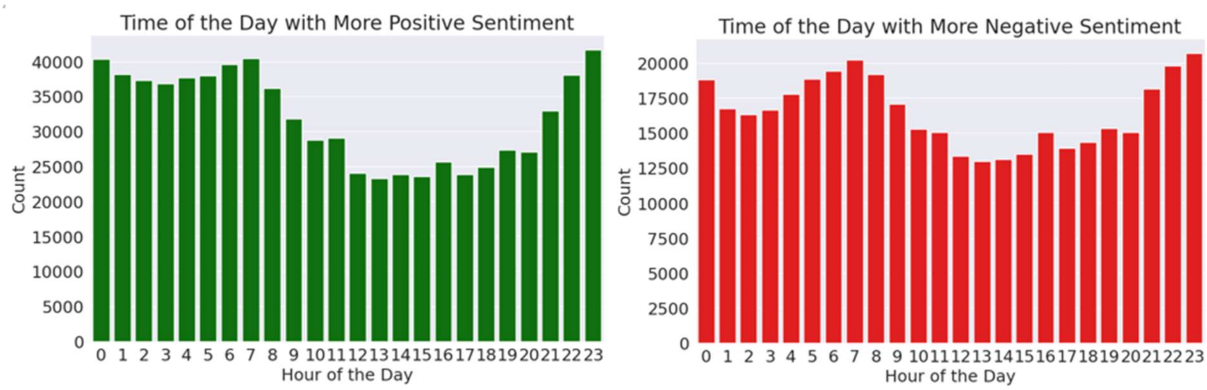


Figure 6: Sentiment Count Distribution Across 24 Hours of the Day.

The analysis examined the outcomes of sentiment grouping and counting across morning, afternoon, evening, and night sessions.

In various daily sessions (morning, afternoon, evening, and night), negative sentiments occurred 125,223, 68,025, 43,641, and **160,143** times. Neutral sentiments appeared 139,415, 69,729, 43,331, and **180,158** times. Positive sentiments were observed 243,754, 120,331, 76,078, and **330,172** times. Particularly, **night** sessions saw increased online expression of thoughts and moods, leading to prevalent sentiments (Figure 6).

The analysis shows that hour 23 (11:00 PM) has the highest count of positive sentiment, indicating content that elicits positivity during late evenings. Hour 23 also sees the highest count of negative sentiment, possibly due to the content's nature or context. Hour 7 (7:00 AM) records the highest count of neutral sentiment, suggesting common neutrality in early mornings. These patterns imply temporal trends in sentiment distribution. The shared peak at hour 23 suggests content evoking **strong emotions**, and the neutral peak at hour 7 could relate to morning engagement dynamics. From the above diagram, it could be stated that the positive sentiment tends to be more dominant throughout the sessions of the day.

Extra efforts were made to detect the locations with the most positive, neutral and negative sentiment counts considering the given dataset as highlighted in Figure 7. *Mexico* is the most positively dominated country while *Cuba* tends to be most pessimistic.

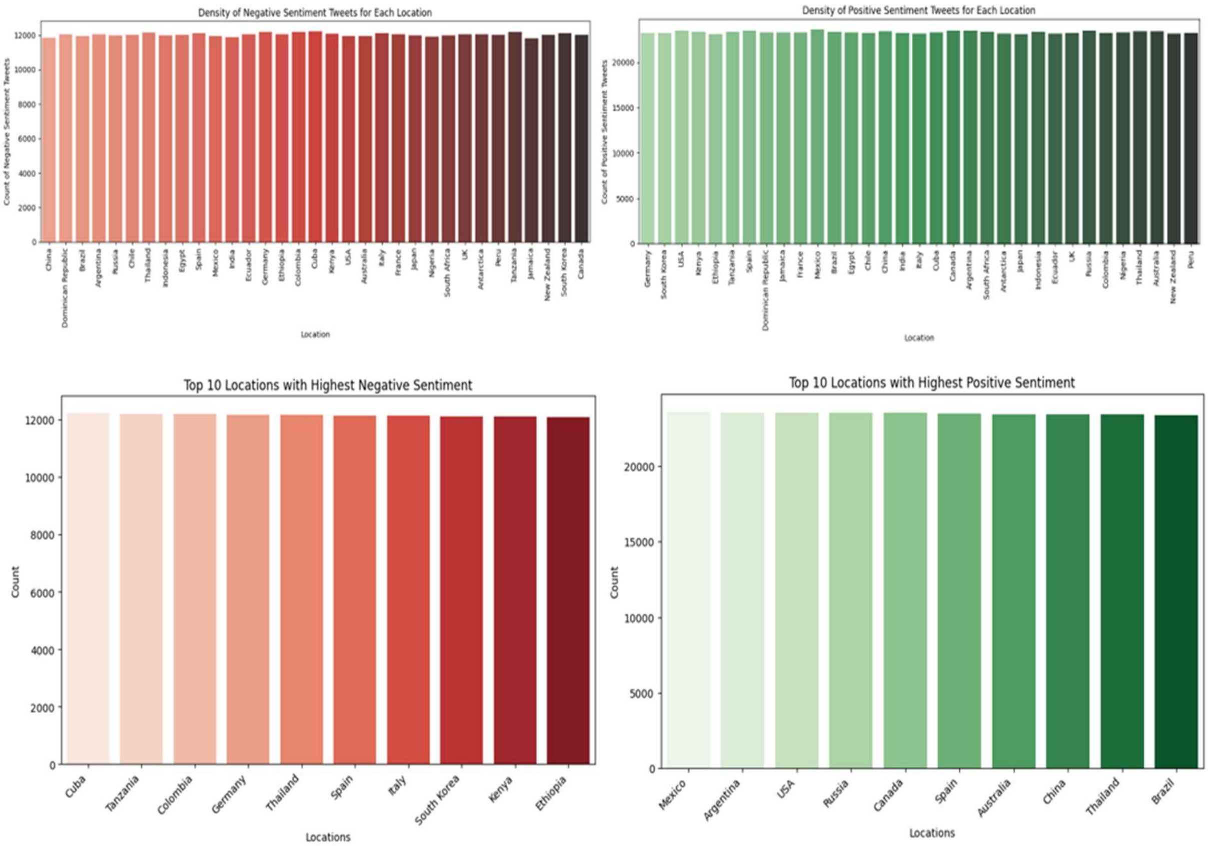
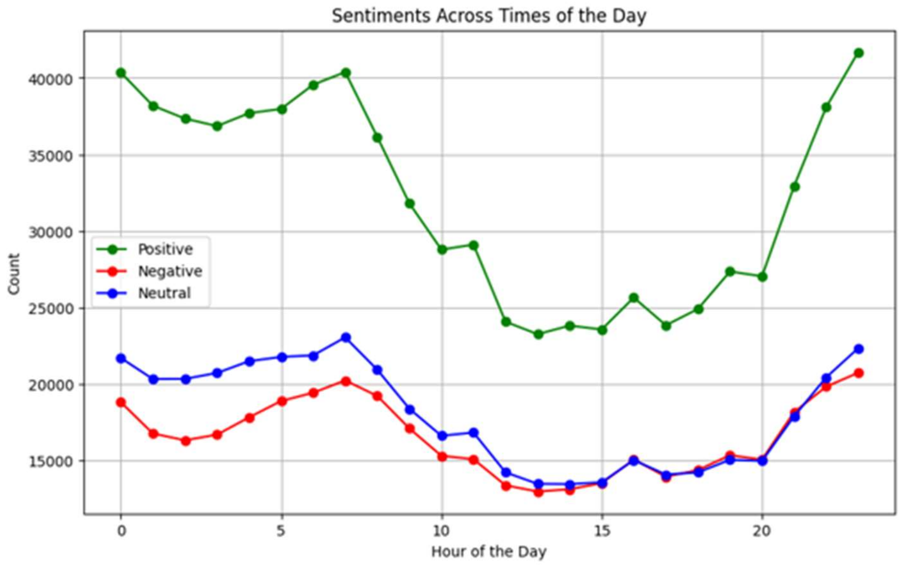


Figure 7: Sentiment Distribution by Locations.

Peak Positive Sentiment Hour (Hour 1) - The analysis indicates that the hour with the highest average positive sentiment is hour 1. Peak Negative Sentiment Hour (Hour 16) - The analysis also shows that the hour with the lowest average sentiment (highest average negative sentiment) is hour 16. These imply that, on average, users tend to express more positive and negative sentiments when interacting with the content or data at these hours. Content creators or marketers can use this information to time their content delivery for better engagement, considering that positive sentiment is highest at hour 1 and negative sentiment is highest at hour 16. This *reaffirms* Burch et al. (2009) which proposed that employees are more optimistic in the morning and early afternoon than in the late afternoon and evening. Figure 8 further illustrates sentiment trends, indicating **morning positivity** and **evening negativity.**



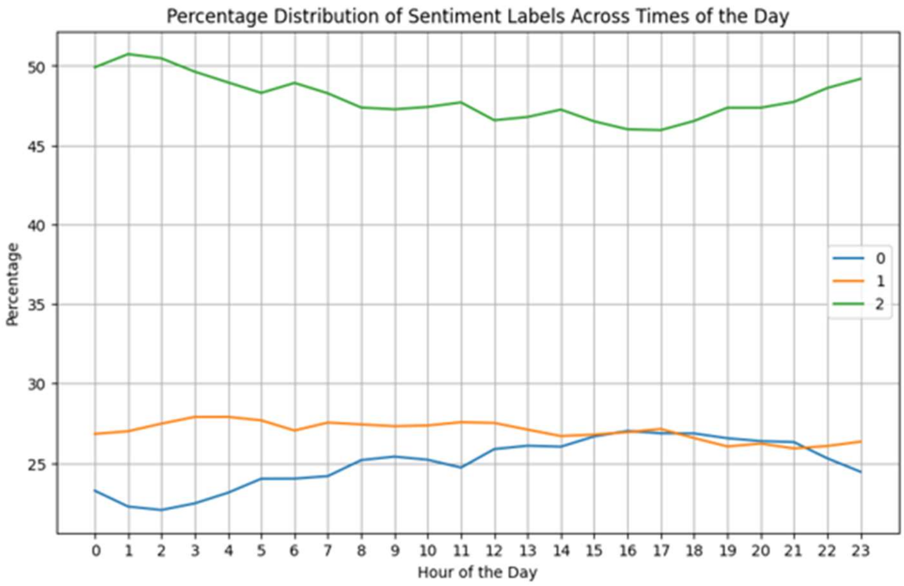


Figure 8: Sentiment trends and Percentage Distributions Across 24 hours of the Day.

If hour 1 consistently yields elevated positive sentiment in its content, exploring and comprehending the contributing elements or subjects could prove advantageous. Likewise, for hour 16, delving into the content types that possibly trigger more negative sentiment offers insights. These findings hint at potential user behaviour trends that impact emotional reactions during distinct hours.

Table 2: Hourly Distributions of the Labelled Sentiments

Hour	Negative Label (0)	Neutral Label (1)	Positive Label (2)	Total	0_percentage	1_percentage	2_percentage
0	18810	21698	40357	80865	23.260991	26.832375	49.906635

1	16760	20321	38187	75269	22.266803	26.997834	**50.734034**
2	16316	20331	37344	73993	22.050735	27.476923	50.469639
3	16682	20716	36855	74256	22.465525	27.898082	49.632353
4	17813	21485	37697	76999	23.134067	27.902960	48.957779
5	18882	21766	37975	78628	24.014346	27.682251	48.297044
6	19424	21869	39559	80858	24.022360	27.046180	48.924040
7	20228	23042	40384	83661	24.178530	27.542104	48.270998
8	19214	20927	36146	76295	25.183826	27.429058	47.376630
9	17094	18380	31804	67287	25.404610	27.315826	47.266188
10	15300	16609	28780	60699	25.206346	27.362889	47.414290
11	15081	16822	29106	61020	24.714848	27.568010	47.699115
12	13370	14218	24065	51665	25.878254	27.519597	46.578922
13	12967	13470	23252	49702	26.089493	27.101525	46.782826
14	13117	13453	23810	50394	26.028892	26.695638	47.247688
15	13512	13571	23560	50658	26.672984	26.789451	46.507955
16	15059	15017	25644	55736	**27.018444**	26.943089	46.009760
17	13934	14075	23834	51860	26.868492	27.140378	45.958349
18	14371	14217	24897	53503	26.860176	26.572342	46.533839
19	15336	15039	27347	57741	26.559983	26.045617	47.361494
20	15057	14966	27036	57079	26.379229	26.219801	47.365931
21	18155	17882	32927	68985	26.317315	25.921577	47.730666
22	19822	20426	38080	78350	25.299298	26.070198	48.602425
23	20728	22333	41689	84773	24.451181	26.344473	49.177214

The table above reveals that the peak positive sentiment, approximately **51%**, appeared at **1:00 a**.m. **(hour 1)** during the day. Conversely, the highest negative sentiment percentage, **27.01**%, was noted at **4:00 p.m. (hour 16)**.

In essence, these findings **validate** the previously mentioned hypothesis: **people exhibit greater positivity in the morning and more negativity in the evening**.

Mean Absolute Deviation **(0.64)**: This signifies the average dispersion or range of sentiment scores around their mean, possibly indicating a broader range of sentiments.

Distribution of Sentiment values and cumulative distribution of sentiment across the tweets

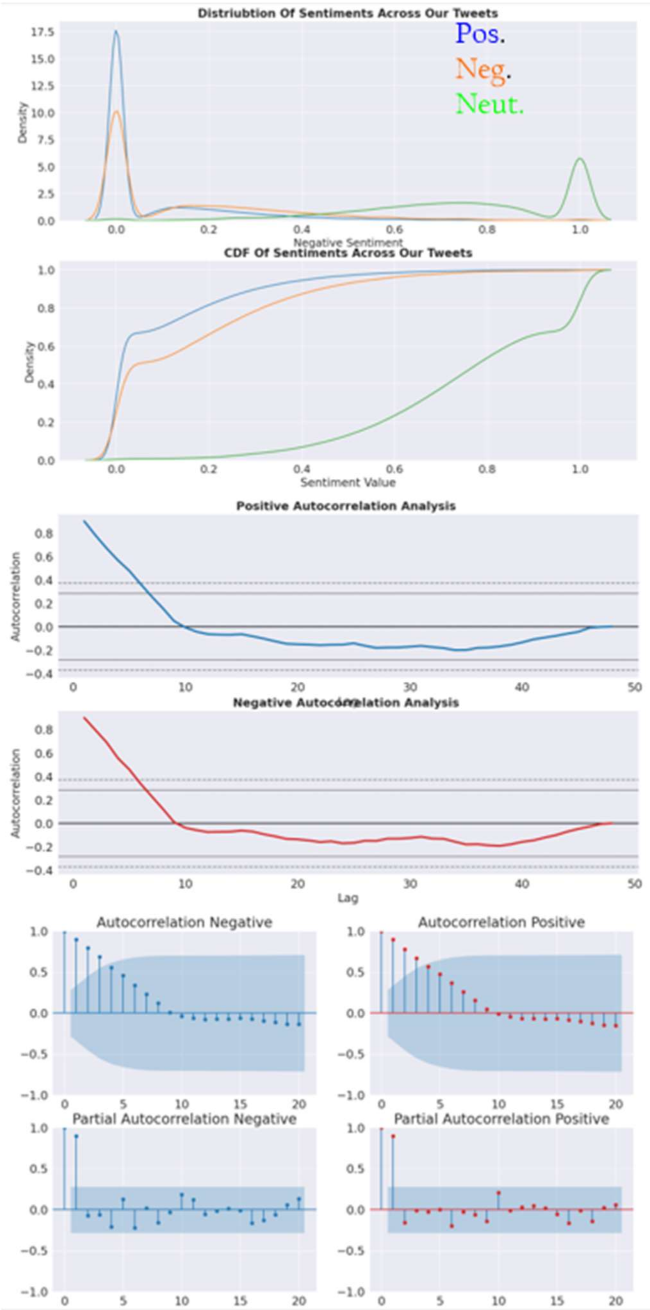


Figure 9: Plots Showing Cumulative Distributions and Autocorrelations.

The sentiment distributions resemble a normal curve, with similar negative and positive patterns suggesting comparable sentiment strengths, differing primarily in counts. Positive sentiment prevails, while neutral sentiment stands out distinctly from positive and negative sentiments. This insight enhances comprehension of sentiment proportions and their dataset-wide distribution.

Table 3: Sentiment Partitions Across Times

Sentiment Labels	Partition_1_mean	Partition_2_mean	Partition_3_mean	Partition_1_SD	Partition_2_SD	Partition_3_SD
Positive sentiment	0.1577	0.1580	0.1437	0.2028	0.2021	0.1935
Negative sentiment	0.0821	0.0816	0.0919	0.1483	0.1481	0.1552

The average negative sentiment rises gradually across partitions, indicating increased negativity over time. A growing standard deviation reflects greater sentiment diversity, offering insights into evolving sentiment trends and potential interventions based on observed patterns.

Autocorrelation across Tweet Sentiments

The Autocorrelation (ACR) and Partial Autocorrelation (PACR) plots show no significant correlation between current sentiment and past values at lag x. Recent sentiment changes are not closely linked to sentiment x periods ago, indicating that external factors influence sentiment fluctuations over time. Analysing sentiment trends requires accounting for other factors, as past sentiment at the specified lag weakly impacts current values.

CHAPTER FIVE CONCLUSIONS

5.1 Summary of Findings

The study explored the intricate link between sentiment expression and time of day across various locations. It reveals a well-defined diurnal sentiment pattern in tweets, with positive sentiment prevailing in the morning and diminishing towards the evening, while negative sentiment exhibits an opposite trajectory, intensifying as the day progresses. Using supervised and unsupervised methods, the research scrutinized sentiment patterns, distributions, and dynamics across hours, days, and places, unveiling insights into evolving sentiments. Hybrid-LSTM excelled in sentiment analysis, capturing contexts within the tweets. Sentiment clusters disclosed distinct temporal patterns—early morning positivity, evening negativity, and day-long trends. Geographically, specific locations exhibit elevated positive and negative sentiments respectively, informing tailored strategies. The study also underscored complex sentiment dynamics shaped by factors beyond history, including news, culture, and personal experiences.

5.2 Contributions to the Field

The study significantly advances sentiment analysis, temporal dynamics, and social media research by exploring sentiment's link with time. It highlights the efficacy of machine and deep learning algorithms, with Hybrid-LSTM as a balanced choice for sentiment classification. It enhances our understanding of evolving human emotions and technology's potential to reveal hidden patterns. The innovative approach, incorporating supervised learning, unsupervised clustering, and temporal analysis, sets a precedent for comprehensive sentiment analysis. Also, the location-based sentiment analysis contributes to a more comprehensive understanding of global sentiment expression.

5.3 Implications and Future Research

The study holds implications across various sectors. Businesses can strategically time content delivery based on diurnal sentiment trends, while governments can align policies with public sentiment dynamics. The research also offers the potential for mental health practitioners to detect mood disorder indicators through sentiment trends. Future research avenues include refining sentiment analysis algorithms for context-specific emotions, exploring cultural influences on diurnal sentiment patterns, and investigating external factors' impact on sentiment trends.

In summary, this research enhances our grasp of daily emotional fluctuations, offering fresh perspectives on studying digital-era human behaviours and actionable insights for communication, decisions, and well-being.

REFERENCES

Agarwal, A., Xie, B., Vovsha, I., Rambow, O. and Passonneau, R.J., 2011, June. Sentiment analysis of Twitter data. In *Proceedings of the workshop on language in social media (LSM 2011)* (pp. 30-38).

Amplayo, R.K., Kim, J., Sung, S. and Hwang, S.W., 2018. Cold-start aware user and product attention for sentiment classification. *arXiv preprint arXiv:1806.05507*.

Appel, O., Chiclana, F., Carter, J. & Fujita, H. (2016) A hybrid approach to the sentiment analysis problem at the sentence level. Knowledge-Based Systems, 108 110-124.

Bird, S., Klein, E. and Loper, E., 2009. *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.".

Burch, J.B., Tom, J., Zhai, Y., Criswell, L., Leo, E. and Ogoussan, K., 2009. Shiftwork impacts and adaptation among health care workers. *Occupational medicine*, *59*(3), pp.159-166.

Choi, Y., Cardie, C., Riloff, E. and Patwardhan, S., 2005, October. Identifying sources of opinions with conditional random fields and extraction patterns. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing* (pp. 355-362).

Dou, Z.Y., 2017, September. Capturing user and product information for document level sentiment analysis with deep memory network. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 521-526).

Hatzivassiloglou, V. and McKeown, K., 1997, July. Predicting the semantic orientation of adjectives. In *35th annual meeting of the association for computational linguistics and 8th conference of the european chapter of the association for computational linguistics* (pp. 174-181).

Hogenboom, A., Heerschop, B., Frasincar, F., Kaymak, U. and de Jong, F., 2014. Multi-lingual support for lexicon-based sentiment analysis guided by semantics. *Decision support systems*, *62*, pp.43-53.

Jasim, Y.A., Saeed, M.G. and Raewf, M.B., 2022. Analyzing Social Media Sentiment: Twitter as a Case Study. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, *11*(4), pp.427-450.

Jeon, H.J., Baek, J.H., Ahn, Y.M., Kim, S.J., Ha, T.H., Cha, B., Moon, E., Kang, H.J., Ryu, V., Cho, C.H. and Heo, J.Y., 2016. Review of cohort studies for mood disorders. *Psychiatry Investigation*, *13*(3), p.265.

Jin, Y., Zhang, H. and Du, D., 2017. Incorporating positional information into deep belief networks for sentiment classification. In *Advances in Data Mining. Applications and Theoretical Aspects: 17th*

Industrial Conference, ICDM 2017, New York, NY, USA, July 12-13, 2017, Proceedings 17 (pp. 1-15). Springer International Publishing.

Kaklauskas, A., 2023. Internet and Biometric Web Based Business Management Decision Support.

Kahneman, D., Krueger, A.B., Schkade, D.A., Schwarz, N. and Stone, A.A., 2004. A survey method for characterizing daily life experience: The day reconstruction method. *Science*, *306*(5702), pp.1776-1780.

Kolchyna, O., Souza, T.T., Treleaven, P. and Aste, T., 2015. Twitter sentiment analysis: Lexicon method, machine learning method and their combination. *arXiv preprint arXiv:1507.00955*.

Lai, P., 2010. Extracting strong sentiment trends from Twitter. *Nlpstanfordedu*.

Li, W., Qi, F., Tang, M. and Yu, Z., 2020. Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification. *Neurocomputing*, *387*, pp.63-77.

Mäntylä, M.V., Graziotin, D. and Kuutila, M., 2018. The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*, *27*, pp.16-32.

Μαριοσ, M. K. (2017). Sentiment140 dataset with 1.6 million tweets: Updated Kaggle version. Retrieved from https://www.kaggle.com/datasets/kazanova/sentiment140 [Accessed July 17, 2023].

Mostafa, R., Mehedi, M.H.K., Alam, M.M. and Rasel, A.A., 2022, December. Bidirectional LSTM and NLP Based Sentiment Analysis of Tweets. In *International Conference on Soft Computing and Pattern Recognition* (pp. 647-655). Cham: Springer Nature Switzerland.

Nasukawa, T. and Yi, J., 2003, October. Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2nd international conference on Knowledge capture* (pp. 70-77).

O'Connor, B., Balasubramanyan, R., Routledge, B. and Smith, N., 2010, May. From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the international AAAI conference on web and social media* (Vol. 4, No. 1, pp. 122-129).

Pang, B., Lee, L. and Vaithyanathan, S., 2002. Thumbs up? Sentiment classification using machine learning techniques. *arXiv preprint cs/0205070*.

Pang, B. and Lee, L., 2008. Opinion mining and sentiment analysis. *Foundations and Trends® in information retrieval*, *2*(1–2), pp.1-135.

Pak, A. and Paroubek, P., 2010, May. Twitter as a corpus for sentiment analysis and opinion mining. In *LREc* (Vol. 10, No. 2010, pp. 1320-1326).

Prabowo, R. and Thelwall, M., 2009. Sentiment analysis: A combined approach. *Journal of Informetrics*, *3*(2), pp.143-157.

Saha, A., Al Marouf, A. and Hossain, R., 2021, June. Sentiment analysis from depression-related user-generated contents from social media. In *2021 8th International Conference on Computer and Communication Engineering (ICCCE)* (pp. 259-264). IEEE.

Salinca, A., 2017. Convolutional neural networks for sentiment classification on business reviews. *arXiv preprint arXiv:1710.05978*.

Syakur, M.A., Khotimah, B.K., Rochman, E.M.S. and Satoto, B.D., 2018, April. Integration k-means clustering method and elbow method for identification of the best customer profile cluster. In *IOP conference series: materials science and engineering* (Vol. 336, p. 012017). IOP Publishing.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D. and Kappas, A., 2010. Sentiment strength detection in short informal text. *Journal of the American society for information science and technology*, *61*(12), pp.2544-2558.

Tripathy, A., Agrawal, A. and Rath, S.K., 2016. Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications*, *57*, pp.117-126.

Quercia, D., Ellis, J., Capra, L. and Crowcroft, J., 2012, February. Tracking" gross community happiness" from tweets. In *Proceedings of the ACM 2012 conference on computer supported cooperative work* (pp. 965-968).

Wu, L., Li, J., Hu, X. and Liu, H., 2017, June. Gleaning wisdom from the past: Early detection of emerging rumors in social media. In *Proceedings of the 2017 SIAM international conference on data mining* (pp. 99-107). Society for Industrial and Applied Mathematics.

Wu, Z., Dai, X.Y., Yin, C., Huang, S. and Chen, J., 2018, April. Improving review representations with user attention and product attention for sentiment classification. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32, No. 1).

Yang, S. and Ghose, A., 2010. Analyzing the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence?. *Marketing science*, *29*(4), pp.602-623.

Zainuddin, N. and Selamat, A., 2014, September. Sentiment analysis using support vector machine. In *2014 international conference on computer, communications, and control technology (I4CT)* (pp. 333-337). IEEE.