AMS 3910 – Mathematical Modelling Project:

Weekly Patterns in Social Media Engagement.

Instructor

Shafiqul Islam

Author

Chizurum Ewelike – 340157

Contents

Introduction

In today's digital age, social media platforms have become integral to our daily lives, offering communication, entertainment, and information dissemination avenues. With billions of users worldwide, these platforms host an immense amount of data, providing a rich source for analysis and exploration. One fascinating aspect of social media dynamics is the temporal patterns of user engagement, which fluctuate over time, exhibiting distinct trends across different days of the week.

Understanding these weekly patterns in social media engagement is crucial for individuals, businesses, and content creators alike. Identifying the optimal times to post content can significantly impact its reach, visibility, and effectiveness in capturing audience attention. However, the complexities inherent in social media behaviour, influenced by diverse factors such as user demographics, cultural norms, and platform algorithms, pose challenges for accurate prediction and analysis.

This project seeks to delve into the mathematical modelling of weekly patterns in social media engagement, aiming to unravel the underlying dynamics governing user interactions on platforms like Instagram, Facebook, Twitter, and TikTok, striving to discern recurring trends, peak activity periods, and influential factors driving weekly engagement fluctuations by harnessing data analytics techniques and mathematical modelling approaches.

The insights gained from this study can inform content creators, marketers, and social media strategists in optimizing their posting schedules and content strategies for enhanced audience engagement and impact. The methodologies and findings generated by this research contribute to the broader understanding of social media dynamics and pave the way for further exploration into the intricacies of online social interactions.

Data Collection & Analysis

Data Collection:

For the data collection phase, it is paramount to acknowledge the prevalent privacy concerns linked with accessing personal social media engagement data. Acquiring such data necessitates a private research process involving interventions and surveys to obtain individuals' consent and corresponding data. However, given the scale and scope of this project, the route of personal interventions and surveys was not pursued. Instead, alternative methodologies were explored to acquire the requisite dataset. Despite the initial intent to collect real-world social media engagement data, accessing such datasets directly took time due to privacy constraints and resource limitations, so an alternative approach was adopted.
Extensive online searches were conducted to locate publicly available datasets related to social media engagement. Unfortunately, this effort yielded limited success, as suitable datasets did not align with the project's specific requirements.

In response to these challenges, creative solutions were pursued to generate a dataset suitable for analysis. Leveraging artificial intelligence (AI) tools such as ChatGPT and the randomized datasets I created, synthetic data was generated to simulate social media engagement patterns. While this approach may only replicate real-world data partially, it provides a foundational basis for exploring and understanding potential trends and patterns in social media engagement. It is essential to note that while the dataset utilized for this research may be unconventional and potentially unrealistic compared to real-world data, rigorous methodologies were employed to ensure its relevance and validity within the study context. It is an illustrative example that the dataset can consist of various numerical representations while retaining relevance and applicability for research purposes.

The synthetic dataset enables the exploration and analysis of potential trends and patterns in social media engagement without being constrained by the limitations of real-world data availability. Using a synthetic dataset underscores the adaptability and resilience of mathematical modelling approaches in capturing the essence of dynamic systems like social media engagements.

Data Analysis:

The numerical data extracted from the dataset serves as the foundation for the mathematical aspects of the model. These numbers are essential for quantifying and analyzing the relationship between the day of the week and social media engagements. Graphing this relationship allows for a visual representation of the patterns and trends observed, providing valuable insights into the dynamics of social media engagement throughout the week.

It is crucial to highlight the significance of using data-driven approaches instead of relying solely on popular research results or anecdotal evidence from influencers or marketers. While such insights may offer valuable guidelines, they need more substantiation of empirical evidence derived from rigorous data analysis. For instance, a standard recommendation from influencers or marketers suggests that consistent posting and optimal timing are vital strategies for boosting social media engagement. However, without concrete data analysis and mathematical modelling to support these claims, placing complete trust in such research findings becomes challenging.

Therefore, in this project, emphasis is placed on leveraging a mathematical modelling technique to validate or refute these assertions. By examining the engagement data from social media platforms, we aim to uncover patterns that corroborate or challenge existing assumptions. Through statistical analysis and mathematical modelling, we seek to provide a robust framework for understanding the factors influencing social media engagement and optimizing posting strategies.

It is essential to consider various factors influencing social media engagement beyond the day of the week. These factors could include the type of content posted, the demographics of the audience, the timing of posts, and external events or trends. Incorporating these variables into the analysis allows for a more comprehensive understanding of the dynamics at play. It enables the development of more nuanced and effective strategies for maximizing social media engagement.

Mathematical Modelling

TABLE 1

Day of the Week	Likes	Comments	Shares	Platform
Monday	150	30	20	INSTAGRAM
Tuesday	180	35	25	TWITTER
Wednesday	160	30	20	FACEBOOK
Thursday	175	37	27	TIKTOK
Friday	185	40	30	THREADS

GRAPH 1

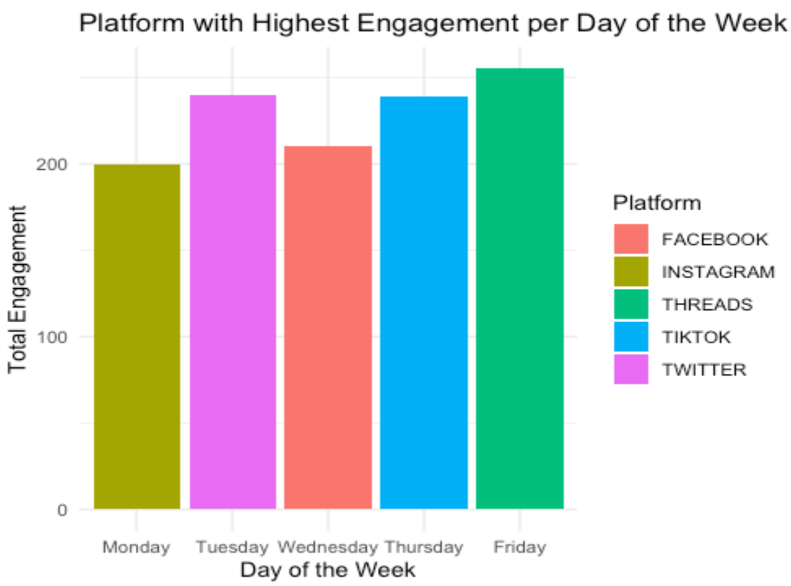


Table 1 presents an illustrative example of a social media engagement dataset, showcasing the volume of engagements across diverse social media platforms. Graph 1 visually represents the trends in engagement across these platforms. For instance, it allows us to observe that on Mondays, this user typically receives the highest engagement on social media. It is crucial to acknowledge the dynamic nature of social media engagements. A few or numerous engagement reports may not yield a consistent pattern due to many influencing factors, particularly across different platforms. In this model, I will utilize a Recurrent Neural Network (RNN), a powerful type of neural network specialized in processing sequential data, to analyze and predict the temporal dynamics of social media engagements.

A Recurrent Neural Network (RNN) is an artificial neural network designed to effectively process and analyze sequential data, making it particularly suited for tasks involving time-series data or sequences. Unlike traditional feedforward neural networks, which process inputs independently and do not have memory, RNNs maintain an internal state, allowing them to capture temporal dependencies within sequential data.

The foundation of RNNs lies in their ability to loop the information from one step of a sequence to the next. This looped structure enables RNNs to retain information about past inputs while processing current inputs, effectively incorporating context from earlier steps into later ones. This mechanism makes RNNs well-suited for language modelling, speech recognition, time-series prediction, and natural language processing. The concept of recurrent neural networks dates to the 1980s, with seminal work by Paul Werbos and others. However, it was only in the mid-1990s and early 2000s that significant advancements were made in developing and training RNNs.

Notably, the Long Short-Term Memory (LSTM) architecture, introduced by Sepp Hochreiter and Jürgen Schmidhuber in 1997, addressed the issue of vanishing gradients during training, making it possible to train RNNs on long sequences effectively. The LSTM architecture and other variants, such as Gated Recurrent Units (GRUs), have since become fundamental in developing modern RNNs, enabling them to handle complex sequential data with improved performance and efficiency. Today, RNNs are widely used across various domains, ranging from natural language processing and machine translation to time-series forecasting and gesture recognition.

RNNs were explicitly designed to capture and model sequential dependencies and temporal dynamics in data, making them well-suited for analyzing dynamic systems like social media engagements. While it may not perfectly capture all nuances of real-world dynamics, it can provide valuable insights and estimates to inform decision-making and strategy development in social media marketing, content creation, and optimizing user engagement.

Mathematical Formulation

A basic RNN, as represented by the formula below, is well-suited for processing sequential data like social media engagements. By leveraging the hidden state, activation functions, and learned parameters, RNNs can effectively capture the temporal dependencies and complex patterns inherent in sequential data, making them valuable tools for analyzing and predicting engagement trends on social media platforms.

$$h_t = f(W_{hx}x_t + W_{hh}h_{t-1} + b_h)$$

Where,

- h_t : This represents the hidden state of the RNN at time step *t*. It captures the memory or internal representation of the network at that particular time step. In the context of social media engagements, the hidden state can be thought of as encoding information about the engagement patterns observed up to that point in time. This hidden state serves as a condensed representation of the input sequence, enabling the network to retain information about past interactions while processing current ones.

- f : This denotes the activation function applied to the weighted sum of inputs and the previous hidden state. Common activation functions include the hyperbolic tangent (tanh) or rectified linear unit (ReLU). These activation functions introduce nonlinearity into the model, allowing the network to capture complex relationships and patterns in the sequential data. In the context of social media engagements, nonlinear activation functions are essential for capturing the diverse and dynamic nature of user interactions across different platforms.

- W_{hx} and W_{hh}: These are weight matrices that determine how much influence the current input (*xt*) and previous hidden state (*ht*−1) have on the current hidden state (*ht*). They are learned parameters during the training process. The weights *Whx* control how the current input influences the hidden state, while the weights *Whh* govern the influence of the previous hidden state. In the context of social media engagements, these weight matrices allow the network to learn and adapt to the sequential patterns observed in user interactions over time.

- b_h: This is the bias vector added to the weighted sum. It allows the model to adjust for unexplained shifts or offsets in the data. The bias term provides flexibility to the model, enabling it to account for systematic differences in engagement levels across different platforms or days of the week. In the context of social media engagements, the bias term helps the network capture platform-specific or time-specific trends that may not be captured by the input data alone.

Mathematical Analysis

TABLE 2

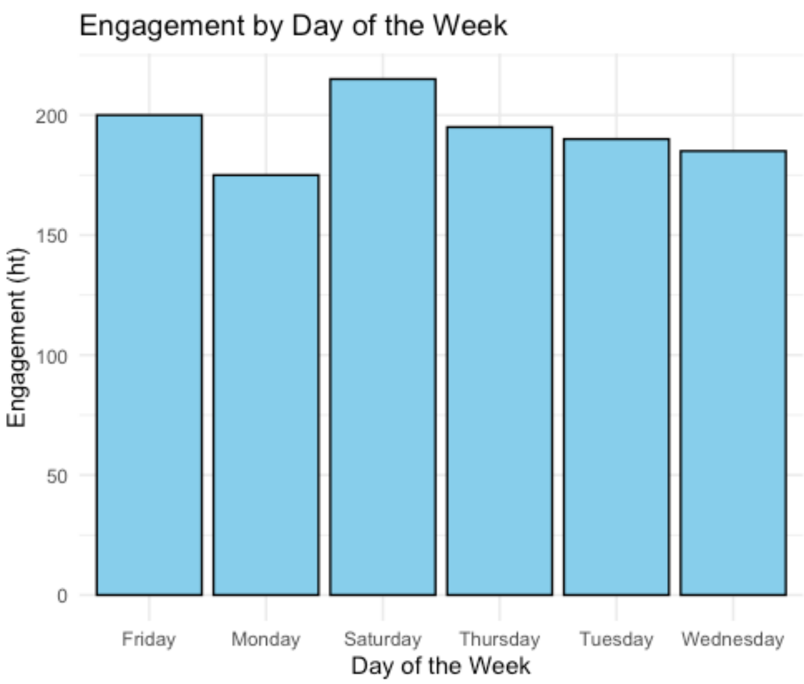
Day of the Week	**Likes**	**Comments**	**Shares**	***ht***
Monday	100	20	15	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$
Tuesday	110	25	20	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$
Wednesday	105	22	17	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$
Thursday	115	28	22	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$

Friday	120	30	25	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$
Saturday	125	35	27	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$

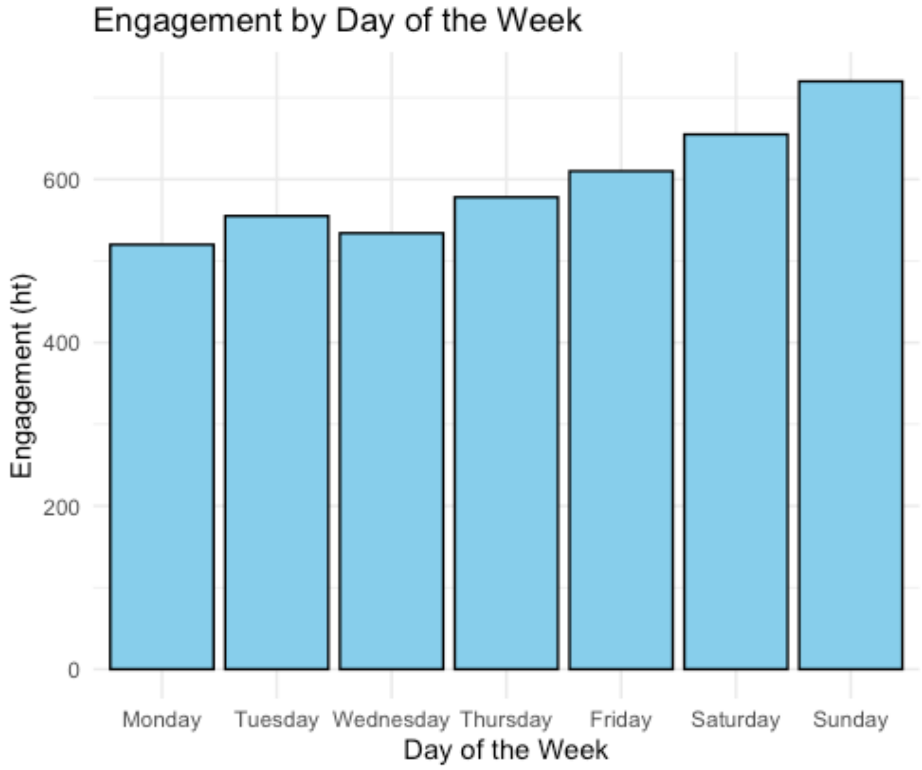
TABLE 3

Day of the Week	**Likes**	**Comments**	**Shares**	***ht***
Monday	250	40	30	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$
Tuesday	270	45	35	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$
Wednesday	260	42	32	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$
Thursday	280	48	38	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$
Friday	290	50	40	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$
Saturday	300	55	45	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$
Sunday	310	60	50	$f(W_{hx}x_1 + W_{hh}h_0 + b_h)$

GRAPH 2



GRAPH 3



Graphs 2 and 3 depict the correlation between engagements (*ht*) and the days of the week, capturing the patterns sought after in this study. According to prevalent research findings, social media engagements tend to be subdued on Mondays. The outcomes derived from these two visual representations offer compelling evidence for this assertion. A consistent observation emerges through the utilization of two distinct randomized datasets within the mathematical model: Mondays exhibit the lowest engagement rates.

Pros and Cons:

Pros:

1. **Predictive Capabilities:** By leveraging mathematical modelling techniques like RNNs, the model can forecast future engagement levels based on historical data, aiding in proactive decision-making.

2. **Adaptability:** The model can accommodate various factors influencing social media engagement, allowing for a comprehensive analysis of dynamic interactions over time.

3. **Innovative Approach:** Utilizing synthetic datasets generated through AI tools demonstrates creativity in overcoming data limitations, enabling exploration and analysis of engagement patterns.

Cons:

1. **Accuracy Estimation:** While social media engagements are dynamic and fluctuate, the model provides an estimated accuracy for predicting engagement patterns. This estimation offers valuable guidance despite the inherent variability in social media dynamics.

2. **Complexity:** Mathematical modelling approaches like RNNs require expertise in data science and computational techniques, posing challenges for implementation and interpretation for those without specialized knowledge.
3. **Ethical Considerations:** Accessing and utilizing social media data raises privacy and ethical concerns, necessitating careful handling of sensitive information and adherence to ethical guidelines.

4. **Generalization Challenges:** The dataset used for analysis may be specific to particular contexts or platforms, limiting the generalizability of the model's findings to broader social media landscapes. Adjustments may be required to apply the model effectively across diverse scenarios and platforms.

Improvements

1. **Real-world Data Collection:** Although synthetic datasets were used due to data availability constraints, collecting real-world social media engagement data would significantly improve the model's accuracy and generalizability. Collaborating with social media platforms or accessing anonymized datasets could provide more representative data for analysis.

2. **Feature Engineering:** Incorporating additional features beyond likes, comments, and shares could enrich the dataset and enhance the model's predictive capabilities. Factors such as post type (e.g., image, video, text), posting time, user demographics, and external events could be considered to capture more nuanced patterns in engagement.

3. **Temporal Analysis:** Expanding the analysis to include longer time intervals, such as monthly or seasonal patterns, could provide a broader perspective on social media engagement dynamics. Understanding how engagement trends evolve over extended periods can offer valuable insights for long-term strategy planning.

4. **Fine-tuning and Optimization:** Conducting thorough experimentation and optimization of model parameters, including learning rates, batch sizes, and regularization techniques, can improve the model's performance. Fine-tuning the model on relevant metrics, such as mean squared error or accuracy, can lead to better results.

5. **Cross-validation and Validation Metrics:** Implementing cross-validation techniques and using appropriate validation metrics, such as mean absolute error or root mean squared error, can provide a more reliable assessment of the model's performance. Ensuring robust validation procedures can enhance confidence in the model's predictive capabilities.

6. **Interpretability and Explainability:** Enhancing the interpretability of the model's predictions by incorporating techniques such as feature importance analysis or attention mechanisms can help stakeholders understand the factors driving engagement patterns. Providing transparent and interpretable insights can facilitate decision-making and strategy refinement.

Conclusions

In conclusion, this project has offered valuable insights into the temporal dynamics of social media engagement patterns, leveraging mathematical modelling techniques to analyze and predict weekly engagement fluctuations. By examining data from diverse social media platforms and utilizing innovative approaches such as synthetic dataset generation, the project has provided a foundation for understanding the complex interplay between various factors influencing user interactions online.

The model's pros include its predictive capabilities, adaptability to diverse factors, and innovative approach to overcoming data limitations. The model can inform decision-making processes for content creators, marketers, and social media strategists by forecasting engagement levels and accommodating multiple variables. There are also notable cons, including challenges in accuracy estimation, complexity in implementation, and ethical considerations regarding data privacy. Despite these challenges, the model's estimated accuracy provides valuable guidance for navigating the dynamic landscape of social media engagements.

Several improvements can enhance the model's effectiveness and robustness. These include collecting real-world data, incorporating additional features for richer analysis, expanding temporal analysis to longer time intervals, optimizing model parameters, implementing rigorous validation procedures, and enhancing stakeholder interpretability.

By addressing these improvements, the model can become more accurate, generalizable, and actionable, empowering stakeholders to optimize their social media strategies effectively.

References

1. https://www.clickworker.com/ai-glossary/recurrent-neural-networks/#:~:text=Recurrent%20neural%20networks%20were%20first,by%20John%20Hopfield%20in%201982. - History of RNN

2. AI Tools – ChatGPT – Creation of Datasets

3. AI Tools – Grammarly – Editing grammar errors.

4. Graphs - RStudio