Decoding the Impact on Influencer Marketing: An Analysis of Sponsored Content Performance

Ping-Lun Lai
Department of Statistics
University of Michigan
Ann Arbor, MI
alanlai@umich.edu

Yonnie Chan
Department of Statistics
University of Michigan
Ann Arbor, MI
yonniech@umich.edu

Nicole Chen
Department of Information
University of Michigan
Ann Arbor, MI
nicolesy@umich.edu

Jingjie Wan
Department of Information
University of Michigan
Ann Arbor, MI
iriswan@umich.edu

***Abstract*—This research utilizes Instagram posts as data to investigate the impact of influencers' sponsored content on electronic word-of-mouth (eWOM) in social commerce. We aim to uncover key factors driving post-virality and audience engagement within the context of influencer marketing. We implement exploratory data analysis, natural language processing, and hypothesis testing, investigating useful features and identifying meaningful metrics that can effectively predict post popularity. This study perform comprehensive analysis to provide actionable insights for businesses, enabling them to optimize their influencer marketing strategies. Furthermore, we conducted several experiments to build predictive models utilizing machine learning, and neural networks techniques, with a mean square error of 0.00106, enabling influencers to better foreseen their post's possible popularity.**

***Index Terms*—Social media, Instagram, trend analysis, popularity prediction, natural language processing, machine learning, neural network.**

The following questions are addressed in this research, i.e., (R1) How does the presence of sponsored content by influencers on Instagram affect electronic word-of-mouth (eWOM) dynamics in social commerce? (R2) What are the key factors affecting the ER? (R3) What is the best method for popularity prediction?

Our study will identify predictive metrics that drive post popularity, encompassing a broad spectrum of factors such as influencer characteristics, post-content attributes, and post-image. By harnessing these predictive metrics, we seek to develop robust models that can accurately forecast the impact of influencer marketing efforts, thus empowering businesses to optimize their strategies and enhance their digital presence effectively. The contributions of this research are as follows:

I. Introduction

In today's digital age, the landscape of commerce has undergone a significant transformation, shifting towards social media where influencers hold substantial sway over consumer behavior. This shift has led to an unprecedented surge in social commerce, blending the realms of social media and e-commerce. Among the plethora of social platforms, Instagram emerges as a dominant force in influencer marketing, boasting a high Engagement Rate [1] and visually captivating content, making it an ideal platform for sponsored content.

The rise of social networks and user-generated content has significantly transformed advertising as consumers increasingly rely on word-of-mouth (WOM) [2]. Electronic word-of-mouth (eWOM) is the digital form of WOM where consumers share opinions on products or services online [3]. Social media influencers, especially on platforms like Instagram, have built reputations and personas by promoting values and products [4]. For example, Instagram influencers often share fashion tips, makeup tutorials, and personal stories to engage followers and endorse products. Therefore, brands seek to collaborate with influencers to showcase products and engage with audiences [5].

The following questions are addressed in this research, i.e., (R1) How does the presence of sponsored content by influencers on Instagram affect electronic word-of-mouth (eWOM) dynamics in social commerce? (R2) What are the key factors affecting the ER? (R3) What is the best method for popularity prediction?

Our study aims to explore the impact of influencers' sponsored content on eWOM in social commerce, uncovering the potential drivers behind post-virality and audience engagement, with Instagram posts as the primary data source. Utilizing natural language processing (NLP) and hypothesis testing, we compare sponsored and non-sponsored posts, investigate a wide range of features from post information(i.e. likes, comments, captions, hashtags, post images), then extract meaningful insights. Through machine learning and neural networks techniques, we identify key metrics for predicting post popularity to capture post features with high engagement rates. The contributions of this research are as follows:

- Identify predictive metrics and explore key factors that drive post popularity on Instagram.
- Provide feasible suggestion through comprehensive trend and hashtags analysis.
- Develop neural networks method, combining text and image decoder, to accurately predict post popularity.

The result of this study is beneficial for both brand and Instagram influencers. This study aims to provide actionable insights for businesses looking to optimize their influencer marketing strategies. In addition, the extensive features engineering work also help general public to understand deeper about the mechanism of post popularity on Instagram.

II. Related Work

A. Electronic Word-of-Mouth (eWOM)

In 2014, Barreto defined WOM as an "oral or written communication process, between a sender and an individual or group of receivers, regardless of whether they share the same social network, to share and acquire information, on an informal basis." as a recent framework of word-of-mouth communication. The main concepts of WOM are structured to informal communication between individuals, such as senders and receivers, regarding products, brands, organizations, or experiences, aimed at sharing and acquiring information without any commercial motives [6].

With the emergence of online social networks, WOM has transformed into electronic word-of-mouth (eWOM) through the new communication format. An online consumer review serves a dual purpose, acting as both an informant and a recommender. In its role as an informant, it offers user-centered product details, while as a recommender, it delivers suggestions from past consumers in the form of eWOM [7]. Social networks on social media platforms provide consumers with friends' and influencers' information [8]. eWOM has been an advertisement strategy to leverage merchandise' benefit of brand building, customer acquisition, and retention [9].

B. Social Media Influencers Attributes

Social media influencers reflect a part of their followers' beliefs and values. Influencers' followers choose their followees by comparing their identity with the influencers' image, which is shaped by various followers' backgrounds [10]. Interaction frequency between influencers and followers on platforms is significantly related to followers' loyalty [11]. For example, participation in the influencer's community encourages meaningful interaction with the influencer and this phenomenon is measurable by an increased engagement rate. Mutually, influencer images in followers are constructed by the influencers' personal thoughts, information, and feelings intentionally shared contents and those images create imaginary intimacy with the influencers, strong connections, and emotional exchanges [12]. Our study aligns with the following definition for SMIs: influencers social media users who, acting as third-party individuals, have attained micro-celebrity status with substantial followings on social platforms and hold sway over their audience. This attained status enables them to convey marketing messages for brands and shape consumer opinions [13].

C. User Engagement Rate on Social Media

Previous studies have extensively explored various facets of Instagram's popularity, employing statistical analyses and predictive modeling. In the Thomas VL. et al. [14] study in 2020, popularity metrics for Instagram posts typically include likes, engagement level, intrinsic image appeal, growth of likes, and categorized output (e.g., popular vs. unpopular). These studies predominantly rely on metadata such as follower count, post count, captions, and hashtag count. Some studies have also incorporated diverse attributes and methods such as image content analysis and image quality assessment [15]. When a post becomes highly popular, it expands the audience for each hashtag it utilizes. Consequently, other posts employing these same hashtags will reach a broader audience, resulting in increased social popularity. This underscores the advantage of recommending hashtags with a substantial audience, indicating that leveraging trending hashtags can elevate social engagement [16]. Notably, hashtag count has been integrated into prediction models to examine its impact on popularity. While past research has primarily focused on hashtag count, recent studies have underscored the significance of hashtag choice in boosting likes.

D. Token Transformers and Pre-training for NLP

Transformers, as introduced by Vaswani et al. [17], have sparked a revolution in natural language processing (NLP) tasks by enabling efficient parallel processing of text data. This architectural innovation has resulted in significant improvements across various NLP domains. A pivotal advancement in NLP lies in pre-training large language models (LLMs) on vast text corpora. These models acquire general language representations that can be finely tuned for specific tasks. Among the prominent pre-training approaches are masked language modeling (MLM).

In masked language modeling, a portion of the input text is masked, and the model is trained to predict the masked tokens. BERT [18] stands out as a successful example of an MLM-based pre-trained model, achieving state-of-the-art performance across a wide range of NLP tasks. On the other hand, denoising auto-encoding involves presenting the model with corrupted input text and tasking it with reconstructing the original clean text. This approach encourages the model to learn robust representations.

While Transformers have excelled in NLP, recent efforts have explored their adaptation for computer vision tasks. The Vision Transformer (ViT), introduced by Dosovitskiy et al. [19], has shown promising results in image recognition. ViT breaks down an image into patches, processes them through a Transformer encoder akin to text, and utilizes a classification head for image categorization.

Our approach focuses on leveraging the Instagram influencers dataset, utilizing pre-trained ViT for image understanding, and BERT for textual analysis. By combining these powerful models, we aim to capture a more comprehensive representation of the post's content, leading to more accurate predictions of user engagement rates.

III. Data

A. Data Collection

Access to the Instagram influencers dataset was granted through the studies "Multimodal Post Attentive Profiling for Influencer Marketing" [20] and "Discovering Undisclosed Paid Partnership on Social Media via Aspect-Attentive Sponsored Post Learning" [21] by requesting the [1]Instagram Influencer

[1]https://sites.google.com/site/sbkimcv/dataset/instagram-influencer-dataset?authuser=0

Dataset & Influencer Brand Dataset from the SeungBae Kim. Our dataset includes sponsored and non-sponsored Instagram posts, comprising 1,601,074 posts from 38,113 influencers and 26,910 brands from 2012 to 2019. Sponsored posts were identified either through hashtags or branded content tools.

The dataset includes two types of files, post metadata and image files. Post metadata files are in JSON format and contain the following information: caption, user tags, hashtags, timestamp, sponsorship, likes, comments, etc. (Detailed information about the data structures of those features can be found in Appendix B) Image files are in JPEG format and the dataset contains 12,933,406 image files since a post can have more than one image file. Also, we obtain a JSON-Image_mapping file that shows a list of image files that correspond to post metadata.

To refine the scope of our research while maintaining a substantial dataset, we have chosen to concentrate on posts authored by influencers belonging to the 'fashion' category, which is the largest category with 638,741 posts from 9,290 influencers. Overall, the dataset encompasses 88,691 sponsored posts[2] and 550,050 non-sponsored posts[3] with partial post images. For this research, an influencer is defined as a user who has created at least one sponsored post within our dataset. We found that 1,011 influencers had not contributed any sponsored posts within the dataset. To narrow our focus to sponsored content, we filtered the main posts of DataFrame to exclusively include 425,611 social media posts from users identified as influencers. Consequently, this filtering process yielded a refined dataset consisting of 605,963 posts for subsequent analysis.

B. Data Preprocessing

1) Engagement Rate Calculation: To assess the popularity, our targeted feature, of each post within our dataset, we employed an ER metric. This metric is calculated for each post by dividing the sum of the number of likes and the number of comments by the corresponding influencer's follower count. Mathematically, this can be expressed as:

$$ER = (\#Likes + \#Comments)/\#Followers$$

Following this formula, we integrated the information on post engagement[4] from the post-DataFrame with the influencer data[5] from the influencer DataFrame based on their shared column, and JSON file name. This process allowed us to calculate the ER for each post-influencer combination within the merged dataset.

2) Outlier Detection and Removal: During the data cleaning process, we encountered two data points with an ER of 0, accompanied by zero values for both likes and comments. Such occurrences often indicate potential anomalies, such as system-generated errors, spam attempts, or posts containing inappropriate content. Due to the ambiguity surrounding the validity of these data points, we decided to exclude them from the dataset. This action aimed to mitigate the potential influence of outliers or invalid entries on subsequent analyses.

Moreover, to ensure the robustness of our analysis, we utilized a box plot visualization to identify potential outliers in terms of ER. The visualization highlighted eight data points with ER values exceeding 2, far beyond the range of other data points. Seven of these outliers originated from non-sponsored posts, while only one belonged to a sponsored post.

The presence of such high ER values can stem from two primary factors: a remarkably low follower count for the influencer or the presence of inauthentic engagement practices, such as scams or bot activity. While the detection of scams was beyond the scope of our research, we chose to remove these outliers due to their potential to skew subsequent analyses. Given the limited number of outliers in proportion to the dataset size and the rationale for their removal, we believe this decision will have minimal impact on the overall trends observed within the data.

IV. Methodology

A. Hypothesis Tests on the Effect of Sponsorship

In the dynamic landscape of social media marketing, influencer collaborations have become a cornerstone for brands seeking to reach wider audiences and build authentic connections. However, the effectiveness of sponsored content in driving engagement remains a subject of debate. To shed light on this matter, we embarked on a comprehensive statistical analysis to explore whether sponsorship affects influencer posts' popularity, as measured by ER. By examining a merged dataset of influencer and post-data, we employed a series of statistical tests to unravel the intricate relationship between sponsorship and ER, considering various dimensions and potential confounding factors.

To assess the impact of sponsorship on the effectiveness of influencer marketing, we conducted a rigorous statistical analysis using a merged dataset comprising influencer and post information. Our primary focus was to determine whether sponsored posts exhibited a significant difference in ER compared to non-sponsored posts[6]. We began by employing a T-test to compare the average ER between these two groups, revealing a p-value of 0.049. This finding suggests a statistically significant, albeit marginal, effect of sponsorship on ER when considering the entire dataset.

However, recognizing the limitations of the T-test in accounting for potential confounding variables, we proceeded to conduct several Chi-square tests. These tests aimed to examine the relationship between sponsorship and ER while considering factors such as individual influencers, time of posting, and the presence of specific brand hashtags.

Firstly, to examine for each influencer whether sponsorship would affect the ER of their posts, we grouped data by

[2] approximately 13.89%

[3] approximately 86.11%

[4] Likes and Comments

[5] Follower Count

[6] H0: There is no significant difference in the average engagement rate (ER) between sponsored and non-sponsored posts; H1: Sponsored influencer posts have a significantly different average engagement rate compared to non-sponsored posts.

'Username' and 'sponsor', calculated the average ER for each group, and then performed Chi-square test of independence on the grouped data. The resulting P-value is 1.0, showing that sponsorship wouldn't affect the ER of a post at all for a specific influencer.

Similar Chi-square tests were then performed to examine the relationship between sponsorship and other variables. In all instances, the resulting p-values either equaled 1 or approached 1, suggesting that sponsorship had no significant influence on the ER of posts, regardless of the posting timeframe or the presence of specific hashtags.

While our initial analysis did not reveal statistically significant differences in engagement rate between sponsored and non-sponsored posts, we opted to further explore potential temporal trends within this metric.

B. Exploratory Data Analysis - Trend Analysis

In our exploration to uncover temporal trends within the ER value, we undertook a meticulous approach known as time series feature engineering. This methodological step encompassed the extraction of pertinent time features from the post-timestamps embedded within our dataset. These features were carefully selected to encapsulate various temporal dimensions, including the year, month, day, hour, and weekday. By integrating these time features into our analysis, we aimed to discern any discernible patterns or fluctuations in ER over different temporal scales.

1) Analysis of Engagement Rate by Yearly: The scatter plot in Fig. 1 visually represents the distribution of post likes about their timestamps. General Upward is observed in the ER over time. Earlier posts, around 2013-2014, exhibit considerably lower ER compared to those in later years, especially as we approach 2019. As the ER increases, the spread or variance of the data points also seems to expand. This suggests that while the overall ER is rising, individual post-performance becomes more varied in later years. While not explicitly evident, the plot hints at potential seasonality or cyclical patterns within the data. More granular analysis with the extracted features (month, day, etc.) would be needed to confirm this.

2) Analysis of Average Engagement Rate by Month and Sponsorship: The line plot in Fig. 2 illustrates the average ER for posts across each month, differentiating between sponsored and non-sponsored content. The ER starts relatively high in January, dips to its lowest point around June, and then rises again towards the end of the year. While both categories follow a similar general trend, sponsored posts consistently maintain a higher average ER compared to non-sponsored ones throughout the year. Both sponsored and non-sponsored posts experienced peak ER in January. This could be attributed to New Year's resolutions, fresh starts, or increased social media engagement during the holiday season. The lowest ER for both categories occurs in June. This might be due to summer holidays, decreased social media usage, or competition with other forms of content during this period. Interestingly, the ER for both categories converges around June, suggesting that external factors influencing engagement might overshadow the impact of sponsorship during this month.

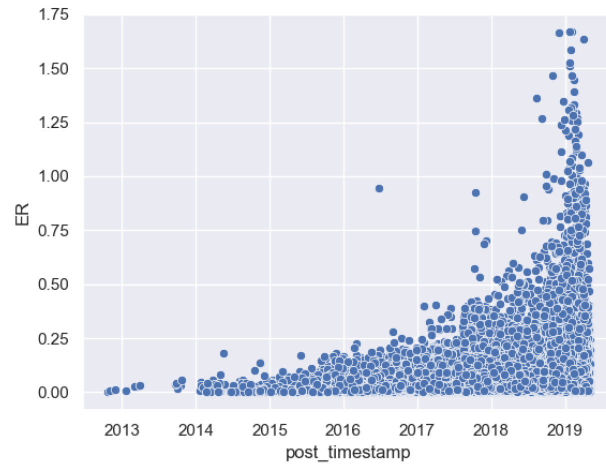

Fig. 1. Distribution of ER rate over time

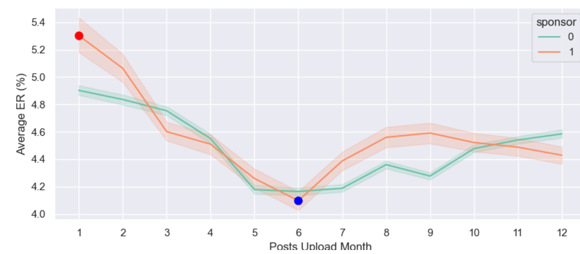

Fig. 2. Average ER by month (75% confidence interval)

3) Analysis of Average Engagement Rate by Season and Sponsorship: Our violin plot in Fig. 3 presents the distribution of average ER for posts categorized by the season in which they were uploaded and whether they were sponsored or not. It may be limited in its ability to reveal significant differences between seasons due to the inherent imbalance in the dataset. Despite these limitations, the ER distribution for all seasons and sponsorship categories exhibits high variability, as indicated by the wide range of the violins. This suggests that while there may be central tendencies, individual post-performance varies significantly.

4) Analysis of Average Engagement Rate by Weekday and Sponsorship: Across weekdays, our line plot in Fig. 4 differentiates the average ER for posts between sponsored and non-sponsored. Both sponsored and non-sponsored posts fluctuate. However, sponsored and non-sponsored posts on Sunday stand out with the highest average ER, which suggests that weekends might be optimal days for uploading content to maximize engagement. On the other hand, while not as pronounced as the weekday peak, some variations are observed during weekdays. For instance, both categories are slightly lower on Friday,

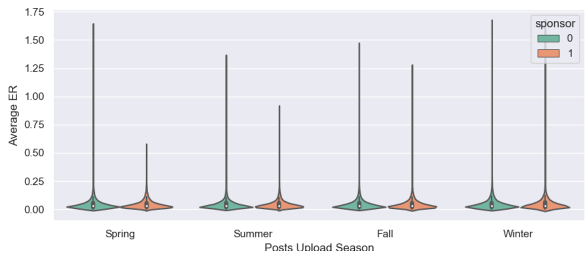

Fig. 3. Average ER by season (75% confidence interval)

nonetheless, Monday and Tuesday show relatively higher than Friday.

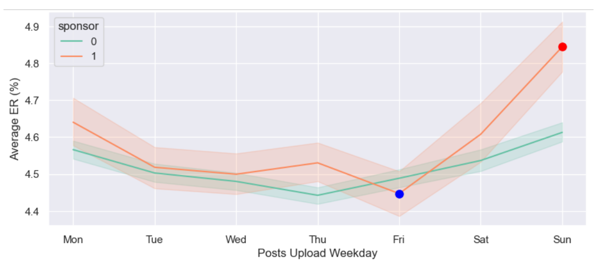

Fig. 4. Average ER by weekday (75% confidence interval)

5) Analysis of Average Engagement Rate by Hour and Sponsorship: Our line plot in Fig. 5 illustrates the average ER for posts across different upload hours, differentiating between sponsored and non-sponsored content. ER starts relatively low in the early morning hours, increases significantly during the late morning and early afternoon, and gradually declines throughout the evening and late night. While both categories follow a similar general trend, sponsored posts consistently maintain a higher average ER compared to non-sponsored throughout most of the day. However, the difference between the two categories appears to be most pronounced during the peak engagement hours. Observing trends in a day, both categories experience their lowest ER during these hours, suggesting that early morning[7] might not be the optimal time for posting content during early morning. In the late morning[8], this period witnesses the peak ER for sponsored and non-sponsored posts, indicating the most effective time frame for uploading content to maximize engagement.

The violin plots illustrate the distribution of ER for posts categorized by whether they were uploaded during working hours or non-working hours, further differentiated by sponsorship status. The width of each violin represents the density of data points at different ER values, providing insights into the spread and concentration of the data. The ER distribution for working and non-working hours exhibits a similar pattern, characterized by a high concentration of posts with low ER and a long tail extending towards higher ER values. This suggests

[7]Early Morning is indicating 12am.-6am.

[8]Late Morning is indicating 10am-3pm.

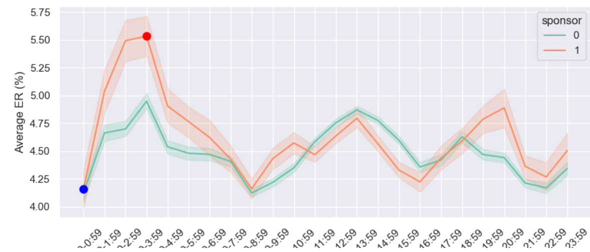

Fig. 5. Average ER by hour (75% confidence interval)

that while most posts achieve relatively low engagement, there are some outliers with exceptionally high ER.

The ER distribution for working and non-working hours exhibits a similar pattern, characterized by a high concentration of posts with low ER and a long tail extending towards higher ER values. This suggests that while most posts achieve relatively low engagement, there are outliers with exceptionally high ER. Interestingly, within the category of sponsored posts, there appears to be a subtle difference between working and non-working hours. The violin plot for sponsored posts during working hours has a slightly higher concentration of posts with lower ER compared to the violin plot for sponsored posts during non-working hours. This suggests a potential trend where sponsored posts uploaded during working hours might be associated with a lower likelihood of achieving high engagement.

C. Hashtag Analysis and Selection

We undertook a comprehensive analysis of hashtags to investigate thematic patterns within this content category. From Word Clouds visualization in 6, our findings indicated that the top 20 most commonly utilized hashtags were consistently aligned with themes pertinent to the fashion domain. Notably, the hashtag "#danielwellington" was included within this group. This brand, a watch company, has achieved significant growth by leveraging influencer marketing as its primary marketing strategy.

Fig. 6. Word Cloud of Hashtags (Sponsored Posts)

To explore the relationship between hashtag usage and engagement rate (ER), we built two regression models: one examining the effect of the total number of hashtags and another focusing on "important hashtags." From fig. 7, our analysis revealed a positive association between the total number of hashtags and ER, suggesting that including more hashtags might lead to a slight increase in engagement. However, it is crucial to consider the relevance of included hashtags. The model focusing on "important hashtags" can help identify which hashtags are most effective in driving engagement.

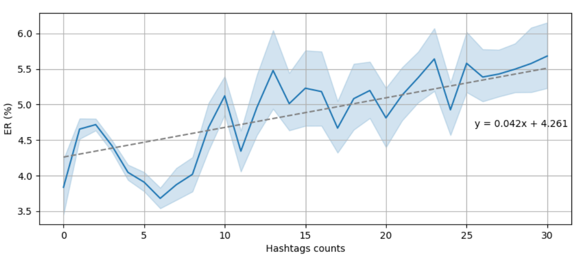

Fig. 7. Regression Analysis: Hashtag Count vs. Engagement Rate (Sponsored Posts)

Given the impracticality of including every hashtag due to its sheer volume, we focused our analysis on identifying the top 10[9] most commonly used hashtags across the dataset[10].

D. Supervised Learning

Our initial approach to predicting ER and identifying key features involved the utilization of supervised learning algorithms. Each entry within our dataset represents one post, from which we extract a variety of features sourced from both the post and influencer data. These features encompass essential metrics such as whether the post is sponsored, the influencer's post count, post time, the post's caption, as well as hashtags.

E. Feature Extraction and Encoding

Time series features that were previously identified as correlated with ER are derived from the timestamp. These features include year, month, day, hour, weekday, season, hour category (grouped into 3-hour periods), and working hour status. Among these, variables such as season, hour category, and working hour status are categorical. To ensure compatibility with a wider array of Machine Learning algorithms, such as XGBoost, these categorical variables are converted into numerical representations.

Given the impracticality of encoding the existence of each hashtag due to its large volume, we opt to identify the top 10[11] most popular hashtags across the dataset. Hashtags such as 'ad' and 'sponsor', which are already represented by the 'sponsor' variable, are excluded from this analysis. The features[12] are utilized in our Machine Learning models.

F. Model Selection and Optimization for Engagement Rate Prediction

To select an optimal model for predicting ER, we evaluated a diverse range of supervised learning algorithms. Our exploration contained several commonly employed models[13]. To objectively assess and compare the performance of these models, we employed mean squared error (MSE) as our primary evaluation metric. To ensure every feature on the same scale, we standardized the training and testing datasets. By fitting each model on the training data and calculating the MSE on the test data, we concluded that Random Forest Regressor emerged as the clear leader, achieving the lowest MSE of 0.026, significantly outperforming the linear regression benchmark of 0.046. This indicated the superior ability of Random Forest to capture the complex relationships between various features and ER. The MSEs of all the employed models can be found in Appendix C .

To further enhance the performance of our chosen model, we employed GridSearchCV with 5-fold cross-validation to fine-tune its hyperparameters. With these optimized hyperparameters, our final Random Forest Regressor model achieved an impressive MSE of 0.0012, demonstrating a substantial improvement over the initial results and solidifying its position as the most effective model for ER prediction.

G. Feature Importance

To find out the critical features in determining ER in supervised learning, we calculated the importance of each feature, as shown in fig. 8.

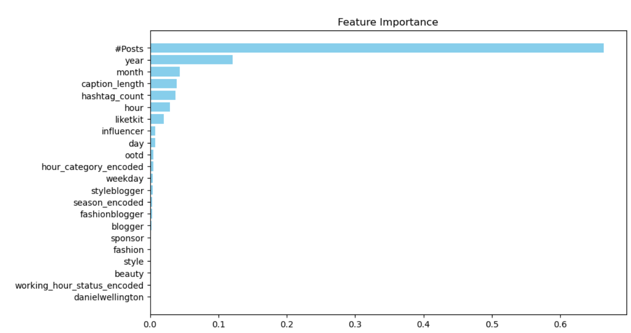

Fig. 8. Feature Importance of the Final Supervised Learning Model)

[9] 'liketkit', 'ootd', 'fashion', 'fashionblogger', 'danielwellington', 'blogger', 'style', 'styleblogger', 'beauty', and 'influencer'

[10] For those hashtags already represented by the 'sponsor' variable ('ad' and 'sponsor') were excluded from consideration.

[11] 'liketkit', 'ootd', 'fashion', 'fashionblogger', 'danielwellington', 'blogger', 'style', 'styleblogger', 'beauty', and 'influencer'

[12] '#Followers', '#Posts', 'sponsor', 'year', 'month', 'day', 'hour', 'weekday', 'season_encoded', 'hour_category_encoded', 'working_hour_status_encoded', 'hashtag_count', 'caption_length', 'liketkit', 'ootd', 'fashion', 'fashionblogger', 'danielwellington', 'blogger', 'style', 'styleblogger', 'beauty', and 'influencer'

[13] Linear Regression: Serving as a benchmark for comparison. Regularized Linear Models: Lasso and Ridge Regression to address potential overfitting. K-Nearest Neighbors Regressor: Leveraging the similarity between data points. Ensemble Methods: Random Forest, Gradient Boosting AdaBoost and XGBoost to utilize the power of multiple models. Multi-layer Perceptron: Exploring the capabilities of a neural network architecture.

The post count of an influencer emerges as the most important feature in ER prediction. Year and month of the post time, along with the length of captions, also exhibit a noteworthy degree of importance within our model. Conversely, the role of hashtags, while present, was found to be comparatively less substantial than initially hypothesized, exerting only marginal effects on the final model outcome. Also, sponsorship has almost 0 feature importance, showing that sponsorship is not a determinant of ER's value. This outcome shows that among features that are not image or text vectors, the intrinsic attributes characterizing an influencer play a pivotal role in determining the ER of a post.

H. Unsupervised Learning

The neural network model structure, depicted in Fig. 9, is designed to predict the engagement rate without considering the post's release time. Comprising four primary components, the model integrates visual and textual information alongside temporal features to form a comprehensive representation.

1) Visual Encoder: Responsible for feature extraction from images, the visual encoder leverages a pre-trained Vision Transformer (ViT) model to convert images into vector representations.

2) Text Encoder: Dedicated to extracting features from text, the text encoder utilizes pre-trained Bidirectional Encoder Representations from the Transformers (BERT) model to transform textual content into vector representations.

3) Time Features Input: Derived from each post's release timestamp, seven distinct features are extracted, including month, day, hour, weekday, season, hour category, and working hour. These temporal features provide additional context for the prediction task.

4) Fusion Layer: At the model's core, the fusion layer merges the features extracted by the visual and text encoders. Implemented as a fully connected neural network (FCNN) [22]with one hidden layer, the fusion layer synthesizes the information from both modalities into a single vector representation. This vector serves as the comprehensive feature representation of the combined image and text inputs.

This structured approach enables the model to effectively integrate visual, textual, and temporal information, facilitating accurate predictions of engagement rate in influencer marketing scenarios.

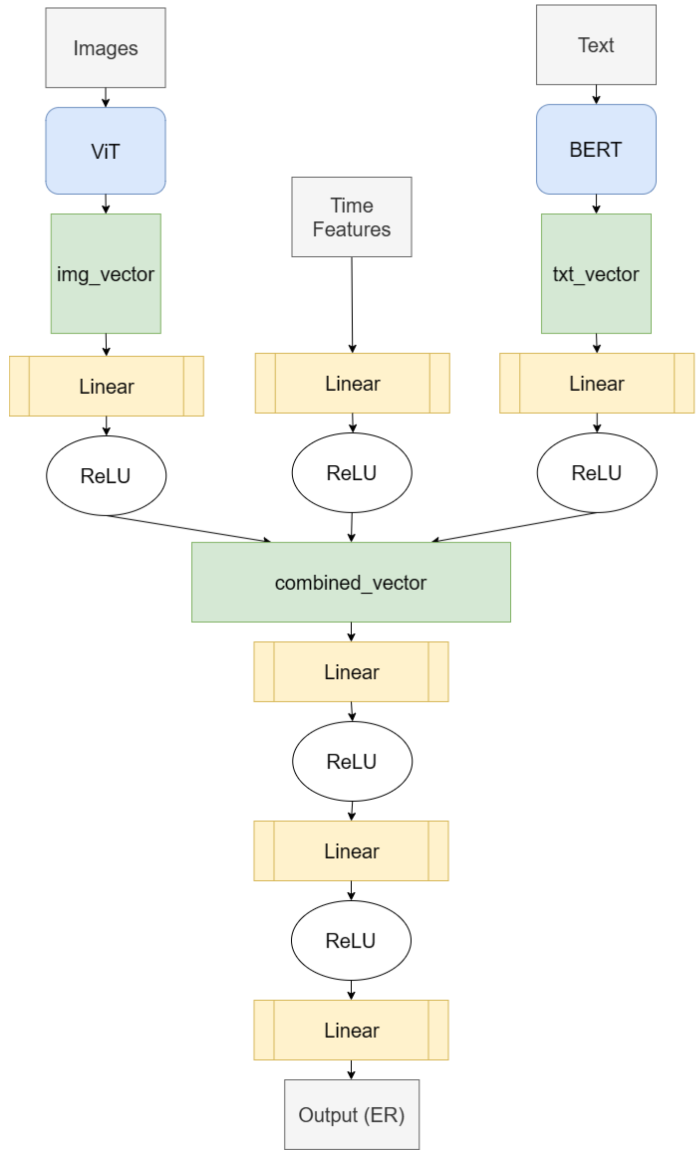

Fig. 9. The unsupervised learning model structure

V. Discussion

A. Impact of Sponsorship on Engagement Rate

While the overall analysis revealed a statistically significant difference in engagement rate (ER) between sponsored and non-sponsored posts, the effect size was marginal. This suggests that sponsorship alone may not be a guaranteed driver of increased engagement. The finding that sponsored posts exhibit a statistically significant, yet marginal, difference in ER compared to non-sponsored posts presents several potential explanations. The effectiveness of sponsored content depends on the alignment between the influencer and the target audience. An influencer whose values, aesthetics, and niche expertise resonate with the audience is more likely to generate genuine engagement, regardless of sponsorship. Audiences are increasingly discerning and value authenticity. If a sponsored post feels forced or lacks transparency, it can negatively impact engagement. Influencers who integrate sponsored content organically into their authentic voice and maintain transparency about their partnerships can achieve better results.

The potential benefits from the marginal effect are further compounded when considering the long-term impact on brand loyalty and customer retention. Sponsorship can provide access to a broader audience through the influencer's existing follower base and potential promotional strategies employed by the brand. This increased reach can contribute to overall brand awareness and visibility, even if the direct engagement on the sponsored post is not significantly higher. On the other hand, partnering with relevant and respected influencers can enhance brand association and credibility by leveraging the influencer's positive reputation and influence within their niche. This can indirectly contribute to increased engagement with the brand's content and marketing efforts.

From the perspective of future marketing strategy, instead of solely relying on sponsorship as a driver of engagement, brands should develop a holistic influencer marketing strategy that considers content quality, influencer selection, audience alignment, and platform-specific dynamics. To enhance the eWOM conversation between influencer and audience, merchandise cultivating long-term partnerships with influencers can foster authenticity and trust with the audience, leading to more organic and sustainable engagement over time.

B. Time-series Trends and Engagement Rate Pattern

Observed temporal trends in ER offer valuable insights into audience behavior and guide for optimizing content scheduling and influencer marketing strategies. The steady increase in ER over the years aligns with the growth trajectory of Instagram as a platform. As the user base expands and engagement becomes more ingrained in user behavior, the potential for higher ER naturally increases. Scoping into seasonal posts analysis, the peaks in ER during specific months, particularly around January and the holiday season, suggest that users are more active and engaged during these periods. Brands can leverage this trend by tailoring content to holiday themes, promotions, and festivities. The dip in ER during summer months could be attributed to vacation periods and increased outdoor activities, leading to decreased social media usage. However, this also presents an opportunity to create content that aligns with summer themes and travel experiences. More specifically, the fluctuations in ER throughout the day reflect the typical daily routines and social media usage patterns of the audience. Mornings and evenings, when users are commuting or winding down, often see higher engagement. Understanding when the target audience is most active and receptive to content is crucial for optimizing posting schedules. Experimenting with different posting times and analyzing engagement data can help identify the optimal windows for maximum reach and impact.

Aligning content themes and campaigns with seasonal trends and holidays can capitalize on periods of higher user engagement and interest and adapting posting schedules to account for daily fluctuations in user activity ensures that content reaches the target audience during their peak engagement times are essential points to arrange a new post. After posting, continuously monitoring engagement data and making real-time adjustments to posting times and content strategies allows for optimization based on actual audience behavior.

In future research, the manual assessment values in this study can be changed to similar automated values in order to reduce subjectivity. Other features, such as user history, can still be tuned to get better results. Text analysis features, such as concept semantic similarity [26], can also be added to distinguish between popular or less popular posts.

VI. LIMITATION

The large size and complexity of the dataset presented challenges in terms of computational resources required for certain analyses, particularly those involving deep learning models for image and text processing. While we employed efficient techniques and algorithms, exploring more computationally intensive methods could potentially yield further insights and refine the predictive models. Therefore, due to constraints in the dataset and computational resources, our image analysis focused solely on the first image of each post. However, many Instagram posts include multiple images, which could collectively contribute to the overall message and engagement potential. Analyzing the full set of images within each post might reveal additional insights into visual content strategies and their impact on engagement. In addition, our analysis relied primarily on data extracted directly from Instagram posts and user profiles. This limited our ability to explore deeper insight into the influencers' backgrounds, such as their demographics, professional experiences, or other social media activities. Such information could provide further context and potentially reveal additional factors influencing their effectiveness and engagement rates. Moreover classification tasks on influencers can be undertaken to unveil additional patterns regarding the effectiveness of sponsorship, both across and within categories, which would provide insights to inform brand sponsorship decisions.

REFERENCES

[1] B. Feehan, "2019 social media industry benchmark report," Nov 2021. [Online]. Available: https://www.rivaliq.com/blog/2019-social-media-benchmark-report/

[2] "Consumer trust in online, social and mobile advertising grows," Apr 2012. [Online]. Available: https://www.nielsen.com/insights/2012/consumer-trust-in-online-social-and-mobile-advertising-grows/

[3] G. Aerts, T. Smits, and P. Verlegh, "How online consumer reviews are influenced by the language and valence of prior reviews: A construal level perspective," *Computers in Human Behavior*, vol. 75, pp. 855–864, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0747563217303941

[4] P. Torres, M. Augusto, and M. Matos, "Antecedents and outcomes of digital influencer endorsement: An exploratory study," *Psychology & Marketing*, vol. 36, no. 12, pp. 1267–1276, 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/mar.21274

[5] W.-H. S. Tsai and L. R. Men, "Social ceos: The effects of ceos' communication styles and parasocial interaction on social networking sites," *New Media & Society*, vol. 19, no. 11, pp. 1848–1867, 2017. [Online]. Available: https://doi.org/10.1177/1461444816643922

[6] S. Verma and N. Yadav, "Past, present, and future of electronic word of mouth (ewom)," *Journal of Interactive Marketing*, vol. 53, no. 1, pp. 111–128, 2021. [Online]. Available: https://doi.org/10.1016/j.intmar.2020.07.001

[7] J. L. Do-Hyung Park and I. Han, "The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement," *International Journal of Electronic Commerce*, vol. 11, no. 4, pp. 125–148, 2007. [Online]. Available: https://doi.org/10.2753/JEC1086-4415110405

[8] A. M. Barreto, "The word-of-mouth phenomenon in the social media era," *International Journal of Market Research*, vol. 56, no. 5, pp. 631–654, 2014. [Online]. Available: https://doi.org/10.2501/IJMR-2014-043

[9] S. Verma and N. Yadav, "Past, Present, and Future of Electronic Word of Mouth (EWOM)," *Journal of Interactive Marketing*, vol. 53, no. C, pp. 111–128, 2021. [Online]. Available: https://ideas.repec.org/a/eee/joinma/v53y2021icp111-128.html

[10] M. Kim and J. Kim, “How does a celebrity make fans happy? interaction between celebrities and fans in the social media context,” *Computers in Human Behavior*, vol. 111, p. 106419, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0747563220301722

[11] B. C. Maecker, O. and J. Becker, “The effect of social media interactions on customer relationship management.” *BuR - Business Research*, vol. 9, pp. 133–155, 2016.

[12] J. Kim and M. Kim, “Rise of social media influencers as a new marketing channel: Focusing on the roles of psychological well-being and perceived social responsibility among consumers,” *International Journal of Environmental Research and Public Health*, vol. 19, no. 4, 2022. [Online]. Available: https://www.mdpi.com/1660-4601/19/4/2362

[13] M. Delbaere, B. Michael, and B. J. Phillips, “Social media influencers: A route to brand engagement for their followers,” *Psychology & Marketing*, vol. 38, no. 1, pp. 101–112, 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/mar.21419

[14] V. L. Thomas, M. Chavez, E. N. Browne, and A. M. Minnis, “Instagram as a tool for study engagement and community building among adolescents: A social media pilot study,” *DIGITAL HEALTH*, vol. 6, p. 2055207620904548, 2020, pMID: 32215216. [Online]. Available: https://doi.org/10.1177/2055207620904548

[15] R. Rogers, “Visual media analysis for instagram and other online platforms,” *Big Data & Society*, vol. 8, no. 1, p. 20539517211022370, 2021. [Online]. Available: https://doi.org/10.1177/20539517211022370

[16] M. E. B. S. e. a. Chakrabarti, P., “Hashtag recommendation for enhancing the popularity of social media posts,” *Soc. Netw. Anal. Min.*, vol. 13, no. 21, 2023.

[17] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” 2023.

[18] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” 2019.

[19] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, “An image is worth 16x16 words: Transformers for image recognition at scale,” 2021.

[20] S. Kim, J.-Y. Jiang, M. Nakada, J. Han, and W. Wang, “Multimodal post attentive profiling for influencer marketing,” in *Proceedings of The Web Conference 2020*, ser. WWW ’20. New York, NY, USA: Association for Computing Machinery, 2020, p. 2878–2884. [Online]. Available: https://doi.org/10.1145/3366423.3380052

[21] S. Kim, J.-Y. Jiang, and W. Wang, “Discovering undisclosed paid partnership on social media via aspect-attentive sponsored post learning,” in *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, ser. WSDM ’21. New York, NY, USA: Association for Computing Machinery, 2021, p. 319–327. [Online]. Available: https://doi.org/10.1145/3437963.3441803

[22] L. F. S. Scabini and O. M. Bruno, “Structure and performance of fully connected neural networks: Emerging complex network properties,” 2021.

APPENDIX A
CODE

The code we implemented in this research is available in GitHub repository

APPENDIX B
SUMMARY OF DATA STRUCTURE

User-based features:

- JSON file: A unique identifier derived from the original JSON file name of each post.
- Username: The name of the user who created the post.
- Sponsor: An indicator denoting whether the post was sponsored by a brand or not.
- Image files: A reference to the image file name associated with the post.
- Number of followers: The count of users following the influencer at the time of the post.
- Number of followees: The count of users the influencer follows at the time of the post.
- Number of posts: The total number of posts created by the influencer.

Post-based features:

- JSON: A unique identifier derived from the original JSON file name of each post.
- Post Timestamp: The date and time when the post was created.
- Number of Likes: The total number of likes received by the post.
- Number of Comments: The total number of comments associated with the post.
- Edge Media to Caption: Information about the caption of the post, including content or hashtags.
- Edge Media to Comment: Information about the comments associated with the post, including user IDs or comment content.
- Edge Media to Sponsor User: An indicator denoting is sponsored and may include details about the sponsoring user.
- Edge Media to Tagged User: Information about users tagged within the post, including user IDs and usernames.

APPENDIX C
MEAN SQUARED ERROR OF SUPERVISED LEARNING MODELS

Model	Mean Squared Error
Linear Regression (benchmark)	0.046
Linear Regression with Ridge Regularization	0.046
Linear Regression with Lasso Regularization	0.048
Gradient Boosting Regressors	0.039
AdaBoost Regressor	0.045
XGBoost Regressor	0.034
Multi-Layer Perceptron	0.044
Random Forest Regressor (not tuned)	0.026
Random Forest Regressor (tuned)	0.0012