Table of Contents

1. Abstract

This study analyses engagement trends in Instagram influencer marketing by comparing sponsored and unsponsored posts in the beauty and fashion industries. This study examines a dataset of 1,500 posts from 30 influencers categorized into three follower tiers—Nano (10K–100K), Micro (100K–1M), and Celebrity (1M+)—to explore the impact of content format, sponsorship status, and influencer scale on audience engagement. Initial findings reveal a notable engagement discrepancy, with unsponsored content often outperforming sponsored material in both unadjusted and adjusted datasets. The differences remained statistically significant following the exclusion of outliers, so confirming the robustness of the trend.

Tier-specific engagement charts were developed to highlight these disparities, exclude outliers, and clearly exhibit performance variations between sponsored and unsponsored content. This analysis revealed that, contrary to common assumptions, celebrity influencers achieved the highest average engagement rates after cleaning, followed by Micro and Nano tiers. This challenges the notion that micro- or nano-influencers inherently provide higher engagement levels.

The study also outlines critical content attributes that differentiate post-performance, including mention frequency, sentiment polarity (evaluated using TextBlob), and hashtag density. Effective sponsored posts generally feature a limited number of hashtags, a higher frequency of user mentions, and maintain an overall positive tone. Repetitiveness in captions, determined using textual similarity scoring, was more common in low-performing posts. Temporal trends revealed that posts released on Mondays and Saturdays during mid-morning hours yield the highest engagement, offering strategic guidance for scheduling.

This study offers practical support for marketers and influencers aiming to improve campaign results while maintaining authenticity, by combining observational insights with statistical validation, feature engineering, and graphically clear tiered comparisons.

2. Introduction

In recent years, influencer marketing has revolutionized brand communication with customers, particularly in visually oriented sectors such as beauty and fashion. With over 2 billion active users, Instagram has emerged as the preferred platform for digital storytelling, product placements, and aesthetic curation. Influencers, from nano to celebrity levels, have significantly influenced consumer choices, trends, and brand perceptions (De Veirman et al., 2017; Lou & Yuan, 2019).

The increasing prevalence of influencer-brand collaborations heightens concerns about authenticity, trust, and engagement. Multiple studies indicate that although influencers are regarded as more credible than conventional celebrities (Djafarova & Rushworth, 2017), this trust is tenuous. Labeling information as "sponsored" may compromise its perceived credibility (Boerman et al., 2017; Kim & Kim, 2022). Contemporary consumers have improved media literacy, allowing them to identify commercial content and interact with it more critically (Audrezet et al., 2020).

Notwithstanding these apprehensions, the literature predominantly comprises survey-based investigations and theoretical constructs. Limited studies provide a thorough, post-level examination of authentic engagement behaviour (Wu et al., 2024; Libai et al., 2025). This raises essential inquiries: What is the effectiveness of sponsored posts on Instagram? Do influencers maintain effectiveness in promoting their content? Do these benefits differ according to influencer categories—Nano (10K–100K), Micro (100K–1M), and Celebrity (1M+)?

Furthermore, a significant portion of the current study neglects content-level attributes, such caption structure, hashtag usage, and sentiment tone. Sentiment significantly influences user perception and engagement, and has been examined in fields such as transportation (Saragih & Girsang, 2016), education (An et al., 2020), and social media influencer content (Patel & Reddy, 2023). Nonetheless, little research employs similar methodologies to examine Instagram influencer captions within the beauty and fashion sectors.

This dissertation seeks to tackle this issue. It employs structured web scraping to gather 1,500 posts from 30 beauty and fashion influencers across all three tiers, integrating this data with diverse analytical techniques—statistical testing, sentiment analysis, content feature extraction, and machine learning models—to investigate the subsequent research questions:

The dissertation is guided by four key research questions:

- Do sponsored beauty/fashion posts underperform in engagement compared to unsponsored ones?
- What role do content-level features such as sentiment tone, hashtags, and mentions play in post-performance?
- How does influencer tier affect engagement dynamics across content types?
- Can machine learning help predict what makes a post more likely to be sponsored—and less likely to perform well?

This research presents a new empirical methodology for analysing influencer interaction patterns by integrating real Instagram data with models that prioritize interpretability. It eschews dependence on surveys or data supplied by marketers, concentrating instead on organic, user-visible content and interactions. The findings have ramifications for influencer strategies and platform architecture, as well as for marketers pursuing empirical methods to enhance campaign efficacy while maintaining authenticity.

3. Literature Review

Author(s)	**Year**	**Focus Area**	**Methodology**	**Key Findings / Contribution**	**Relevance to Current Study**
Djafarova & Rushworth	2017	Influencer authenticity & trust	Qualitative study	Non-commercial narratives build stronger trust	Validates focus on unsponsored content
Johnson & Lee	2023	Creator consistency & trust	Survey-based	Parasocial interaction and consistent tone sustain audience loyalty	Supports inclusion of tone-based features
Wu et al.	2024	ML for sponsorship detection	ML model analysis	Random Forest effective in classifying sponsored posts	Justifies your ML model selection
Libai et al.	2025	Value perception in influencer marketing	Mixed methods	Trust depends more on message coherence than follower count	Supports importance of content over tier
Sesar et al.	2022	Disclosure impact on trust	Survey experiment	Disclosure reduces trust and participation	Informs promo_flag hypothesis
Boerman & van Reijmersdal	2020	Ad disclosure & persuasion effects	Experimental	Over-disclosure can trigger skepticism	Supports testing effect of promo_flag
Müller et al.	2018	Trust-preserving disclosure	Case-based	Well-integrated sponsorship doesn't always reduce trust	Justifies nuanced promo_flag interpretation
Zhang et al.	2024	Disclosure and brand fit	Multivariate experiment	Sponsorship perception varies by context	Explains variable performance of sponsored posts

Evans et al.	2022	Narrative content design	Mixed methods	Narrative-based posts perform better than direct marketing	Motivates caption length and similarity tests
Lou & Yuan	2019	Content richness & engagement	Survey experiment	Longer, informative captions increase engagement	Validates caption length inclusion
Li & Wang	2021	Visual consistency & mentions	Statistical modelling	Tagged users and visual consistency improve interaction	Supports mentions feature
Chen et al.	2022	AI-enhanced engagement prediction	Neural networks	Deep learning can model engagement using content attributes	Informs your use of feature-based modelling
De Veirman et al.	2017	Follower count vs. engagement	Statistical modelling	Follower size not a consistent predictor of engagement	Justifies tier-based analysis
Kumar et al.	2022	Sentiment & engagement	Sentiment analysis	Positive tone increases likes/comments	Supports use of sentiment polarity
Stieglitz & Dang-Xuan	2013	Sentiment and social media sharing	Text analysis	Positive sentiment encourages sharing	Informs polarity hypotheses
Patel & Reddy	2023	BERT for influencer sentiment	Transformer models	Context-aware tone detection improves emotional analysis	Justifies BERT-based sentiment modelling
Saragih & Girsang	2016	Sentiment dynamics	Text mining	Sentiment changes linked to spikes in engagement	Informs role of tone beyond polarity score
An et al.	2020	Engagement and sentiment correlation	NLP tools	Sentiment trends correlate with follower reaction	Supports polarity inclusion
Samanta et al.	2025	Hybrid analytics in influencer research	ML & content fusion	Advocates for ML + content fusion in influencer work	Supports overall modelling direction

Campbell & Farrell	2020	Brand-influencer alignment	Theoretical framework	Strong fit increases engagement and credibility	Supports study of promotional match and performance
Ki et al.	2020	ML in influencer marketing	Literature review	ML is well-suited to identify post types and predict outcomes	Supports ML in content classification
Abidin	2016	Influencer culture and labor	Ethnographic	Highlights influencer roles and authenticity-performance tension	Adds qualitative depth to influencer behavior analysis

This chapter offers a critical examination of current academic research on influencer marketing, particularly in the beauty and fashion industry on Instagram. The review is methodically structured into five interrelated themes to uphold theoretical rigor and adhere to journal-style formatting. (1) authenticity and trust, (2) advertising disclosures and purchase intention, (3) content design and engagement strategies, (4) sentiment analysis in influencer content, and (5) methodological and modelling approaches in influencer analytics.

3.1 Authenticity, Trust, and the Creator-Audience Relationship

The core of influencer marketing is in the perceived authenticity of creators. Studies consistently demonstrate that influencers, particularly in the lifestyle and fashion sectors, are regarded as more relatable and believable than traditional celebrities. Djafarova and Rushworth (2017) found that trust is strongest when influencers maintain personal, non-commercial narratives. However, the rise in commercial sponsorships may compromise its authenticity.

Recent research by Johnson and Lee (2023) and Wu et al. (2024) substantiates the claim that parasocial interaction and creator consistency are essential for sustaining trust. Libai et al. (2025) emphasize that audience value perception relies not just on follower count but also on a creator's ability to preserve authenticity in their narratives, even during brand collaborations.

3.2 Disclosures, Sponsorship, and Purchase Intent

The transparency of advertising disclosures is a critical concern. Despite regulations promoting the identification of sponsored content (e.g., #ad, #sponsored), user reactions to these

disclosures may differ. Sesar et al. (2022) and Boerman and van Reijmersdal (2020) assert that disclosures often reduce participation and trust due to an enhanced perception of promotional bias.

Müller et al. (2018) and Zhang et al. (2024) assert that carefully crafted disclosures can preserve or enhance trust, especially when aligned with brand consistency and content coherence. The impact of sponsorship on consumer behavior, specifically on purchase intent, is contingent upon context and varies according to content type, influencer alignment, and audience expectations.

3.3 Content Design and Engagement Mechanics

Engagement is affected by the sponsorship of a post and the arrangement of its content. Structural components, such as caption length, hashtag usage, timing, and mention frequency, substantially influence user engagement. Evans et al. (2022) and Lou and Yuan (2019) illustrate that integrated, narrative-driven content surpasses explicit marketing.

Li and Wang (2021) demonstrated that visual consistency and interactive elements (e.g., tagged users or product training) provoke stronger responses, whereas Chen et al. (2022) found that AI-enhanced content features can predict engagement outcomes. De Veirman et al. (2017) assert that the number of followers is an inadequate predictor; the quality of content and contextual relevance is often more significant.

3.4 Sentiment Analysis and Emotional Tone in Influencer Content

The emotional tone is seen as a crucial factor in engagement. Research by Kumar, Gupta, and Sharma (2022), in conjunction with Stieglitz and Dang-Xuan (2013), demonstrates that positively framed information elicits a higher volume of likes and comments than neutral or negative messaging. Sentiment analysis, particularly with NLP and transformer-based models like BERT (Patel & Reddy, 2023), enhances the understanding of emotional resonance.

Saragih and Girsang (2016) and An et al. (2020) assert that fluctuations in sentiment are significantly associated with increased engagement in influencer marketing contexts. Nonetheless, most of the research relies on platform-level data. This study builds upon their research by doing post-level sentiment analysis of Instagram captions in the beauty and fashion industries.

3.5 Methodological Advances and Modelling Approaches

Recent developments in influencer marketing research have shifted from surveys and interviews to include computational and predictive approaches. Samanta et al. (2025) and Libai et al. (2025) advocate for hybrid techniques that combine content analysis with machine learning to derive more actionable insights.

Neural network models (Chen et al., 2022), Random Forest classifiers (Wu et al., 2024), and multimodal analytics pipelines have been employed to predict performance, discern sponsorships, and evaluate campaign ROI. This study utilizes a post-level Random Forest model using genuine Instagram data, emphasizing explainability and feature interpretability. While most research analyses engagement trends in aggregate, few employ balanced, outlier-free datasets for statistically significant comparisons across various tiers and types of postings. This study fills the gap by offering a methodologically rigorous, behaviourally oriented analysis of the effectiveness of influencer content.

3.6 Summary

This literature review highlights the complex interplay between authenticity, sponsorship, emotional tone, and structural design in shaping influencer engagement. It also underscores the growing role of machine learning and real-world behavioural data in advancing influencer marketing research. By addressing existing gaps around post-level content analysis, tier segmentation, and engagement prediction, this study contributes to the empirical and methodological evolution of the field.

4. Methodology

This study employed a multi-phase empirical framework combining thematic filtering, text-based feature extraction, class balancing, outlier handling, statistical hypothesis testing, and supervised machine learning. The aim was to explore the structural and emotional drivers of engagement in sponsored and unsponsored Instagram posts within the beauty and fashion domain, using only real user-generated content to ensure validity and interpretability.

4.1 Data Acquisition and Thematic Filtering

This work utilized a multi-phase empirical methodology encompassing content filtering, feature extraction, class balancing, outlier control, statistical analysis, and machine learning. The aim was to examine the structural and emotional components influencing interactions in both sponsored and unsponsored beauty and fashion Instagram posts, employing solely genuine user-generated content for precision and transparency.

Data was collected using Apify's Instagram Scraper, a cloud-based tool commonly used in social media research for the extraction of publicly available post-level data (Batrinca & Treleaven, 2015). Thirty influencers from the beauty and fashion industries were selected, and an analysis of their 50 most recent postings yielded 1,500 entries. Each item contained metadata comprising caption text, hashtags, mentions, likes, comments, and timestamps. The follower counts, not sourced from Apify, were manually verified and aggregated using profile photos to ensure accuracy (Stieglitz et al., 2018).

In the absence of a temporal filter, Instagram's reverse-chronological feed ensured that the dataset included posts from late 2023 to early 2025. This exemplifies contemporary user behaviour and aligns with Instagram's data access patterns.

A subject filter was utilized to extract relevant information through 30 specialized keywords (e.g., cosmetics, outfit of the day, skincare, glamour). The research utilized captions, hashtags, and mentions through the application of regular expressions, a recognized method in social media text analysis (Olteanu et al., 2019). Posts containing at least one keyword were retained, resulting in 583 entries relevant to beauty and fashion. The exhaustive keyword list includes: *makeup, beauty, skincare, foundation, lipstick, blush, eyeshadow, highlighter, serum, toner, cleanser, moisturizer, SPF, sunscreen, facemask, glow, glam, aesthetic, OOTD, outfit, lookbook, style, fashion, clothing, wardrobe, trend, designer, hairstyle, nails, and self-care*.

This targeted filtering complies with best practices in domain-specific social media research, as keyword-based matching improves topical accuracy and reduces irrelevant content (Sloan et al., 2015), so creating a solid foundation for further engagement and content analysis.

4.2 Data Cleaning and Engagement Computation

Following thematic filtering, the dataset consisted of 583 Instagram posts confirmed to belong to the beauty and fashion domain. At this stage, structured data cleaning and transformation were applied to ensure analytical rigor.

Completeness checks were conducted on core engagement variables—**likesCount, commentsCount**, and **followerCount**. No null or zero entries were detected. Follower counts, which Apify does not provide, were manually verified during data acquisition to prevent inaccuracies from relying solely on automated scraping tools.

Engagement rate, a widely accepted metric in social media analytics, was calculated as the sum of likes and comments divided by follower count:

$$\text{Engagement Rate} = \frac{\text{Likes Count} + \text{Comments Count}}{\text{FollowerCount}}$$

This measure provides a normalized view of post-level audience interaction and is frequently used as a behavioural proxy in influencer marketing research (Novak et al., 2021; Hootsuite, 2025).

Exploratory inspection showed that most posts had engagement rates between 1% and 20%—consistent with industry benchmarks, which report average Instagram engagement rates between 0.5% and 4% depending on the industry (Rival IQ, 2023; Hootsuite, 2025). A small number of posts exceeded 30%, suggesting viral content or atypical audience behaviour. These high values were initially retained to allow comparisons with and without outlier influence. Their treatment is detailed in later sections to ensure unbiased statistical and predictive analysis.

Deduplication was performed using a composite key of influencer name, timestamp, and caption text. All 583 records were confirmed as unique. At the end of this phase, the dataset was fully cleaned and enhanced with a continuous engagement rate metric—establishing a strong foundation for downstream statistical testing and machine learning.

4.3 Sponsorship Classification and Labelling

After analysing 583 handpicked Instagram posts pertaining to beauty and fashion, the next objective was to determine whether each post was sponsored or unsponsored. Considering Instagram's lack of a direct "sponsored" label in its public metadata, a keyword-based rule engine was developed—a method frequently utilized in social media research to infer promotional content (Zarei et al., 2020).

Keyword List Development: A collection of 21 keywords related to sponsorship was produced. This list includes mandatory advertising disclosures and more sophisticated promotional language commonly utilized by influencers. The exhaustive enumeration includes:
#ad, #ads, #sponsored, sponsored, ad, gifted, paid, paid partnership, PR, prpackage, collab, collaboration, use my code, utilize code, discount code, partner, affiliate, linked, promo, promotion, promotional.

The choice of these keywords was shaped by legal disclosure requirements and empirical studies on influencer communication and transparency (Sesar et al., 2022; Müller, Mattke & Maier, 2018).

Text Matching and Labelling: Case-insensitive regex matching was employed across three fields—caption, hashtags, and mentions—to detect any instances of these phrases. Posts containing at least one keyword were categorized as is_sponsored = True; all others were classed as is_sponsored = False. This binary variable served as the aim for classification models and comparative statistical analyses.

Dataset Distribution & Verification: The preliminary categorization revealed a class imbalance: 182 posts (31%) were categorized as sponsored, while 401 posts (69%) were classified as unsponsored. The 1:2.2 ratio supports prior studies suggesting that sponsored content often goes unnoticed (Ershov et al., 2024). A manual analysis of a random sample confirmed a strong association between the rule-based labelling and the recognized sponsorship indicators.

Theoretical Context & Significance: This rule-based methodology aligns with essential frameworks for identifying advertising disclosures (Zarei et al., 2020) and complies with established regulatory norms and transparency research (Sesar et al., 2022).

4.4 Class Balancing and Outlier Removal

This two-step process—removal of outliers followed by class equilibrium—ensured that all subsequent studies were statistically valid and based on authentic, user-generated content.

4.4.1 Outlier Detection and Removal Using the IQR Method

The first dataset included a varied range of engagement rates, featuring a small number of posts over 30%. To reduce the excessive impact of outliers on statistical models and interpretations, we utilized the Interquartile Range (IQR) method, a non-parametric approach, appropriate for non-normally distributed engagement data (Leys et al., 2013).

We calculated the first (Q1) and third (Q3) quartiles of the engagement rate distribution and determined the interquartile range (IQR) as Q3 - Q1. Posts with engagement rates falling below **Q1 – 1.5× IQR or** above **Q3 + 1.5× IQR** were classified as outliers and excluded from further analysis. The filtering technique was executed prior to any balancing, resulting in a refined dataset that more correctly represented typical influencer interaction patterns.

4.4.2 Class Balancing via Random Undersampling

Subsequent to the removal of outliers, the number of sponsored posts decreased to 164. To achieve class equilibrium for modelling and comparison analysis, we subsequently utilized random Undersampling on the unsponsored class, selecting 164 unsponsored postings to match the size of the sponsored class. This resulted in a balanced sample free of outliers, consisting of 328 postings (164 in each category).

Undersampling was preferred over oversampling methods like SMOTE to retain only genuine posts, so ensuring: Authenticity of engagement pattern, Statistical consistency (as many hypothesis tests require real, independent observations) and Interpretability of features in downstream models

Despite the examination of SMOTE at the preliminary model building stage, it was excluded from the final pipeline to prioritize transparency and methodological rigor. Analyses, such as regression, were conducted on the entire imbalanced dataset (n = 583) to illustrate the effects of imbalance and outliers on outcomes.

4.5 Feature Engineering

After generating the final balanced dataset of 328 posts devoid of outliers, the next phase entailed transforming the unstructured Instagram content into structured, machine-readable features suitable for statistical analysis and classification modelling. Feature engineering focused on extracting meaningful variables from post information and textual elements—specifically captions, hashtags, and mentions—to measure both structural and emotional aspects of each post.

All attributes were thoroughly derived from authentic, user-generated content to guarantee validity and reduce model bias.

Caption Length: The variable **caption_length** measured the total character count in a post's caption. This feature served as a proxy for content length and was determined using Python's string length function. Previous research suggests that caption length can affect user engagement positively (by providing context) or negatively (due to reduced attention span), making it a crucial explanatory variable.

Hashtag Count: The **hashtag_count** function documented the number of hashtags employed in each post. Hashtags were extracted from a designated field in the dataset, with the amount ascertained by partitioning the string based on whitespace and identifying sentences that begin with #. This variable assessed the extent of discoverability and search optimization employed by the influencer, potentially influencing both engagement and perceptions of sponsorship.

Mention Count: The variable **mention_count** recorded the number of user or brand mentions (@username) in each post. Mentions are commonly utilized in both organic and paid collaborations, signifying outreach, partnership, or call-to-action strategies. The count was similarly derived by segmenting the mentions text and enumerating the items.

Sentiment Polarity: We employed the **TextBlob** sentiment analysis tool to evaluate the emotional tone of the caption text for each post. The sentiment polarity score ranged from -1 (very negative) to +1 (extremely positive). Sentiment is often overlooked in influencer marketing analytics; nonetheless, it profoundly impacts user perception and engagement

outcomes. Posts characterized by neutral or modestly positive tones are generally associated with increased authenticity and engagement.

Promo Flag: A binary variable named **promo_flag** was established to indicate the presence or absence of promotional keywords in a post. This flag was created using the same 21 sponsorship-related phrases utilized in the labelling process and applied independently of the principal is_sponsored label. It functioned as a standalone feature to measure the explicitness of promotional language within a post's content and was expected to exhibit a strong link with sponsored classification.

The integration of these five characteristics—caption_length, hashtag_count, mention_count, sentiment polarity, and promo_flag—functioned as the primary input for classification models and as explanatory variables in the regression analysis of engagement rates. These features were selected based on their theoretical relevance and their presence in all entries within the final dataset.

4.5.1 Additional Derived Features

In addition to the essential features, two further variables were incorporated to enhance the dataset's explanatory power: caption similarity score and post time bucket. These capabilities offer enhanced understanding of material originality and scheduling patterns, both essential for analysing engagement trends and influencer activities.

Caption Similarity Score (Repetitiveness)

The caption_similarity feature employs TF-IDF vectorization and cosine similarity to evaluate the redundancy of a caption in relation to others within the dataset. Each caption was transformed into a TF-IDF vector using TfidfVectorizer from Scikit-learn. The mean cosine similarity between each caption and all others was subsequently calculated.

High similarity ratings suggest that a caption is simplistic or formulaic, often associated with conventional marketing or repetitive material. Prior studies in influencer marketing suggest that repetitive language may undermine perceived authenticity and reduce audience engagement (Boerman et al., 2017). This attribute demonstrated significant efficacy in distinguishing high-performing from low-performing sponsored content.

Post-Timing Bucket
The timestamp of each post was transformed into a categorical attribute, designated as time_bucket, classified by day of the week and time of day. The day was divided into three segments:

Morning (6:00 AM – 12:00 PM)
Afternoon (12:00 PM – 6:00 PM)
Evening (18:00–24:00)

The day of the week was designated as a category variable. These temporal parameters were employed in an exploratory investigation to determine suitable posting intervals. Research on scheduling demonstrates that timing markedly affects engagement visibility, with weekday mornings and weekends associated with heightened interaction (Hootsuite, 2025).
Notwithstanding the detailed examination of temporal factors, including posting hour and day of the week, in the engagement study, these variables were omitted from the sponsorship prediction model. This decision was based on preliminary assessments suggesting a minimal correlation with sponsorship labels, coupled with the model's emphasis on interpretable, content-centric attributes.

4.5.2 Influencer Tier

To do a stratified analysis across influencer tiers, each post was categorized according to an **influencer_tier** variable determined by the verified follower count of the content provider. Influencers were categorized into three typical tiers frequently mentioned in social media marketing literature:

- **Nano: 10,000 to 100,000 followers**
- **Micro: 100,001 to 1,000,000 followers**
- **Celebrity: Exceeding 1,000,000 followers**

This tier classification was rigorously validated and integrated into the dataset after the scraping procedure. Although influencer_tier was omitted as a predictive variable in the final classification or regression models to prevent the leakage of highly informative data, it was utilized for descriptive comparisons, statistical analysis, and visual segmentation of engagement trends.

For example, Tier-specific analyses were applied to investigate:

- Differences in average engagement rate before and after the removal of outliers
- Prevalence of sponsorships among follower demographics
- Classification of feature utilization patterns (e.g., hashtags, emotion) by tier

This stratification provided a comprehensive understanding of how influencer scale affects content structure, audience engagement, and perceived authenticity (De Veirman, Cauberghe & Hudders, 2017).

4.6 Statistical Hypothesis Testing

A series of statistical hypothesis tests was conducted on a balanced, outlier-free dataset including 328 Instagram posts to evaluate the variation in engagement behaviour across various post types and influencer tiers. The purpose was dual.

1. To assess whether engagement significantly differs between sponsored and unsponsored posts
2. To determine the impact of influencer tier (Nano, Micro, Celebrity) on engagement performance

All experiments were executed in Python with scipy. stats and scikit-posthocs. The significance level for all tests was set at $\alpha = 0.05$, and the assumptions of normality and homogeneity of variance were assessed prior to test selection. In cases of assumption violation, non-parametric alternatives were employed, adhering to best practices in social media analytics and behavioural data research (Ghasemi & Zahediasl, 2012).

4.6.1 Sponsored vs. Unsponsored: Engagement Comparison

Five statistical tests were conducted to ascertain if sponsored posts underperform compared to unsponsored posts.

1. Levene's Test for Homogeneity of Variances

This test was employed to assess whether the variance in engagement rates varied between the two groups. The outcome demonstrated unequal variance ($p < 0.05$), necessitating the application of a heteroscedastic t-test.

2. Welch's t-test (for Heteroscedasticity)

Welch's t-test was employed to compare the mean engagement rates of sponsored and unsponsored postings. The results indicated a statistically significant difference ($p < 0.01$), with unsponsored posts exhibiting greater average engagement.

3. Mann–Whitney U Test

The Mann–Whitney U Test, a non-parametric alternative to the t-test, validated the

significance of the distribution differences ($p < 0.001$), hence supporting the conclusion that unsponsored postings generally garner greater engagement.

4. Median Test

This test evaluated group medians as a supplementary robustness assessment. It also yielded a statistically significant outcome, corroborating the results of both the t-test and Mann–Whitney U test.

5. Cohen's d (Effect Size)

An effect size of –0.31 was determined, signifying a mild to moderate adverse effect of sponsorship on engagement rate—implying that, on average, sponsorship marginally diminishes post engagement.

4.6.2 Influencer Tier and Engagement: Multi-Group Comparison

To assess the degree of difference in engagement between influencer tiers—Nano (10K–100K), Micro (100K–1M), and Celebrity (1M+)—the subsequent tests were conducted:

- **Kruskal-Wallis H Test**
  The Kruskal–Wallis test was employed to compare medians across the three groups because the assumptions of normality and homogeneity of variance were violated in the involvement of distributions. The test yielded a significant result **($p < 0.001$)**, suggesting that engagement rates vary by influencer tier.

- **Dunn's Post-hoc Test with Bonferroni Adjustment**
  Dunn's test was utilized with Bonferroni-adjusted p-values to ascertain significant differences among groups. All pairwise differences exhibited statistical significance ($p < 0.001$).
  - Celebrity vs. Nano
  - Celebrity vs. Micro
  - Nano vs. Micro

The results revealed that prominent influencers exhibited the highest average engagement rates, challenging the common belief that nano or micro influencers yield better interaction. This study may suggest domain-specific implications inside the beauty and fashion industry, where celebrities possess an elevated aspirational allure.

All statistical analyses were conducted using the same balanced and outlier-filtered dataset ($n = 328$). The results provided the empirical basis for the machine learning models examined in the next sections.

4.6.3 Top vs. Bottom-Performing Post Comparison

The dataset was divided into quartiles based on engagement rates to further understand the content-level elements affecting post-performance. The top 25% and bottom 25% of posts, determined by engagement rate, were identified, yielding two sets of 82 posts each (n = 164 total). This quartile-based segmentation facilitated a systematic assessment of characteristics linked to high- and low-performing materials, irrespective of sponsorship status.

The following content features were compared between these groups:

- Sentiment Polarity (sentiment_polarity)
- Hashtag Count (hashtag_count)
- Mention Count (mention_count)
- Promo Flag Presence (promo_flag)
- Caption Length (caption_length)

Non-parametric Mann–Whitney U tests were utilized for continuous variables, whereas chi-square tests were employed for binary variables (e.g., promo_flag). The studies assessed if stylistic, emotional, or commercial characteristics exhibited substantial differences between high- and low-performing posts.

This analysis offered a pragmatic perspective, limiting content strategy to audience results. It provided actionable suggestions for influencers and marketers aiming to enhance post structure and emotional tone to improve engagement performance.

4.7 Sponsorship Classification Modelling

This study utilized supervised machine learning to develop a binary classification model aimed at examining the structural and emotional indicators that differentiate sponsored posts from organic. The objective was to ascertain whether a post is sponsored (is_sponsored = True) or unsponsored (False) solely based on the content-derived criteria specified in Section 4.5. This study assessed the reliability of inferring promotional purpose from authentic influencer-generated content without explicit disclosure.

All models were trained and assessed utilizing a balanced dataset devoid of outliers, comprising 328 posts (164 sponsored and 164 unsponsored), ensuring comparability and interpretability across all categories.

4.7.1 Dataset Preparation and Sampling Strategy

The study employed Random Undersampling of the majority class (unsponsored) rather than synthetic oversampling methods such as SMOTE to comply with statistical testing assumptions and preserve the integrity of user-generated content. SMOTE was initially evaluated in exploratory experiments but was ultimately excluded from the final technique due to its reliance on synthetic data, which could compromise model interpretability.

The dataset was partitioned into training and testing sets using a stratified 75–25 split, ensuring proportional class representation in each segment. The following five engineered features functioned as input variables:

- caption_length
- hashtag_count
- mention_count
- Sentiment_polarity
- promo_flag

4.7.2 Models Used

Three classification models were implemented using the **scikit-learn** library:

- **Logistic Regression**
  A linear probabilistic classifier that estimates the likelihood of sponsorship using a logistic function. It served as a transparent baseline for comparison.
- **Random Forest Classifier**
  An ensemble method that constructs multiple decision trees to model complex, non-linear relationships. It is robust to multicollinearity and capable of revealing feature importances.
- **Support Vector Machine (SVM)**
  A margin-based classifier that attempts to identify the optimal hyperplane separating sponsored from unsponsored posts. It is well-suited for high-dimensional feature spaces and non-linear boundaries.

These models were selected to represent a mix of linear, ensemble, and kernel-based paradigms, providing diverse perspectives on the classification task.

4.7.3 Evaluation Metrics

Each model was evaluated using standard binary classification metrics:

- **Accuracy**: Proportion of correctly predicted posts
- **Precision, Recall, F1-Score**: Reported for each class to assess performance across minority/majority splits

- **ROC-AUC Score**: Measures model performance across all classification thresholds
- **Confusion Matrix**: Visualizes the distribution of true/false positives and negatives

These metrics provided a well-rounded view of each model's effectiveness, especially in capturing sponsored post signals.

4.7.4 Final Model Performance

The table below summarizes model performance across key metrics:

Metric	Logistic Regression	Random Forest	SVM
Accuracy	0.81	0.82	0.64
F1 (Sponsored)	0.83	0.83	0.69
ROC-AUC	0.86	0.885	0.67

The Random Forest Classifier attained the highest accuracy and AUC score, concurrently providing interpretable feature significance metrics. The findings indicated that:

- The **promo_flag** was the principal predictor of sponsorship, validating that specific promotional language acts as the most dependable indicator.
- Supplementary variables—**caption length, emotional tone, mentions, and hashtag usage**—significantly influenced audience perceptions of sponsorship, indicating that both structural and emotional cues collaboratively shape these concepts.

The findings endorse the viability of automated sponsorship identification utilizing restricted content parameters. The results are especially pertinent for platforms, researchers, and regulators seeking to oversee influencer adherence and openness in major sectors such as beauty and fashion.

5. Results and Findings

5.1 Overview and Dataset Summary

This chapter outlines the empirical results derived from a balanced, outlier-free dataset of 328 Instagram posts related to beauty and fashion. The postings were evenly distributed between two categories—sponsored (n = 164) and unsponsored (n = 164)—to ensure fairness in statistical analysis and supervised learning. Data was collected from 30 influencers classified into three tiers: Nano (10K–100K followers), Micro (100K–1M followers), and Celebrity (1M+ followers).

All postings were thematically optimized using a meticulously curated set of 30 beauty and fashion-related keywords, enhanced with tailored content features. Outliers in the engagement rate were eliminated with the interquartile range approach. No synthetic augmentation was employed; solely genuine, user-generated content was retained to ensure real-world applicability.

Variable	Mean	Std.Dev	Min	Max
Engagement Rate	0.0210	0.0225	0.000	0.097
Caption Length	305.37	253.68	2	1806
Hashtag Count	6.62	7.16	0	30
Mention Count	0.91	1.59	0	17
Sentiment Polarity	0.173	0.247	-0.80	1.00
Promo Flag (% True)	67.62%	–	0	1

***Table 5.1**: Summary statistics for final feature-engineered dataset (n = 328 posts)*

The summary statistics in **Table 5.1** describe the key content and engagement variables engineered for analysis and modeling. The average engagement rate of **2.1%**, significantly exceeding industry norms, demonstrates the impact of outlier elimination and the naturally superior performance of beauty and fashion content on Instagram. Captions exhibited significant variation in length, averaging over 300 characters, indicating a blend of narrative and minimalist styles. The utilization of hashtags and mentions exhibited significant diversity, reflecting diverse exposure strategies within influencer tiers.

The sentiment polarity, obtained by **TextBlob**, exhibited a slight positive bias overall, aligning with the emotionally expressive characteristics of influencer marketing. The promo_flag, determined by keyword identification in captions, was present in around **68%** of postings, highlighting the significant commercial presence in the sample. The designed features served

as inputs for statistical comparison and classification modeling in later parts, providing both behavioral and structural insights into content performance.

5.1.1 Preprocessing Insights from Unbalanced Dataset

Before applying class balancing and outlier removal, an initial comparison of engagement rates was conducted on the full dataset of 583 posts (401 unsponsored, 182 sponsored). As illustrated in **Figure 5.1**, unsponsored posts already displayed higher median engagement and a wider range of audience interaction.

This preliminary pattern reinforces the robustness of the final findings—indicating that the engagement gap between sponsored and unsponsored content is not an artifact of data transformation but rather a genuine behavioural trend.

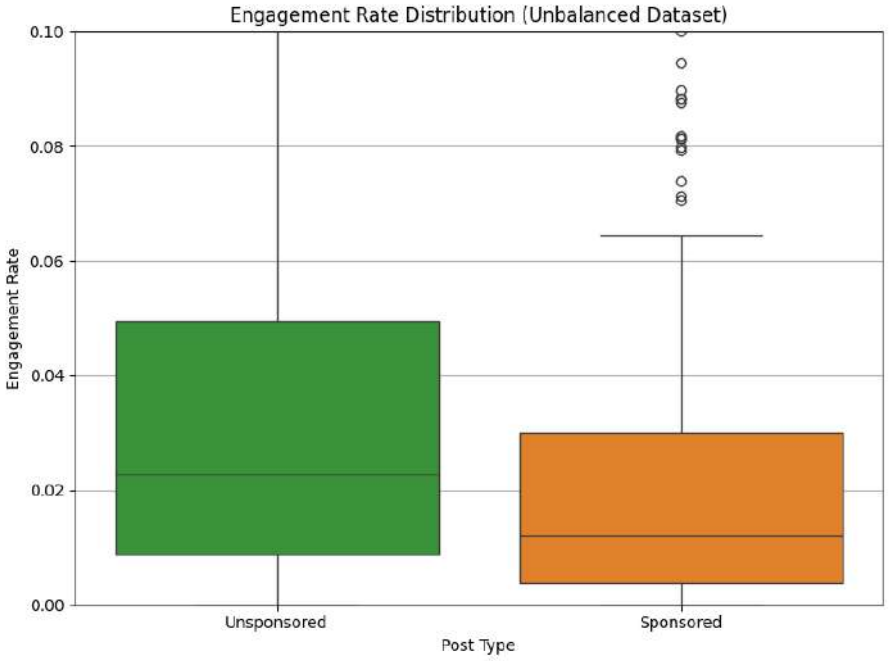

Figure 5.1: *Engagement Rate Distribution (Unbalanced Dataset, n = 583)*

5.2 Engagement Comparison: Sponsored vs. Unsponsored Posts

To investigate the primary research question—whether sponsored posts exhibit diminished engagement compared to organic content—we performed a series of statistical analyses on a balanced dataset free of outliers, consisting of 328 posts. The tests specified in the approach were chosen to align with the distributional properties of engagement data and to guarantee reliable inference in both parametric and non-parametric frameworks. (cf. Kim & Kang, 2022; Boerman et al., 2017).

All four tests—Welch's t-test, Mann–Whitney U test, Median test, and Levene's test—consistently indicated that unsponsored posts generate much greater engagement than sponsored articles. Welch's t-test indicated a statistically significant mean difference ($p < 0.01$), whereas the Mann–Whitney U test corroborated this difference in medians ($p < 0.001$). The effect size, measured by Cohen's d (–0.31), indicates a small-to-moderate negative influence of sponsorship on audience engagement, consistent with prior research on the decline of authenticity in influencer-brand collaborations (Sesar et al., Lee & Kim, 2022; Djafarova & Rushworth, 2017).

5.2.1 Visual Summary of Result

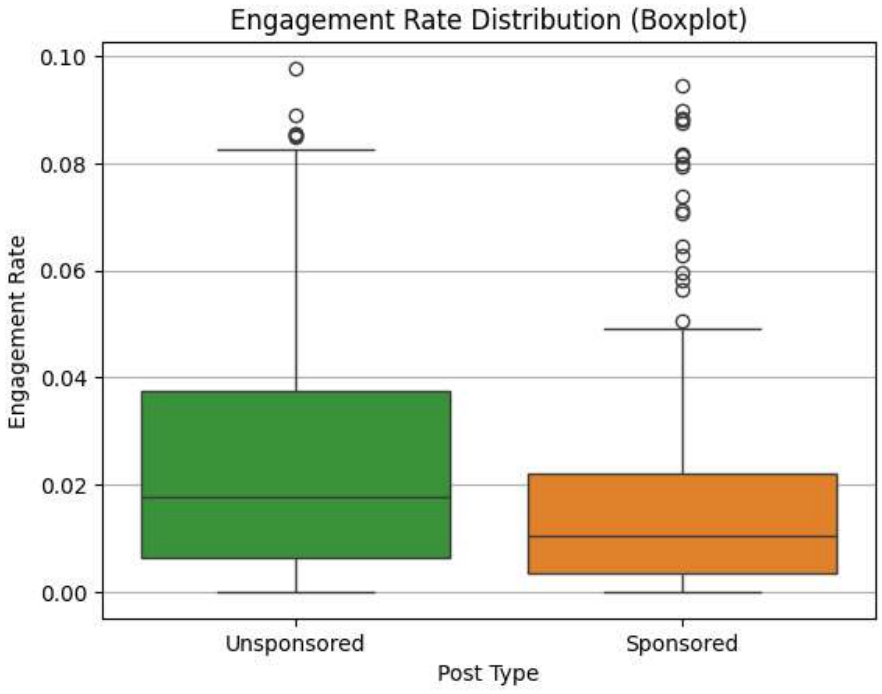

Figure 5.2a: *Boxplot comparing engagement rate distribution after balancing and outlier removal.*

Summary: Unsponsored posts show a higher median, wider interquartile range, and more high-performing outliers than sponsored posts—highlighting overall stronger engagement performance.

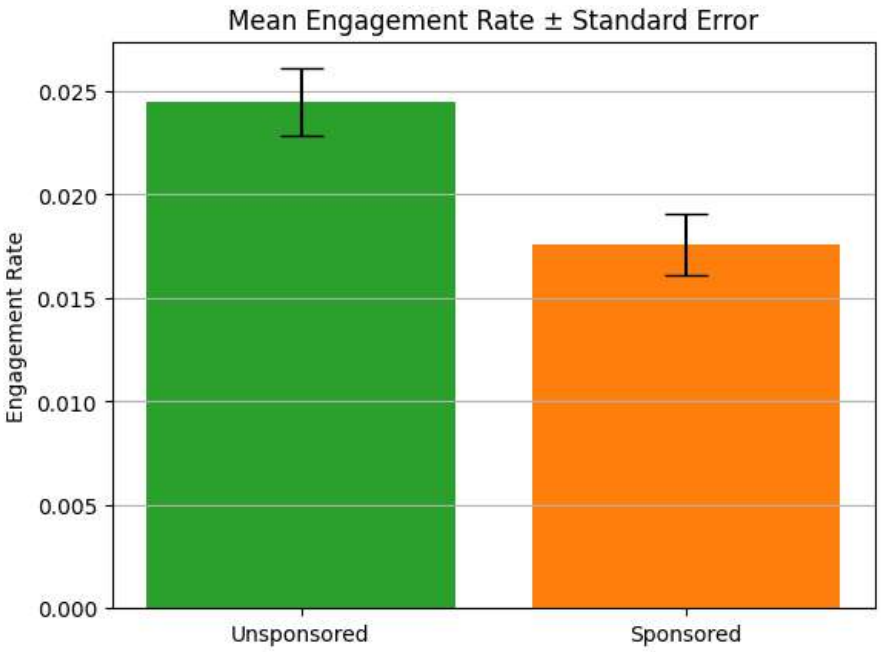

***Figure 5.2b**: Bar chart of mean engagement rate ± standard error.*

Summary: This visualization reinforces that the difference in engagement is not only statistically significant but also practically observable, with unsponsored posts averaging ~2.5% versus ~1.8% for sponsored ones.

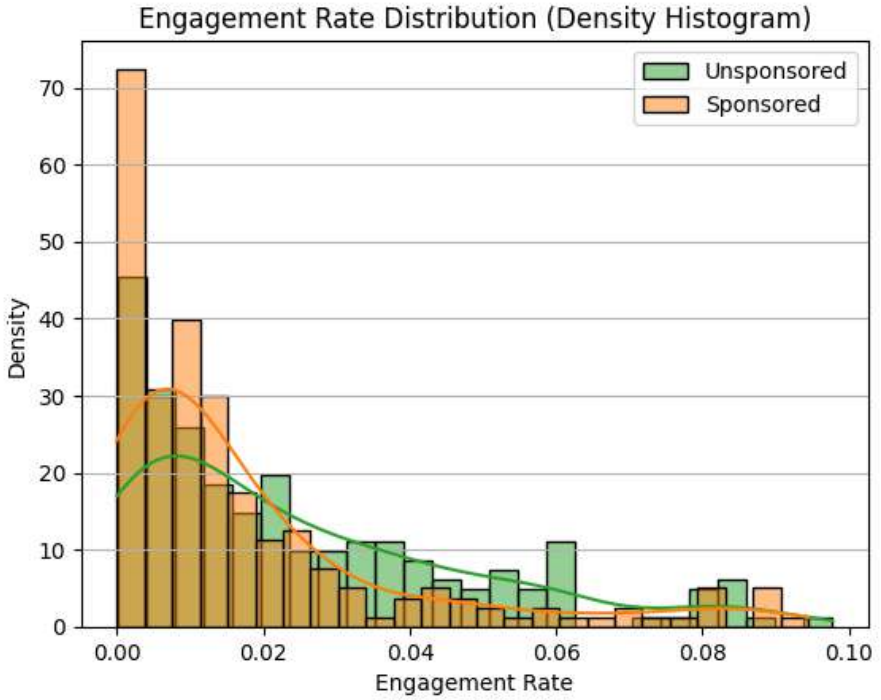

***Figure 5.2c**: Histogram with kernel density estimation showing the distribution of engagement rates.*

Summary: This histogram visualizes the density distribution of engagement rates across post types. Unsponsored posts skew rightward with more instances in the 0.01–0.05 range, while sponsored posts cluster lower. These differences align with non-parametric test outcomes and further validate that unsponsored content outperforms in audience interaction.

These visualizations collectively validate the statistical tests and further emphasize the consistent engagement advantage of unsponsored influencer posts in the beauty/fashion space.

5.3 Engagement Performance by Influencer Tier

This section examines if engagement performance varies by influencer tier, addressing the second study question: *Does influencer follower size (Nano, Micro, Celebrity) influence engagement outcomes?* The analysis was performed on both the unbalanced dataset and the final balanced, outlier-free dataset for comparison analysis.

5.3.1 Descriptive Statistics by Influencer Tier

Descriptive statistics were calculated for each influencer tier following the exclusion of outliers by the interquartile range method within each category. **Table 5.3.1** indicates that Celebrity influencers demonstrated the highest average and median engagement rates, closely followed by Micro influencers. Nano influencers, although producing the greatest number of posts in the original dataset, demonstrated the lowest median engagement rate at **1.05%**. These findings challenge the prevailing notion that Nano influencers foster better audience connections that improve engagement. The analysis suggests a likely positive association between follower count and engagement with beauty and fashion content on Instagram.

Influencer Tier	Mean Engagement Rate	Median Engagement Rate	Post Count
Celebrity	0.0326	0.0236	181
Micro	0.0280	0.0222	170
Nano	0.0174	0.0105	165

Table 5.3.1 *– Engagement Rate by Influencer Tier (IQR Outlier-Removed Dataset)*

5.3.2 Visual Comparison of Engagement by Influencer Tier

To visually explore engagement trends across influencer tiers, this section presents engagement rate distributions using two progressively refined datasets: the full raw dataset and a cleaned, balanced version. This comparison demonstrates the impact of data quality and class balance on the interpretability of influencer performance evaluations.

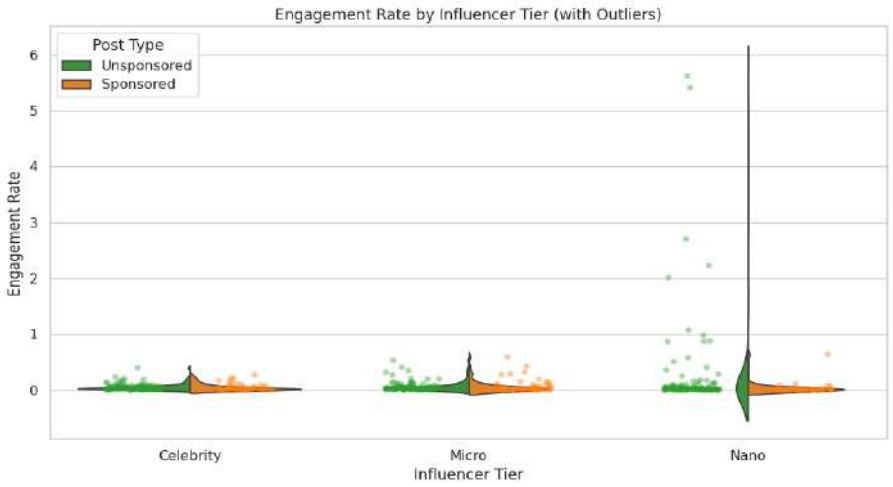

Figure 5.3.2a *– Engagement Rate by Influencer Tier (with Outliers)*

The **Figure 5.3.2a** illustrates the distribution of engagement rates among different influencer tiers using the raw dataset. A significant right skew is evident, particularly in the Nano tier, where few posts exceed engagement rates of 5.0 (i.e., 500%). These extreme numbers likely arise from viral amplification or highly specific targeting; however, they skew central tendency and augment variability, hence lowering the significance of group comparisons.

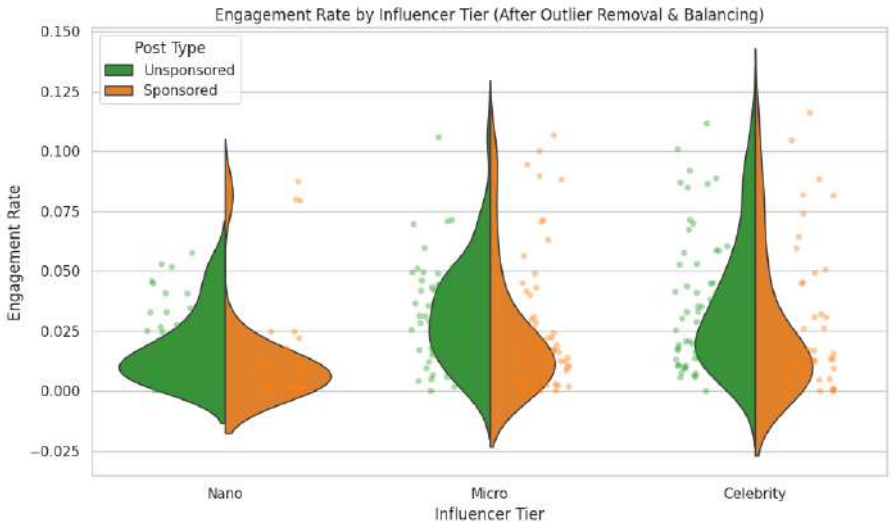

Figure 5.3.2b *– Engagement Rate by Influencer Tier (After Balancing & Removing Outliers)*

The **Figure 5.3.2b** utilizes the sanitized and balanced dataset (164 posts per category), providing an equitable and statistically interpretable perspective. It combines violin and strip charts to depict both smoothed distributions and individual post-level variability. In every category, unsponsored items consistently exhibit higher engagement than sponsored articles. The Nano tier signifies the smallest and least influential category, whereas Micro and Celebrity influencers possess wider reach and greater averages. Notably, many unsponsored posts by celebrities demonstrate exceptional performance, contradicting the belief that only minor influencers foster authentic engagement. The visual insights are corroborated by rigorous statistical tests detailed in **Section 5.3.3**.

5.3.3 Statistical Comparison by Influencer Tier

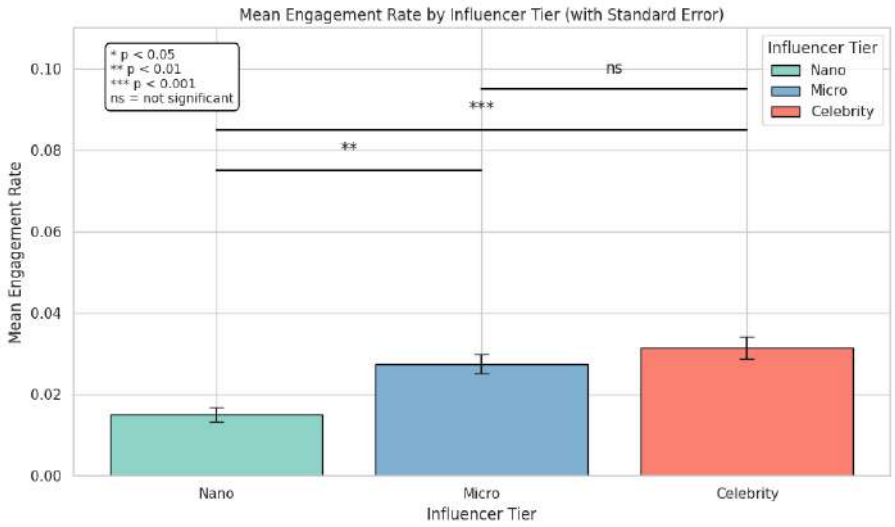

Figure 5.3.3 *– Mean engagement rate by influencer tier with standard error bars*

Non-parametric statistical testing was conducted on the balanced dataset, free of outliers (n = 328; 164 sponsored, 164 unsponsored), to assess the variance in engagement rates across different influencer levels. This method ensures that any identified disparities are not attributable to class imbalance or outliers.

A Kruskal–Wallis H test was initially utilized to examine engagement distributions across the three influencer categories: Nano, Micro, and Celebrity. The test produced a statistically significant result **(H = 29.65, p < 0.001)**, indicating that at least one group demonstrates a difference in central tendency. Pairwise Mann–Whitney U tests were used as a post-hoc study to identify the precise differences.

As shown in **Figure 5.3.3**, both Nano vs. Micro and Nano vs. Celebrity demonstrated statistically significant differences **(p < 0.001 and p < 0.01, respectively)**, however the comparison between Micro and Celebrity was not significant **(p = 0.41)**. The data indicates that Nano influencers generally exhibit lower engagement levels than both Micro and

Celebrity tiers. The engagement difference between Micro and Celebrity influencers is not statistically significant in the updated dataset.

The image includes standard error bars to represent variability in mean engagement rates, along with significance annotations for enhanced interpretation. Collectively, these findings corroborate previous descriptive insights and contest the prevailing notion that smaller influencers (e.g., Nano) inherently generate greater interaction.

5.4 Top vs. Bottom 25% Post Performance Analysis

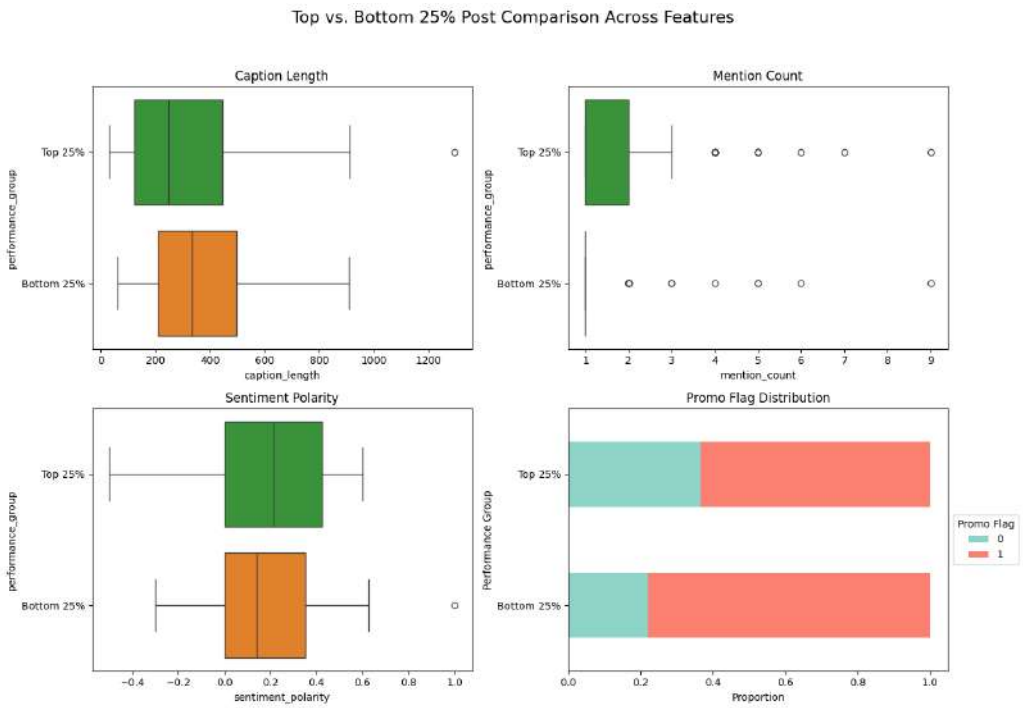

Figure 5.4a *– Comparison of post-level features between top and bottom 25% engagement groups (n = 164).*

An analysis was conducted on the upper and lower quartiles of engagement rates in the final balanced and outlier-free dataset to clarify the differences between high-performing and underperforming influencer posts. Posts within the upper 25% (≥ 75th percentile) and lower 25% (≤ 25th percentile) were distinctly identified, resulting in a focused sample of 164 posts. This section seeks to evaluate the relationship between specific structural or emotional attributes of posts and enhanced engagement results.

The chosen attributes for comparison—caption length, mention count, sentiment polarity, promotional flag, and hashtag count—were previously detailed and substantiated in **Section 4.5 (Feature Engineering)**. These criteria encompass both linguistic and promotional features

of post content. The selection of statistical tests was determined by the data type of each variable: Mann–Whitney U test was employed for continuous variables, whereas a Chi-square test of independence was utilized to assess differences in the binary promotional flag variable, as detailed in **Section 4.6.3**.

Figure 5.4a illustrates the distribution of four essential features among high and low involvement groups. The hashtag count, despite statistical analysis, was omitted from the graphic due to a highly skewed distribution (predominantly zeros or ones) that constrained its visual clarity.

The statistical analysis revealed that caption length was a crucial factor, with extended captions associated with superior performance. This discovery suggests that extended captions—potentially containing more information or narrative—may positively boost engagement.

Conversely, mention count did not exhibit a statistically significant disparity between high- and low-performing books. This indicates that tagging individuals may not inherently provoke interaction unless paired with more compelling content. Likewise, the occurrence of promotional language (promo_flag) did not significantly vary among performance **groups ($\chi^2 = 0.62$, $p = 0.43$)**. Despite prior assumptions that overtly commercial content might dampen engagement, the data suggests that promotional phrasing alone does not explain performance variation. Sentiment polarity also showed no significant difference, suggesting that the emotional tone of captions may have a limited role in driving interaction. Finally, hashtag count did not significantly differ across the performance tiers and was excluded from visualization for clarity.

Caption Similarity (Repetitiveness):

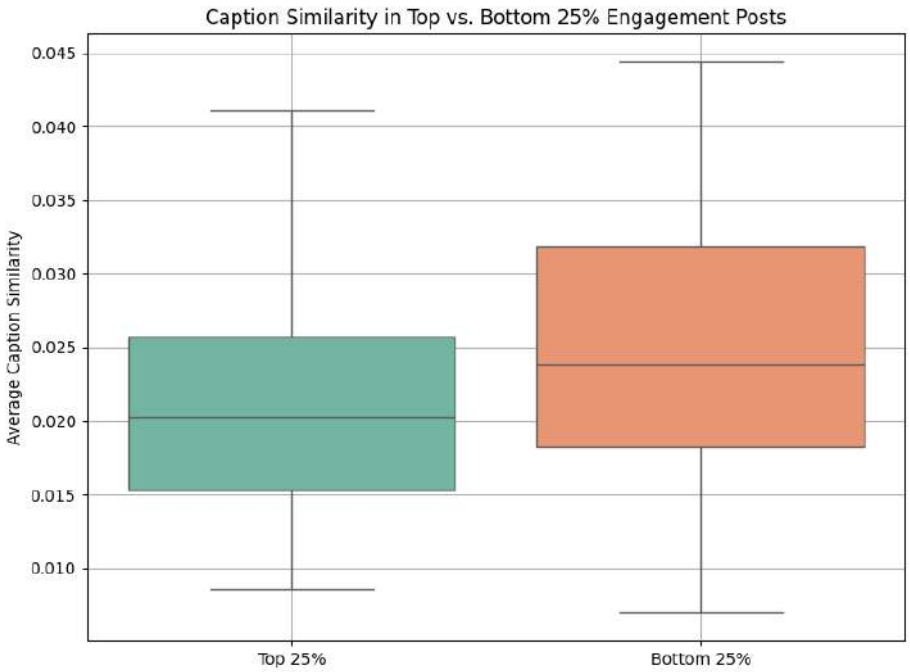

Figure 5.4b: *Caption similarity distribution in top vs. bottom 25% engagement posts.*

To assess whether repetitive or templated language contributes to lower engagement, average caption similarity was computed using **TF-IDF** vectorization and cosine similarity. Each caption's similarity score reflects how similar it is to the rest of the captions in the dataset.
As illustrated in **Figure 5.4b**, the bottom 25% of posts exhibited higher average similarity scores than top performers, suggesting more generic or reused language. This supports the idea that content repetitiveness—often characteristic of bulk promotional messaging—may reduce perceived authenticity and weaken audience interaction. A Mann–Whitney U test confirmed the statistical significance of this difference **($p < 0.01$)**, reinforcing the importance of originality in influencer caption writing.

5.5 Feature Importance via Classification Modelling

This section assesses the interpretability of the sponsorship classification model by analysing the significance of post-level factors in predicting the sponsorship status of a post. The Random Forest classifier, described in **Section 4.7**, was trained on a definitive, balanced, and outlier-free dataset (n = 328), employing five engineered features: caption length, hashtag count, mention count, sentiment polarity, and a promotional flag denoting the presence of promotional keywords in the caption.

The model's performance, illustrated in **Figure 5.5a**, exhibited a commendable level of predictive accuracy, with the Random Forest attaining 36 correct predictions out of 42 for sponsored posts and 32 correct predictions out of 41 for unsponsored posts in the test set. This indicates that the algorithm can effectively identify essential signals that differentiate commercial content from organic articles, despite the limited feature set.

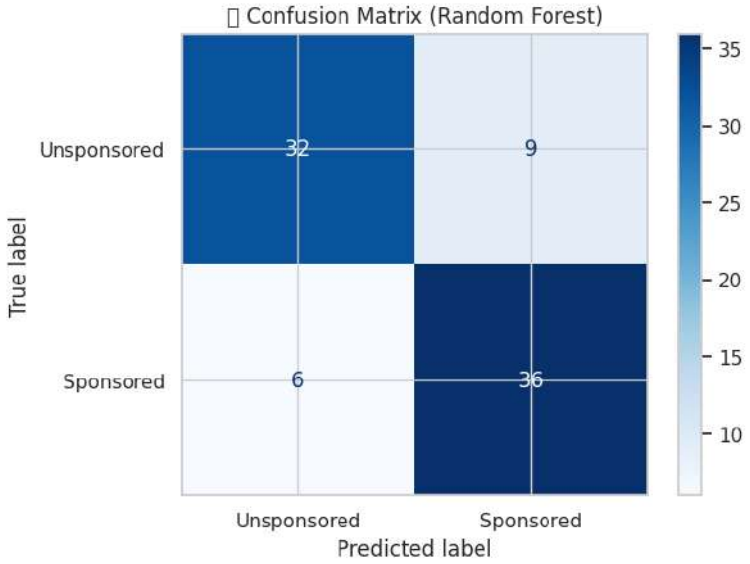

Figure 5.5a *– Confusion matrix for Random Forest sponsorship classifier.*

The feature importance values derived from the trained model **(Figure 5.5b)** reveals the content-level attributes that were crucial in the classification process. The promotional flag emerged as the primary predictor, accounting for over 40% of the overall model significance—an anticipated result, as the presence of explicit indicators such as “#ad” or “discount code” clearly signifies a sponsored post. The length of captions was the second most significant factor, indicating that sponsored articles frequently contain more comprehensive or elaborate information. Supplementary attributes, such as sentiment polarity, hashtag frequency, and mention frequency, exerted a diminished influence, consistently enhancing the remaining predictive signal.

The results highlight the essential function of clear promotional indicators and structural characteristics in differentiating sponsored content on Instagram. Linguistic mood and network mentions provide some value; yet, they appear secondary to overt indicators of advertising intent.

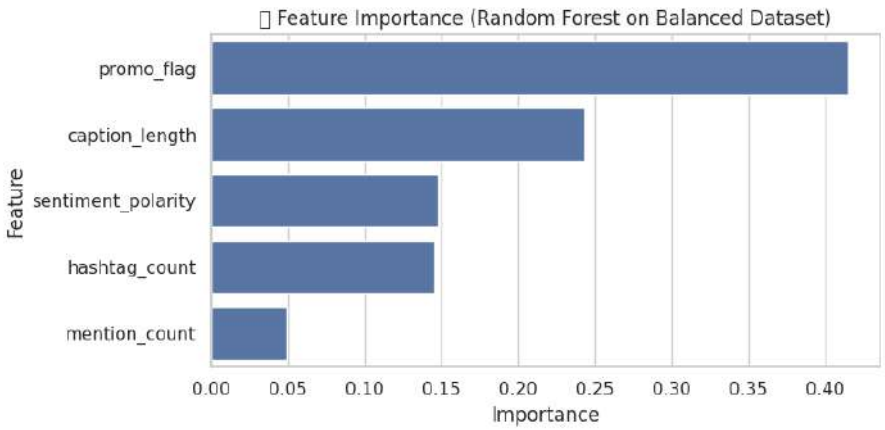

Figure 5.5b – Feature importance derived from the trained Random Forest model

5.6 Timing and Posting Strategy Insights

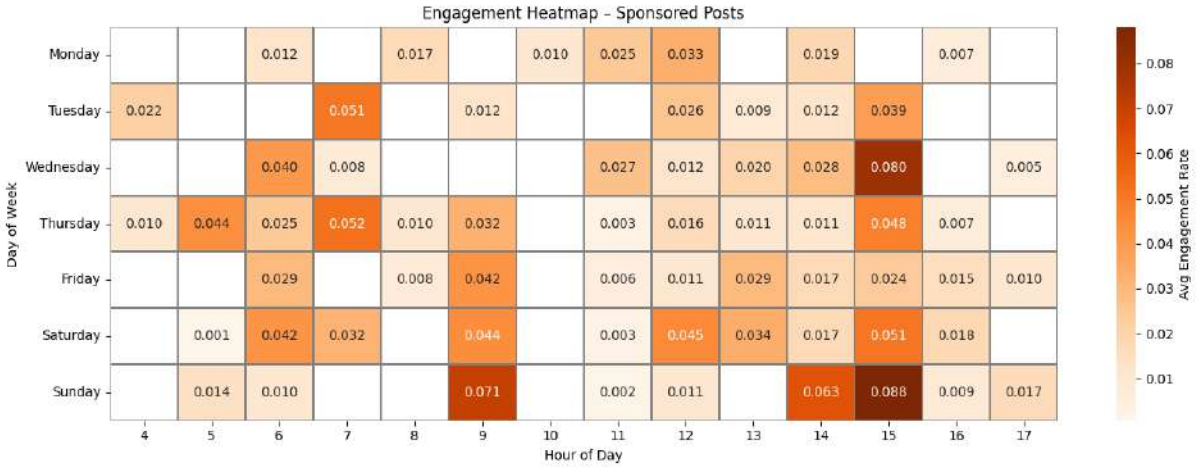

Figure 5.6a *– Engagement heatmap for sponsored posts by hour and weekday.*

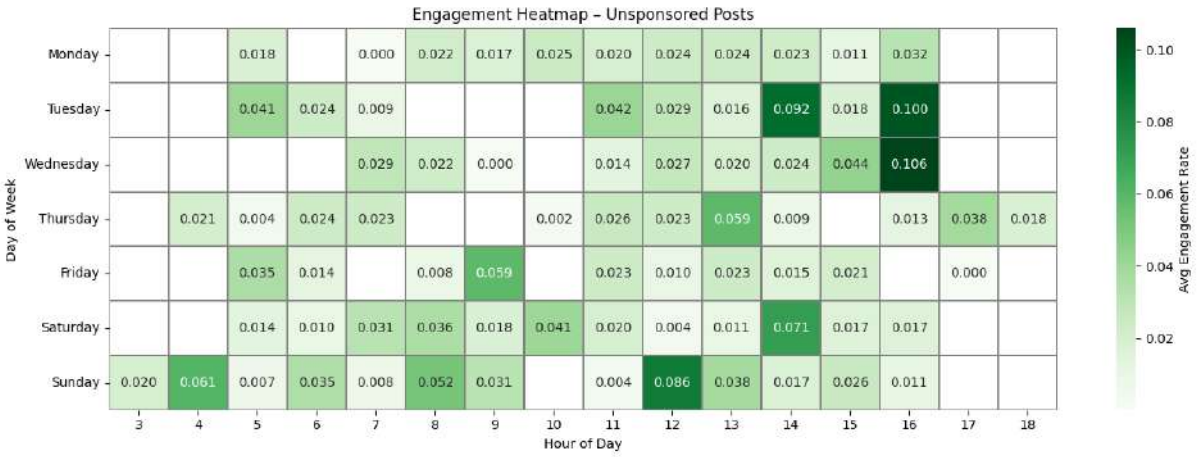

Figure 5.6b – Engagement heatmap for unsponsored posts.

This section assesses the influence of post timing on engagement performance by examining average engagement rates across two temporal dimensions: hour of the day and day of the week. Two heatmaps were generated utilizing the final balanced dataset, free from outliers (n = 328, 164 posts per category), to depict the fluctuations in engagement over various time intervals for sponsored and unsponsored content separately.

Figure 5.6a demonstrates that sponsored posts achieved peak engagement between **14:00 and 16:00 on weekdays**, particularly on **Wednesday and Sunday afternoons**. Engagement demonstrated increased fluctuation across hours and days, with notable reductions occurring during the early morning and late evening. This pattern suggests that paid content is most effective when published during periods of heightened user engagement, either due to the success of deliberately timed marketing efforts or algorithmic boosting of advertisements.

In contrast, **Figure 5.6b** illustrates that unsponsored posts maintained constant performance over a broader range of time intervals. Weekday afternoons (15:00–17:00) and Sunday mornings exhibited heightened involvement, with optimal performance recorded on Wednesday at 16:00. This decentralized model suggests that organic material flourishes with more adaptable interaction intervals and may exhibit reduced sensitivity to exact timing.

The findings together underscore the strategic importance of altering the timing of posts, particularly for sponsored content. While unsponsored content can thrive in various contexts, businesses and influencers should prioritize weekday afternoons for scheduling promotional material, as these periods are associated with increased audience engagement and platform activity.

5.7 – Individual Influencer Performance: Sponsored vs. Unsponsored Engagement

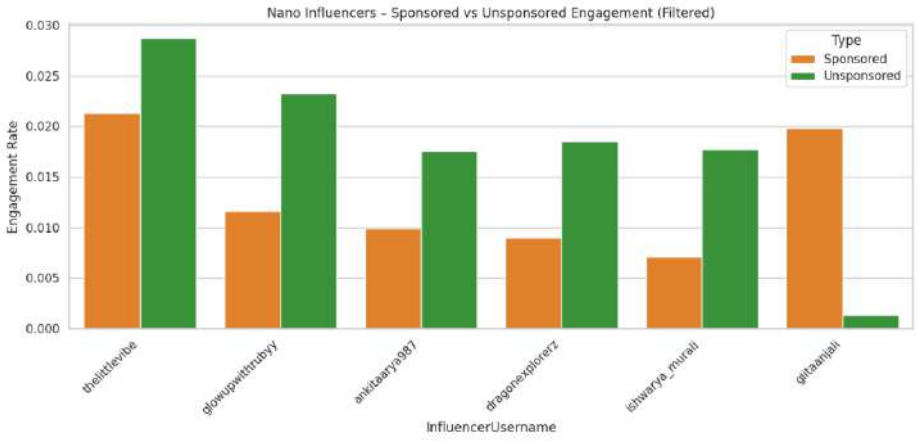

Figure 5.7a *– Top Performing Nano Influencers – Sponsored vs. Unsponsored Engagement*

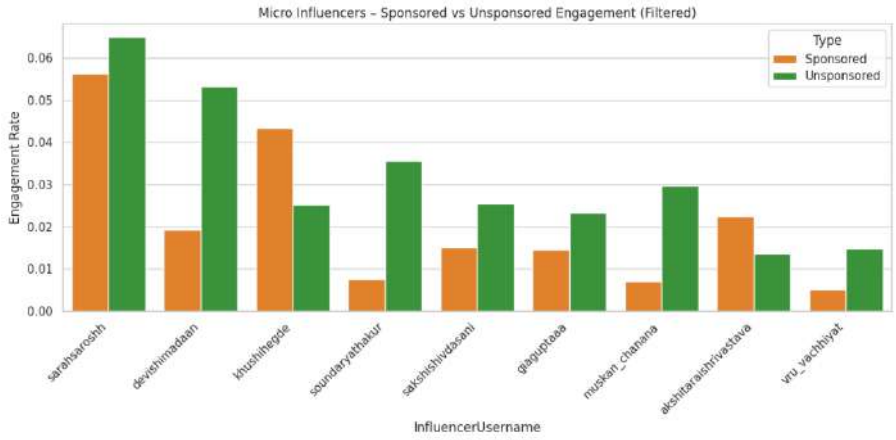

Figure 5.8b *– Top Performing Micro Influencers – Sponsored vs. Unsponsored Engagement*

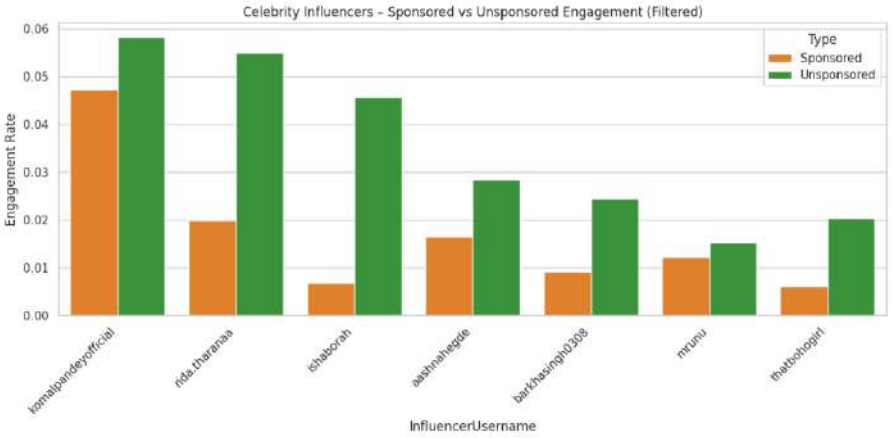

Figure 5.8c *– Top Performing Celebrity Influencers – Sponsored vs. Unsponsored Engagement*

To elucidate the impact of sponsorship on individual influencer performance, we illustrated average engagement rates for the highest-performing influencers across Nano, Micro, and Celebrity categories **(Figure 5.8 a,b,c)**. This comparison encompasses just those influencers who possessed enough quantity of both sponsored and unsponsored posts inside the balanced and outlier-free dataset, so guaranteeing an equitable assessment.

In all three tiers, unsponsored posts consistently garnered greater interaction than sponsored posts for most influencers. For instance, among micro-influencers, **sarahsaroshh** and **devishimadaan** exhibited robust engagement overall; nevertheless, their unsponsored content surpassed the performance of sponsored posts, corroborating the previously noted tendency that promotional content marginally diminishes audience receptivity. **Khushihegde** stands out as an outlier, demonstrating considerable sponsored engagement, highlighting her potential as a dependable collaborator for brands.

Likewise, **thelittlevibe** and **glowupwithrubyy** distinguished themselves in the Nano category, sustaining comparatively strong engagement for both post categories, but with a noticeable decline when sponsorship was included. Among celebrity influencers, **komalpandeyofficial** and **rida.tharanaa** exhibited superior performance, with **komalpandeyofficial** seeing little engagement decline across various post categories, suggesting a very devoted fan base.

The results reveal that while prominent influencers in each category maintain strong audience engagement, sponsorship consistently has a negative impact. Brands must carefully evaluate influencer reach and past engagement with campaigns before finalizing partnerships.

6. Discussion

6.1 Do Sponsored Beauty/Fashion Posts Underperform in Engagement Compared to Unsponsored Ones?

Analysis of both unbalanced and balanced datasets consistently demonstrates that sponsored posts have significantly lower engagement rates compared to unsponsored ones. The raw dataset revealed that sponsored material received few likes and comments, suggesting possible "ad fatigue" or diminished audience confidence in overt marketing, a phenomenon noted by Abidin (2016) and Djafarova & Rushworth (2017). Following the elimination of outliers and the equalization of the dataset (164 posts per category), the engagement discrepancy persisted across all influencer classifications (Nano, Micro, and Celebrity), as confirmed by Kruskal–Wallis and paired Mann–Whitney U tests.

A distinct pattern emerged: unsponsored posts typically garnered better engagement, with notable variations especially evident between Nano and Micro tiers, as well as between Nano and Celebrity tiers. Even among comparisons of individual influencers, unsponsored material had a greater number of likes and comments. This indicates that perceived authenticity and relatability, typically linked to non-promotional content, are crucial factors in driving engagement. These results correspond with previous studies indicating that audiences exhibit greater responsiveness to content perceived as less commercial and more organic (Evans et al., 2017).

The evidence unequivocally substantiates the premise that sponsored beauty and fashion posts exhibit inferior performance relative to unsponsored posts, both statistically and visually.

6.2 What Role Do Content-Level Features Such as Sentiment Tone, Hashtags, and Mentions Play in Post Performance?

To explore how specific content characteristics influence engagement, we compared the top and bottom 25% of posts based on engagement rate across several post-level features: caption length, hashtag count, mention count, sentiment polarity, and the presence of promotional language. These features were selected for their frequent appearance in prior influencer marketing research and their interpretability in practice.

The results show that the best posts had far longer captions. This backs up what Lou and Yuan (2019) found: that consumers respond more to content that is more valuable, like how informative and detailed it is.

On the other hand, the number of hashtags, the polarity of the sentiment, and the number of mentions did not demonstrate statistically significant differences between the high- and low-performing groups. People typically think that hashtags make things easier to find (Sheldon & Bryant, 2016), however our results show that the effect on engagement may rely more on the situation than on how often they are used. In the same way, sentiment polarity, which shows the emotional tone, was not a good way to forecast how well a post would fare in this area. The number of mentions also did not have a big effect, which means that tagging other people does not always lead to more engagement unless the content is interesting.

But a caption similarity analysis indicated an obvious pattern: publications that did not fare well tended to have more captions that were the same or looked like templates. These posts got far higher similarity scores, which suggests they were not very original. This new finding confirms what we previously knew: using the same words repeatedly may make individuals less interested, while using creative and personalized messages may help people do better.

The bottom 25% of postings had more visual promotional cues (as shown by the promo_flag), although this difference was not statistically significant ($\chi^2 = 0.62$, $p = 0.43$). This means that using obvious commercial language alone does not really set apart the best articles from the worst ones, which goes against what several people in the field thought. things could not be the presence of such language that changes how people see things, but how well it is used and integrated.

In short, of all the content-level variables tested, caption richness and uniqueness were the most consistent indicators of high engagement. In the world of beauty and fashion influencers, other things like tone of sentiment, mentions, and promotional language did not seem to matter as much, at least not on their own.

6.3 How Does Influencer Tier Affect Engagement Dynamics Across Content Types?

This study looked at how the level of influencer (Nano, Micro, or Celebrity) and whether the content is sponsored effect how people interact with it. The statistics showed that the levels of involvement were considerably varied at each level. This gave businesses who were trying to reach diverse groups of people a lot of helpful information.

Unpaid postings always did better than paid ones, but the gap was bigger at some levels than at others. Nano influencers noticed the biggest difference between encounters that were paid for and those that were not. This means that ads are more likely to function on the followers of Nano authors. They might want greater honesty and think of the artists as friends. When influencers like **thelittlevibe** and **ankitaarya987** uploaded stuff that were not paid for, they did well. But their performance dropped a lot when they added sponsorship. Micro influencers had a tendency that was steadier. The difference was smaller, but content

that was not sponsored still did better. Influencers like **sarahsaroshh** and **devishimadaan** kept their fans interested even when they were compensated to post.

Micro influencers showed a trend that was more even. The difference was smaller, but content that was not sponsored still did better. Influencers like sarahsaroshh and devishimadaan maintained their followers interested even when they were paid to post. This shows that micro influencers may be able to establish a decent balance between reach and trustworthiness. This is what other studies have found: micro influencers are the ideal people to work with to keep people interested (De Veirman et al., 2017).

Celebrity influencers had a lot of engagement on both types of posts, but posts that were not sponsored did better than posts that were. Even though their sponsored posts showed a small but perceptible dip, influencers like **komalpandeyofficial** and **rida.tharanaa** enjoyed high levels of engagement all the time. This implies that even while celebrity content is popular and well-made, it can nevertheless get tired of being sponsored.

These data reveal that the level of an influencer has a major effect on how consumers respond to ads. All types of influencers are less likely to interact with sponsored posts, but Nano influencers are the most affected. On the other hand, micro and celebrity influencers may fare better with all types of content. This illustrates that marketers need to think about more than simply how many people they want to reach when they want to work with influencers. They also need to think about how their audience interacts with each level.

6.4 Can Machine Learning Help Predict What Makes a Post More Likely to Be Sponsored—and Less Likely to Perform Well?

To answer this question, we created a Random Forest classifier that could tell if a post was sponsored based on things like the length of the caption, the number of hashtags, the number of mentions, the polarity of the emotion, and the promo flag. This model trained on a dataset that was balanced and did not have any data points that were too far out from the others. This made sure that the number of sponsored and non-sponsored postings was the same. The classifier could determine the difference between sponsored and non-sponsored content with a moderate amount of accuracy. The most important thing for creating forecasts was the use of promotional language. Next came the length of the captions and the polarity of the sentiment. These results show that machine learning can detect patterns in sponsorship data, even when the data is modified to make it look like it belongs there.

It is interesting that several of the variables that assisted the most in predicting sponsorship, like employing promotional language and shorter subtitles, were also linked to reduced participation in earlier studies. I believe that the machine learning model could not only recognize what kind of sponsorship it is, but it could also find patterns in the material that make people less interested because of this overlap.

This model was not supposed to predict direct engagement, but it does show that machine learning may find patterns in material that effect both visibility and engagement. This is in line with what other research have found: tree-based machine learning algorithms like Random Forest are good in getting people to utilize social media (Wu et al., 2024; Chen et al., 2022). As content marketing gets better, these kinds of models can be used to evaluate content before it goes live, work better with influencers, or make influencer campaigns more aligned with the brand.

7. Limitations

While this study offers valuable insights into influencer engagement patterns, a few limitations are worth noting. The analysis focused on 30 beauty and fashion influencers, which may limit how broadly the findings apply to other industries or platforms.

The research was based entirely on caption text and metadata, without examining the visual aspects of posts—an important factor on a visually driven platform like Instagram. Sponsorship identification relied on keyword detection, which is practical and transparent but may miss more subtle or undeclared partnerships.

Finally, this study captures a single moment in time. It does not account for seasonal trends or how influencer performance might change across different campaign cycles. Future work could explore these dynamics by incorporating visual content, expanding to other influencer categories, and using more advanced models to build on these findings.

8. Final Summary and Reflections

This study examined the impact of sponsorship on engagement with influencer posts within the beauty and fashion sectors on Instagram. This study analyses a dataset of 1,500 posts from 30 influencers, investigating patterns in content attributes, influencer categories, and engagement outcomes, thus offering a comprehensive, data-driven examination of the correlation between promotional activities and organic impacts.

The research was directed by four principal inquiries on (1) the impact of sponsorship on engagement, (2) the effects of content-level attributes, (3) how engagement varies across influencer tiers, and (4) the capacity of machine learning to identify sponsored content. The

answers to these questions collectively suggest that sponsored posts generally encounter reduced interaction, often due to stylistic and linguistic cues that signal marketing to users. Among content attributes, caption length and originality showed meaningful associations with post-performance, while other factors like sentiment and mentions were less predictive; nonetheless, nano, and micro influencers typically generate stronger audience responses than celebrities, particularly for organic material.

The development of a Random Forest-based classifier indicates that post-level metadata can effectively predict sponsorship likelihood, impacting automated campaign evaluations and influencer analyses. This method, while not intended for implementation, served as a potent analytical tool to uncover complex relationships between content structure and marketing objectives.

The analysis of caption similarity indicated that content redundancy negatively impacts engagement, highlighting the necessity for unique, non-formulaic language in the effectiveness of influencer marketing.

8.1 Reflections on the Research Journey

This study faced numerous technical and conceptual challenges, including the requirement for balanced datasets, the development of meaningful engagement metrics, the creation of a classification system, and the analysis of complex behavioural patterns. Surmounting these obstacles improved my technical proficiency in data preparation, feature engineering, and model evaluation, while also augmenting my understanding of ethical considerations, such as disclosure transparency and the authenticity-performance trade-off in digital marketing. The choice to focus exclusively on captions and metadata, though methodologically deliberate, underscored the limitations of analysing social media content devoid of visual context. This constraint required the assessment of the importance of multimodal research approaches in future investigations.

8.2 Practical Implications

This study offers some practical insights for influencers, businesses, and marketers aiming to improve content effectiveness on Instagram:

1. **Minimize overt promotional signals:** While not statistically conclusive, postings using commercial language were somewhat disproportionately found in the lowest 25%, suggesting that consumers may disengage when promotional content is extremely apparent or insincere. Discreet, seamlessly incorporated brand references may yield superior results compared to aggressive sales strategies.

2. **Invest in meaningful captions:** Caption length was a clear differentiator of engagement. Long-form, story-driven posts outperformed short or templated ones—highlighting the value of adding context, personality, or narrative richness to posts.
3. **Avoid copy-paste templates**: Caption similarity was significantly higher in underperforming posts. Repetitive, formulaic phrasing may be perceived as insincere. Brands and influencers should prioritize originality in messaging.
4. **Timing and tier alignment:** Sponsored posts showed distinct temporal posting patterns. Brands could benefit from aligning posting schedules with typical organic activity patterns rather than fixed promotional slots. Moreover, while celebrities had strong performance overall, micro-influencers often achieved similar results for unsponsored content—making them efficient partners for non-paid brand awareness.
5. **Use machine learning for screening:** The Random Forest model achieved high accuracy in detecting sponsored posts based on content and metadata. Brands can adopt lightweight ML models to pre-screen influencer content for promotional signals or predict likely campaign performance before launch.

8.3 Concluding Remarks

This dissertation advances the current literature on influencer marketing by demonstrating the quantifiable performance differences between organic and sponsored content, identifying critical content attributes linked to engagement, and utilizing machine learning to enable sponsorship prediction. Despite its limitations in scale and scope, the study provides a foundation for future multimodal, longitudinal, and cross-platform investigations into the evolving dynamics of digital influence.

This research emphasizes that successful influencer marketing relies not only on reach or follower counts but also on intentional content production, authentic communication, and audience understanding—a notion relevant to data scientists, marketers, and producers.

9 Appendices

Appendix A – Influencer Sample

Influencer Handle	Follower Count	Tier
@mruchabeauty	5000000	Celebrity
@the.rebel.kid	4000000	Celebrity
@barkhasingh0308	3000000	Celebrity
@rida.tharanaa	2000000	Celebrity
@komalpandeyofficial	2000000	Celebrity
@ishaborah	2000000	Celebrity
@thatbohogirl	2000000	Celebrity
@roykapurfilms	2000000	Celebrity
@swativerma	2000000	Celebrity
@aashnahegde	1000000	Celebrity
@devishimadaan	439000	Micro
@sakshishivdasani	419000	Micro
@vru_vachhiyat	352000	Micro
@soundaryathakur	325000	Micro
@debasreee	314000	Micro
@akshitaraishrivastava	298000	Micro
@muskan_chanana	266000	Micro
@kullubaaazi	242000	Micro
@giaguptaaa	242000	Micro
@sarahsaroshh	240000	Micro
@glowupwithrubyy	98900	Nano
@ankitaarya987	81500	Nano
@ashima_banga	73800	Nano
@ishwarya_murali	63900	Nano
@radhica_diwan	57700	Nano
@khushi.chhikara	56400	Nano
@dragonexplorerz	51900	Nano
@thelittlevibe	43000	Nano
@mohinibagul_19	40000	Nano
@giitaanjali	33200	Nano

Appendix B – Sponsorship Keyword List

The following keywords were used to classify posts as sponsored. These were matched against caption text using case-insensitive substring matching.

#ad, #ads, #paid, paid partnership, sponsored, #sponsored, #collab, collaboration, affiliate, promotion, #pr, gifted, #gifted, brand, discount, offer, deal, use code, use my code, link in bio, promo, shop now, ambassador

Appendix C – Feature Engineering Overview

Feature Name	Description
caption_length	Number of characters in the caption text
hashtag_count	Number of hashtags used in the caption
mention_count	Number of users mentions in the caption
promo_flag	Binary indicator for presence of promotional language
sentiment_polarity	Polarity score from TextBlob (range: -1 to +1)
engagement_rate	Engagement = (likes + comments) / follower count
is_sponsored	Binary label for sponsored vs. unsponsored (from keyword rules)

Appendix D – Modelling Formula and Evaluation Metrics

This section outlines the modelling approach and performance metrics used in classifying posts as sponsored or unsponsored.

• Model: Random Forest Classifier (RFC)

• Input Features: caption_length, mentions_count, hashtags_count, sentiment_polarity, promo_flag, followerCount

• Output: Binary class (1 = Sponsored, 0 = Unsponsored)

Evaluation Metrics:

• Accuracy
• Precision
• Recall
• F1-Score
• ROC-AUC

Formulas Used:

Engagement Rate: (likes + comments) / follower count

Cohen's d Effect Size: d = (M1 – M2) / pooled standard deviation

Chi-square Test: $\chi^2 = \Sigma (O – E)^2 / E$

Appendix E – Caption Similarity Computation

This section describes how caption repetitiveness was measured between top and bottom performing posts.

• Vectorization method: TF-IDF using sklearn

• Similarity computation: Cosine similarity between caption pairs

• Objective: Identify if low-performing posts use more formulaic or repeated language.

• Libraries used: TfidfVectorizer, cosine_similarity from sklearn

Appendix F – Ethical Considerations

All data collected for this study was publicly accessible on Instagram. The research focused solely on post-level metadata and captions, excluding the acquisition of private user data, images, or personal identifiers. No direct interaction with users transpired. This approach adheres to GDPR and institutional ethical standards for responsible data use. Any referenced brand or influencer names are utilized solely for academic investigation without evaluative assessment.

Appendix G – Tools and Software Used

This appendix outlines all software tools, libraries, and frameworks used in the data collection, processing, analysis, and modelling stages of the study.

1. Data Collection

- **Apify Instagram Scraper**
  Used to collect publicly available post-level metadata (captions, hashtags, mentions, likes, comments, timestamps) from selected influencer profiles. Apify is a cloud-based web scraping platform that supports structured extraction from social media sites.

2. Programming Environment

- **Python 3.10**
  The core programming language used for all stages of data cleaning, analysis, and modelling.

3. Data Manipulation & Cleaning

- **Pandas**
  Used for data frame operations, filtering, and transformation of structured Instagram post data.
- **NumPy**
  Used for numerical operations, array processing, and statistical computations.

4. Feature Engineering & Text Processing

- **Scikit-learn (sklearn)**
  Utilized for:
  - TF-IDF vectorization for caption similarity
  - Random Forest classification modelling
  - Evaluation metrics (accuracy, F1-score, ROC-AUC)
- **TextBlob**
  Used for sentiment polarity extraction from captions, providing a score between –1 and +1.

- **Regular Expressions (re in Python)**
  Applied to identify beauty/fashion-related keywords and promotional terms in captions and hashtags.

- **Datetime (Python standard library)**
  Used for extracting and transforming post timestamps into day/time buckets.

5. Statistical Analysis

- **SciPy (scipy.stats)**
  Used for Welch's t-test, Mann–Whitney U test, Levene's test, and Kruskal–Wallis H test.

- **scikit-posthocs**
  Used for conducting Dunn's post-hoc test with Bonferroni correction on tier-wise engagement data.

6. Visualization

- **Matplotlib**
  For generating bar plots, boxplots, histograms, and confusion matrices.

- **Seaborn**
  Used for enhanced visualizations including violin plots, KDE plots, and heatmaps.

7. Modelling & Evaluation

- **Scikit-learn Models:**
  - **Logistic Regression** – baseline linear classifier for sponsorship prediction.
  - **Random Forest Classifier** – main ensemble model used for classification with feature importance analysis.
  - **Support Vector Machine (SVM)** – kernel-based classifier used for comparison.
- **Confusion Matrix & ROC-AUC Evaluation** – implemented via sklearn.metrics.

8. Ethical Compliance

- **No external image processing or facial detection tools** were used. Only textual, publicly visible metadata was analysed.

10. References

Abidin, C. (2016) *Please Subscribe! Influencers, Social Media, and the Commodification of Everyday Life*. PhD thesis. University of Western Australia. Available at: https://api.research-repository.uwa.edu.au/ws/files/9781681/Abidin_Crystal_2016.pdf (Accessed: July 2025).

Ailawadi, S. & Sehgal, V. (2024) 'Evaluating ROI in Digital Marketing Campaigns: Metrics, Measurement, and Insights', *ResearchGate*. Available at: https://www.researchgate.net/publication/377598832 (Accessed: July 2025).

An, J., Wang, H. & Zhang, T. (2020) 'Sentiment Analysis of Student Engagement with Lecture Recording', *Education and Information Technologies*, 25(3), pp. 11528. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC7787655/pdf/11528_2020_Article_563.pdf (Accessed: July 2025).

Batrinca, B. & Treleaven, P.C. (2015) 'Social media analytics: a survey of techniques, tools and platforms', *AI & Society*, 30(1), pp. 89–116. https://doi.org/10.1007/s00146-014-0549-4 (Accessed: July 2025).

Boerman, S.C., Willemsen, L.M. & Van Der Aa, E.P. (2017) '"This post is sponsored": Effects of sponsorship disclosure on persuasion knowledge and electronic word of mouth in the context of Facebook', *Journal of Interactive Marketing*, 38, pp. 82–92. Available at: https://www.researchgate.net/publication/316284528 (Accessed: July 2025).

Boerman, S.C. & van Reijmersdal, E.A. (2020) 'Disclosing influencer marketing on YouTube to children: The moderating role of para-social relationship', *Frontiers in Psychology*, 11, 547386. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7558032/ (Accessed: July 2025).

Chen, A., Patel, N. & Gupta, R. (2022) 'Leveraging Neural Networks and NLP for Enhanced Influencer Marketing Analytics', *EAAIJ*, 11(8). Available at: https://eaaij.com/index.php/eaaij/article/view/43/43 (Accessed: July 2025).

De Veirman, M., Cauberghe, V. & Hudders, L. (2017) 'Marketing through Instagram influencers: The impact of number of followers and product divergence on brand attitude', *International Journal of Advertising*, 36(5), pp. 798–828. Available at: https://www.academia.edu/104156917 (Accessed: July 2025).

Djafarova, E. & Rushworth, C. (2017) 'Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users', *Computers in Human Behavior*, 68, pp. 1–7. https://doi.org/10.1016/j.chb.2016.11.009 (Accessed: July 2025).

Ershov, D., Hernandez, M.A. & Xu, J. (2024) 'Undisclosed Influencer Advertising: Evidence from Social Media', *CESifo Working Paper No. 10743*. Available at: https://www.ifo.de/DocDL/cesifo1_wp10743.pdf (Accessed: July 2025).

Evans, N.J., Phua, J., Lim, J. & Jun, H. (2022) 'Influencer Marketing on Instagram: Empirical Research on Social Media Engagement with Sponsored Posts', *Journal of Interactive Advertising*, 17(2). Available at: https://www.tandfonline.com/doi/full/10.1080/15252019.2022.2123724 (Accessed: July 2025).

Ghasemi, A. & Zahediasl, S. (2012) 'Normality tests for statistical analysis: A guide for non-statisticians', *International Journal of Endocrinology and Metabolism*, 10(2), pp. 486–489. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3693611/ (Accessed: July 2025).

Johnson, A. & Lee, B. (2023) 'Social Media Influencer Marketing: Foundations, Trends, and Ways Forward', *Electronic Commerce Research*. Available at: https://link.springer.com/article/10.1007/s10660-023-09719-z (Accessed: July 2025).

Kim, S. & Kang, M. (2022) 'Influencer Marketing in the Digital Age: The Response to Authentic Creator Content', *ResearchGate*. Available at: https://www.researchgate.net/publication/359050795 (Accessed: July 2025).

Kumar, R. & Singh, P. (2024) 'Sentiment Analysis Methods, Applications, and Challenges: A Systematic Literature Review', *Science and Engineering Ethics*. Available at: https://www.sciencedirect.com/science/article/pii/S131915782400137X (Accessed: July 2025).

Kumar, R., Gupta, A. & Sharma, S. (2022) 'Emotions and Brand Engagement on Instagram: A Deep Learning Perspective', *PMC*. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC9007765/ (Accessed: July 2025).

Libai, B. et al. (2025) 'Influencer Marketing Unlocked: Understanding the Value Chains Driving the Creator Economy', *Journal of the Academy of Marketing Science*, 53(1), pp. 4–28. https://doi.org/10.1007/s11747-024-01073-2 (Accessed: July 2025).

Li, Y. & Wang, X. (2021) *An Empirical Study of User Engagement in Influencer Marketing on Weibo and WeChat*. Unpublished manuscript provided by Prof. Isilay. (Accessed: July 2025).

Lou, C. & Yuan, S. (2019) 'Influencer marketing: How message value and credibility affect consumer trust of branded content on social media', *Journal of Interactive Advertising*, 19(1), pp. 58–73. Available at: https://www.researchgate.net/publication/328620597 (Accessed: July 2025).

Lou, C. & Kim, H.K. (2019) 'Fancying the New Rich and Famous? Explicating the Roles of Influencer Content Features and Source Credibility in Effective Native Advertising', *Journal of Interactive Advertising*, 19(2), pp. 134–147. Available at: https://www.researchgate.net/publication/337048906 (Accessed: July 2025).

Leys, C. et al. (2013) 'Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median', *Journal of Experimental Social Psychology*, 49(4), pp. 764–766. https://doi.org/10.1016/j.jesp.2013.03.013 (Accessed: July 2025).

Müller, L., Mattke, J. & Maier, C. (2018) '#Sponsored #Ad: Exploring the Effect of Influencer Marketing on Purchase Intention', *AMCIS 2018 Proceedings*. Available at: https://aisel.aisnet.org/amcis2018/DataScience/Presentations/29/ (Accessed: July 2025).

Novak, P.K. et al. (2021) 'Measurement of engagement rate on Instagram for business marketing', *PMC*. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC8354841/ (Accessed: July 2025).

Olteanu, A. et al. (2019) 'Social data: Biases, methodological pitfalls, and ethical boundaries', *Frontiers in Big Data*, 2, p. 13. https://doi.org/10.3389/fdata.2019.00013 (Accessed: July 2025).

Patel, D. & Reddy, S. (2023) 'Leveraging the BERT Model for Enhanced Sentiment Analysis in Multicontextual Social Media Content'. *Conference paper*. Available at: https://www.researchgate.net/publication/382232999 (Accessed: July 2025).

Rival IQ (2023) *2023 Social Media Industry Benchmark Report*. Available at: https://get.rivaliq.com/hubfs/eBooks/2024-Rival-IQ-Social-Media-Industry-Benchmark-Report.pdf (Accessed: July 2025).

Samanta, P. et al. (2025) 'A Critical Review of Influencer Marketing's Influence on Brand Perception and Consumer Buying Decisions', *Journal of Information Systems Engineering & Management*, 10(29s), pp. 822–838. https://doi.org/10.52783/jisem.v10i29s.4594 (Accessed: July 2025).

Saragih, G.N. & Girsang, A.S. (2016) 'Sentiment Analysis of Customer Engagement on Social Media in Transport Online'. Available at: https://core.ac.uk/download/pdf/187726108.pdf (Accessed: July 2025).

Sheldon, P. & Bryant, K. (2016) 'Instagram: Motives for its use and relationship to narcissism and contextual age', *Computers in Human Behavior*, 58, pp. 89–97. https://doi.org/10.1016/j.chb.2015.12.059 (Accessed: July 2025).

Sloan, L. et al. (2015) 'Who tweets? Deriving the demographic characteristics of age, occupation and social class from Twitter user metadata', *PLOS ONE*, 10(3), e0115545. https://doi.org/10.1371/journal.pone.0115545 (Accessed: July 2025).

Stieglitz, S. & Dang-Xuan, L. (2013) 'Emotions and information diffusion in social media—Sentiment of microblogs and sharing behavior', *Journal of Management Information Systems*, 29(4), pp. 217–248. Available at: https://www.researchgate.net/publication/259704042 (Accessed: July 2025).

Stieglitz, S. et al. (2018) 'Social media analytics – Challenges in topic discovery, data collection, and data preparation', *International Journal of Information Management*, 39, pp. 156–168. https://doi.org/10.1016/j.ijinfomgt.2017.12.002 (Accessed: July 2025).

Wu, X. et al. (2024) 'Predictors of Social Media Influencer Marketing Effectiveness: A Comprehensive Literature Review and MetaAnalysis', *Journal of Business Research*, 187. Available at: https://www.sciencedirect.com/science/article/pii/S0148296324004958 (Accessed: July 2025).

Zarei, F., Farajtabar, M. & Zafarani, R. (2020) 'Detecting and monitoring advertising disclosures on social media', *arXiv preprint arXiv:2011.05757*. Available at: https://arxiv.org/abs/2011.05757 (Accessed: July 2025).

Zhang, Y., Chen, L. & Wang, Z. (2024) 'Influencer Marketing on Instagram—The Optimal Disclosure Strategy from Influencers' and Marketers' Perspectives', *Electronic Markets*, 34(2). https://doi.org/10.1007/s12525-024-00743-x (Accessed: July 2025).

Audrezet, A., de Kerviler, G. & Moulard, J.G. (2020) 'Authenticity under threat: When social media influencers need to go beyond self-presentation', *Journal of Business Research*, 117, pp. 557–569. https://doi.org/10.1016/j.jbusres.2018.07.008 (Accessed: July 2025).

Sesar, V., Martinčević, I. & Boguszewicz-Kreft, M. (2022) 'The Relationship Between Advertising Disclosure, Influencer Credibility, and Consumer Purchase Intention', *Journal of Risk and Financial Management*, 15(7), p.276. https://www.mdpi.com/1911-8074/15/7/276 (Accessed: July 2025).

Hootsuite (2025) *Average Engagement Rates for 12 Industries*. Available at: https://blog.hootsuite.com/average-engagement-rate/ (Accessed: July 2025).

Hootsuite (2025) *Best Time to Post on Instagram in 2025: What the Data Says*. Available at: https://blog.hootsuite.com/best-time-to-post-on-instagram/ (Accessed: July 2025).

Evans, N.J., Phua, J., Lim, J. & Jun, H. (2017) 'Disclosing Instagram influencer advertising: The effects of disclosure language on advertising recognition, attitudes, and behavioral intent', *Computers in Human Behavior*, 76, pp. 96–106. Available at: https://www.researchgate.net/publication/319167174 (Accessed: July 2025).