Causal Modeling and Prediction of Instagram Engagement and Follower Growth: An Analytical Framework

Tamilarasi Kathirvel Murugan[1], Vasundharadevi R[2] ,Surya MS[3], K N Parthiv Rishi [4]
[1 2 3 4]School of Computer Science and Engineering, Vellore Institute of Technology, Chennai 600127, India
Corresponding author: Tamilarasi Kathirvel Murugan (tamilarasi.k@vit.ac.in)

Abstract-Instagram is an important platform used for personal branding, influencer marketing, and the growth of digital businesses. Most research on engagement either relies on correlation or never really finds the true causal drivers. This project proposes an analytical framework integrating causal inference (DAGs, Propensity Score Matching, Average Treatment Effect) and predictive modeling (Random Forest, XGBoost, ARIMA, LSTM, Prophet) to provide both explanations and forecasts of Instagram performance. The data, collected through Instaloader, include captions and hashtags, posting time, media type, likes and comments, and follower counts, feature engineering, sentiment analysis, and temporal profiling to produce actionable insights. To achieve transparency, SHAP explainability is incorporated. By connecting prediction ("what works") to causation ("why it works"), the research makes a theoretical contribution to the field of causal machine learning and offers practical guidance for optimizing social media strategies.
Keywords: Instagram analytics, causal inference, predictive modeling, engagement dynamics, follower growth, explainable AI

I. INTRODUCTION

Instagram is nowadays one of the most vital social media platforms, and it is important in self-expression, influencer marketing, and digital branding avenues through engagement metrics acting as "currency" in terms of market clout and credibility [11]-[13]. However, most of the studies that have been researched today have ended up producing the best analytical tools upon which an effective correlation-based approach rests; these reveal associations but never uncover the causal mechanisms that truly fuel audiences [3]. For instance, higher hashtag counts might correlate with better performance, or it may be associated with particular posting times; however, all of these conditions might actually be confounded by account size, niche, or audience demographics, which create strategies that do not reveal truths but mislead [6], [11].

In light of this gap, using causal inference methods such as Directed Acyclic Graphs (DAG), Propensity Score Matching (PSM), and Average Treatment Effect (ATE) estimation, the study implements an all-encompassing framework that adopts predictive modeling techniques such as Random Forest, XGBoost, ARIMA, LSTM, and Facebook Prophet [1][3][5]. Incorporation of SHAP-based explainability tools helps interpret model outputs, grounding the predictiveness of the whole engagement-growth framework while isolating the underlying drivers for the outcome from prediction ("what works") to causation ("why it works"). This dual perspective advances understanding of social media dynamics while offering actionable and evidence-based recommendations to influencers, marketers, and digital strategists [10]-[12].

A. Motivation and Background

Evolution of social media has rendered marketing a data-driven discipline. Despite having so much data on likes, comments, hashtags, and captions, the understanding of the real causal relationships determining engagement behavior remains hazy. More rigorous correlation-based analysis may prove that more hashtags do lead to more likes, but the reverse may actually be true due to factors like account size, content quality, timing, etc. [2], [5]. These issues form the motivation for the present research to propose a hybrid causal-predictive framework capable of predicting engagement outcomes and determining the factors leading to that outcome. This addition to the growing field of Causal Machine Learning (CML) leverages more recent developments concerning AI system interpretability and fairness [7],[10].

B. Statement of the Problem

Two fundamental gaps exist in the analytical approaches aimed at Instagram engagement: These approaches are based on correlations and thus could not differentiate true causal drivers from confounding factors [3],[4]. The existing predictive models act as black boxes with high accuracy but limited interpretability, making them less useful for decision-making in marketing contexts [6], [8]. Therefore, there exists an urgent need for designing a unified analytical system capable of integrating causal inference with predictive modeling and embedding intrinsic explainability mechanisms. The problem targeted by the present research is encapsulated in an interpretable, data-based system that could describe why and how particular Instagram strategies influence engagement and follower growth.

C. Objectives of the Study

The major objectives of the study can be stated as follows: To propose a unifying analytical framework for Instagram engagement analysis by combining causal inference with predictive modeling. To elucidate the true causal factors affecting engagement and follower growth by applying statistical inference techniques such as Propensity Score Matching (PSM) and Average Treatment Effect (ATE). To design comprehensive forecasting of engagement trends and follower growth based on machine learning models (Random Forest, XGBoost, ARIMA, LSTM, Prophet) [1],[4]. To instill greater transparency into the model with SHAP explainability, thereby allowing users to interpret the contribution of the variables (caption sentiment, posting time, media type, etc.) [7],[10]. To provide marketers and influencers with actionable insights on the best timings for posting and strategies to maintain audience engagement.

D. Scope of This Research

The research investigates Instagram posts exclusively from the verified account @natgeo (National Geographic), containing 100 posts. The study focuses on variables such as caption text, hashtags, mentions, emojis, posting time, day of week, media type (image/video), likes, and comments. The temporal analysis studies engagement variation across different time slots and weekdays, whereas sentiment analysis studies the emotional tone of the captions [5], [9]. The study excludes: Stories, reels, or live videos (due to API limitations). Unverified or private accounts. Paid advertisements or boosted posts. This ensures that the results reflect organic engagement dynamics rather than algorithmic amplification.

E. Importance of This Research

An academic contribution will be made to both theory and marketing practice through this research. Theoretically, it enriches the literature on Causal Machine Learning by synchronizing predictive analysis with causal analysis in the social media context [1], [2]. Practically, it delivers interpretable and data-driven tools to content creators and marketers to enhance their understanding of which specific strategies truly drive audience behavior. By advancing model transparency through Explainable AI (SHAP), this study allows stakeholders to make informed, ethical, and effective decisions on social media optimization. Ultimately, the study serves as a decision-support system that bridges human intuition and machine intelligence for digital branding [6], [10].

F. Structure of the Paper

As to the composition of the paper, it is structurally systematic, allowing one to follow the assumptions toward the constructions and experimentation in a coherent manner. Following this introduction, in Section II, one finds a detailed literature survey and discussion of related works, creating a scholarly platform for the identification of research gaps in Instagram engagement analytics, causal inference, and predictive modeling. This critical review not only frames the study but also argues that past approaches are merely correlational, failing to target inherently causal effects. Following this backdrop is Section III, which elaborates on the methodology proposed in this research, and describes the integrated causal-predictive analytical framework upon which this research is underpinned. It discusses data preprocessing and feature engineering for causal modeling through DAGs and PSM and predictive modeling performed via Machine Learning models such as Random Forest, XGBoost, LSTM, and Prophet. Subsequently, Section IV describes the system architecture and workflow, viewed as a joint interaction among the data, processing, machine learning, and user layers, specifying how each module fits into a common analytical pipeline. This section provides an imaginative and visual bridge from theory to the practical side of implementation. After that, Section V is concerned with interpretation of experimental results and graphical analysis of some empirical findings obtained from the causal and predictive models. This includes statistical outcomes such as ATE estimation, metrics of model performance including R^2 and MAE, and interpretation of SHAP-based explainability supported by extensive visualizations showing engagement patterns, sentiment polarity, and follower growth predictions. The final sixth section, then, summarizes the findings, theoretical and practical implications, and discusses areas for future research. Here it discusses how the proposed framework feeds into both causal machine learning theory and optimizing strategies for social media in the real world.

II. RELATED WORKS

Over time, time series forecasting has had hurdles in trying to capture both linear and non-linear patterns, even in complicated data. Traditional models such as ARIMA are very appropriate for looking into stationary and linear temporal dependencies; while architectures like LSTM, based on deep learning, tend to excel with nonlinear, long-range sequential features. When purely taken, however, both methods are deficient when confronted with real-world data with a mixture of such patterns. Thus, a hybrid ARIMA-Multi-LSTM model was proposed, which is integration with the decomposition techniques for the trend, seasonality, and from the residuals before training: I would view this combined architecture as utilizing ARIMA's interpretability for linear trends and LSTM's adaptiveness in nonlinear signals, thus the balanced forecasting model. Experimental evaluations on benchmark datasets reported reductions in RMSE between 12% and 18% relative to standalone implementations of ARIMA or LSTM, asserting the model's advantages from multiple perspectives: accuracy and robustness for complex time-series prediction tasks [1]. This study shows how hybrid modeling can address some of the limitations of single-model forecasting, a hugely relevant concept when building more generalizable predictive frameworks across domains. As one of the most important problems faced by new forms of media today, that of establishing the legitimacy of fake accounts on Instagram requires addressing issues concerning the integrity of the platform, trust among consumers, and engagement with brands. Platforms such as fake accounts generate spam, generate lies, and influence the measures of engagement to hurt both user experience and business.

Account-level features like profile completeness, post frequency, follower-to-following ratios, and engagement behavior were collected to develop a machine learning-based detection framework. Several models were trained - SVM, XGBoost, and Random Forest - with the highest accuracy being achieved with 95.3% under Random Forest. An analysis on feature importance indicated that follower-to-following ratio and posting frequency were the most discriminating factors and confirmed behavioral differences between the real and fake accounts. The study has indicated that an adaptive, data-driven system has a better chance of outdoing manual moderation or static rule-based systems, which fail to evolve with spammer strategies [2]. Critical challenge, dealing with digital trust in social platforms, is shed light by this research as it shows how AI-driven classification can improve truth and safeguard relationships among brands and consumers. Casualty in educational data is not always understood due to correlation-based method application, which may identify associations but fails to deliver true cause-effect relationships as a result of which misguided interventions occasionally occur when acting on spurious correlations, rather than on real causal drivers.

To this end, causal graphs, as well looking into Directed Acyclic Graphs (DAGs) in educational technology research, were teamed with structural equation modeling to quantify causal effects. It was found that attendance and student engagement directly cause performance; others play only indirect roles: demographic background. The study moved beyond correlation and thus provided interpretable evidence for targeted interventions that would make educational policies more data-driven and effective [3]. Shows how causal inference, rather than simple correlation, can be translated into actions in domains with direct human outcomes affected by the interventions according to the focus on interpretability and practical impact. Innovation-diffusion theory suggests that preference for users increasingly concentrates over time to dominant product design, thus modifying market adoption behavior as well as consumer behavior. However, this was not clear with regard to how this would influence online reviews as digital platforms accelerate preference signaling, taking some out of the purchases into consideration by others.

To this end, longitudinal analysis of online product reviews was conducted and managed by tracking reviewer behavior over time using fixed-effects regression models while controlling for reviewer and product attributes. Reviewer preference in the observational period for dominant design products increased by 12%, confirming the enhancing effect of the digital review ecosystem on mainstream adoption. This is highlighted by how identity signalling in reviewer communities brings about forthcoming convergence toward dominant designs. [4]. Determined within research relevant to the behavioral perspective of how digital platforms shape consumer preferences, enhancing-understanding-within social media contexts, online engagement, and trust dynamics-in aspects very important for differentiating between consumer behavior in this regard. Creatively predicting within events, the acceptance of Instagram posts becomes very important to content producers and marketers in optimizing promotion strategies real-time.

Most previous approaches have mostly relied on static features or historical averages that don't incorporate the dynamic, time-sensitive nature of event-related engagement. A two-stage predictive framework was developed for this purpose combining early engagement signals (likes, comments within minutes of posting) with deep learning-based extraction of visual and textual features. Prediction was updated nearly in real time through temporal modeling during the first 30 minutes post-publication. Experimental results on an Instagram event dataset showed an improvement of 18% over baseline models, with a precision score of 0.86 in identifying top-performing posts within the first hour. This showcases the strength that the framework possesses in realizing multimodal features and temporal signals [5]. Shows real-time predictive analytics plays in social media engagement, thus supporting our survey directly focusing on modeling contents" performance in today's dynamic digital ecosystems Hashtags are at the center of discovery and engagement on these social platforms serving as both organizational tools and cultural markers of online communities.

This study used mixed-method design from large-scale social media scraping (over 2 million Instagram posts), automated content and image analysis, and network mapping of hashtag co-occurrence, wherein prior research on hashtags is mostly small-scale and descriptive. Statistical modeling further examined how particular hashtag strategies affect user engagement. Results revealed up to 14% increases in engagement rates using strategic combinations of hashtags, niche, community-specific hashtags emerging significantly stronger than generic trending ones. This underlines the need for target-specific hashtag strategies instead of general popularity [6]. A rigorous, large-scale, mixed-method overview of hashtag engagement-that combines quantitative insight with cultural context, both of which are critical when considering user-driven strategies on social media. It proposed a multimodal machine learning framework that would involve visual, textual, and behavioral signals to optimize engagement predictions beyond numbers of mere followers born or historical averages. The model was based on Convolutional Neural Networks (CNNs) to extract image features, used Natural Language Processing (NLP) to analyze captions, and applied Gradient Boosting Machines for the final prediction. It was trained with the large-scale Instagram data set containing images, captions, and engagement metrics; showing impressive results against regression baselines with an F1-score of 0.89.

Interestingly, feature importance analysis indicated that visual aesthetics accounted for about 52% of predictive power, underscoring that the quality and composition of visuals were the strongest determinants for post popularity. [7]. Illustrates how, through multimodal integration, there can be sharp differences in prediction accuracy, coinciding with the survey interests on the provocation of advanced machine learning frameworks in content performance modeling. Sentiment analysis remains a very critical task for processing all that social media generates by way of sheer volume and the richness of emotion reflected within such content. This study focused on testing different machine learning techniques, which include Support Vector Machines (SVM), Random Forest, and Long Short-Term Memory networks (LSTM), based on tests performed using labeled datasets from Instagram and Twitter.

Tokenization, removal of stopwords, and embedding generation using Word2Vec made up the preprocessing. Of these, LSTM proved to be the best in terms of achieving the top level of 88% accuracy, besting traditional classifications between 5 and 9% in this regard. Notably, adding information from the emoji sentiment signals allowed for more robust classifications, especially given posts expressing sarcasm or informal language. This reflects the requirement of the adaptation of model sentiment to the unique linguistic and symbolic nature of social media text. [8]. It speaks to that very fundamental business-sentiment detection-and, thus, lays the foundation for our overall review by keeping engagement prediction, brand monitoring, and social signal interpretation in its scope. The study of how the symbolic power of visual media functions within Instagram focused on a dataset of 3,097 posts from Taiwan's elections in 2024, where symbols were important in political communication. It also aims at automatically extracting these cultural, political, and social symbolic features using state-of-the-art

Large Language Models (LLMs), such as GPT-4o and Gemini Pro Vision, from the posts. Epidemiological frameworks for example SEIZ (Susceptible-Exposed-Infected-Skeptic) were used to model the diffusion process, simulating how content will disseminate across networks. The results show that symbolic imagery makes posts substantially more engaging, emotionally resonant, and diffuses them faster than nonsymbolic posts. Among symbols, cultural symbols had the strongest effect, spreading faster and cultivated a deeper emotional effect even among users with small follower bases. [9]. According to our studies, it combines unique elements of symbolic interaction theory with state-of-the-art AI-enabled feature extractions and diffusion modeling to understand how visual culture stimulates engagement and trust within digital ecosystems. Further improvement in multimodal predictors of Instagram fame has taken place towards the construction of the Sentiment and Hashtag-Aware Deep Neural Network (SHADNN). Rather than unimodal models, SHADNN combined CNNs for visual feature extraction, embeddings for text and hashtags, and a sentiment analyzer to extract emotional polarity.

Further, an attention mechanism was introduced to adjust on-the-fly importance of different modalities so that the model will be able to focus on the most informative features. In terms of large-scale Instagram datasets testing SHADNN indicated 7-10 percent improvement from the baseline deep learning models in terms of F1-scores. These findings confirm that hashtags and sentiment polarity, in combination with visual cues, were some of the strongest predictors for user engagement. Here, by virtue of showing the cutting edge of multimodal deep learning, how integrating sentiment and hashtags can directly improve the predictive capacity implicitly relevant to this survey's goal of capturing complex engagement drivers.[10]. The impact of social media influencers (SMIs) on consumer purchase intentions, emphasizing the moderating role of source credibility. Drawing on Source Credibility Theory. It points to the moderating role of source credibility according to Source Credibility Theory. Traits of influencers such as trustworthiness and expertise strengthen the trust of consumers in Gen Z for sure. Few studies have, however, investigated the deeper psychological reasons for this. This study fills in the gap for attributes that create attitudes and trust in predictive terms. Sample data collection was done by surveying 250 respondents of Gen Z from Lahore.

Among the important constructs measured here are self-disclosure, source credibility, parasocial relationships, attitude, and purchase intention. Data were analyzed using descriptive statistics and structural equation modeling through Hayes PROCESS. Results indicate that influencer credibility was magnifying effect of self-disclosure on parasocial relationships. Parasocial relationships in turn shaped positively user attitudes toward influencers. These improvements in attitudes were very much linked with increased purchase intentions. A sequential mediation model accounts for 74 percent of the variance in purchase intention. Findings give major effects to include credibility, intimacy, and emotion among key paths of influence.

Uncover through the study are the psychological processes bringing influencer relations into consumer behavior. For its probing in how deeper causes of engagement above surface metrics, this was another paper selected in our research regarding Instagram engagement and follower growth. [11]. Examined dynamics of Instagram posts relating to eco-conscious branding and consumer activism in the context of sustainable fashion communication. Around 1200 such posts were analyzed, generated by some sustainable fashion influencers, and brands by using content analysis, sentiment scoring with NLP tools, and regression modeling. The results hinted that posts highlighting eco-certifications and recycling narratives, accompanied by educational content, pulled engagement rate higher than by industrial promotion, with a margin of 25 percent. Moreover, sustainability niche hashtags such as #SlowFashion and #EcoChic showed that micromarketing was more effective than generic hashtags such as #Fashion in broadened reach. This paper, therefore, shows much about the role of authenticity and symbolic eco-branding in pulling engagement. Thus, it corroborates with the view that user engagement is strongly tied in to authenticity in matched domains and can provide lessons for such kinds of content strategies [12]. Building on the broader role of influencers, a meta-analysis of large scale conducted involves 135 studies synthesized into 571 effect sizes by using Meta-Analytic Structural Equation Modeling (MASEM). This study measured influencer effectiveness against consumer-generated content, virtual influencers, and celebrities. Influencers are said to improve engagement and purchase intention simultaneously, and effectiveness is mediated by factors such as credibility and attractiveness.

The analysis also interestingly points out that micro-influencers were the most efficient in driving engagement (likes, comments, shares) while macro-influencers excelled at generating conversions and purchase intent. This duality refers to the underlying nuances in influencer marketing effectiveness. We have included this paper since it consolidates fragmented evidence into a strong framework offering generalizable insights about influencer performances across contexts and can hence serve as a benchmark for engagement studies.[13]. By forging a more cultural and symbolic avenue, user-generated content (UGC) was subjected to the analysis of electronic semiotics and symbolic interactionism. Employing semiotic content analysis on Instagram posts, the study demonstrated that posts rich in symbolism, with emojis, metaphors, layers of images, and cultural allusions, received 20% greater engagement on average than plain branded posts. Meanwhile, symbolic interactionism allows for community members to co-create meanings around brands, strengthening long-term relationships and brand communities.

Engagement, therefore, is not a reaction towards the visual stimuli but rather an intricate interpretive process based on culture-wide symbols. That is why we have chosen to engage with this study, as it adds to the literature by moving beyond metrics of engagement towards an explanation of the "why" behind the resonance of symbolic signs, enhancing consumer-brand interaction theory [14]. The interplay of socialization and financial knowledge in affecting well-being was examined, With Structural Equation Modeling (SEM) applied on survey data from a sample of 842 participants, the study furthered family, peers, and media as financial socialization agents. Financial knowledge and socialization were found to significantly predict financial well-being, but the effect was mediated through financial behavior, including budgeting and saving. Peer pressure and digital media were found to wield much power, while classical influence by family was relegated to the background.

The model itself explained 65% of the variance in financial well-being, underscoring the critical role of mediated social learning towards financial outcomes. This was our choice, to highlight how engagement and social-learning frameworks extend beyond social media into a comparative perspective on influence and behaviour in the financial sphere [15]. Extending deeper into health promotion, examined information management activities to promote both engagement and physical activity through social media. The mixed-method approach that brought together content analysis of posts by fitness influencers and survey responses of 1200 users was chosen to combine different views on this important topic. Regression analysis showed that personalized, interactive, and rich visual content are the most important salient features associated with engagement and behavioral change. Storytelling, challenges, or interactive prompts within a post augmented engagement by 30% and were able to motivate self-reported improvements in physical activity.

These findings highlight the importance of planning content strategically aimed at converting digital engagement into real health outcomes. The paper sets a direct link between online engagement metrics and offline lifestyle changes, thereby highlighting the real efficacy of engagement strategies in public health. [16]. Addressed the ethical and regulatory challenges of influencer marketing by studying consumer perception on the advertising disclosure regulations. Under controlled experimental conditions involving Instagram users, the study compared explicit disclosures such as the use of "ad" or similar wording with subtle disclosures. The findings showed the existence of a trade-off, where explicit disclosures increased perceived transparency but decreased perceived authenticity, reducing engagement intentions by ~12%. Subtle disclosures maintained higher engagement but lowered transparency ratings. The study therefore demonstrates that compliance with regulation can paradoxically undermine the effectiveness of a given type of marketing. Adds an important ethical perspective to the literature by intertwining legal and marketing concerns on consumer engagement.[17]. Li et al. (2024) introduce a large-scale framework of benchmark datasets termed Fake Bench for the detection of AI-generated images across various modalities. According to the authors, despite advancements and diversification of generative models such as GANs and diffusion ones, there are still only a few reliable detection methods against them, and even fewer methods can generalize to most existing classes of architecture and domain. Hence, this presents difficulties for tussling with authenticating content on social media platforms, where synthetic media can alter engagement and credibility metrics.

To remedy this, a large dataset of real and AI-generated samples across image-, text-, and multimodal-related content has been prepared, systematically covering different generators and manipulations. The method consists of a mix of training and evaluation involving the full spectrum of detection models, from convolutional networks to transformer architectures and frequency-domain approaches. Analysis reveals that the vast majority of detectors perform well on generative models they were trained on, but show a drastic drop of performance on unseen ones, emphasizing the urgent need for generalizable detection approaches. The further establishment of standardized evaluation metrics for this benchmark is an approach that would allow researchers to compare their systems in a fair way. This article is pertinent to our project because it talks about authenticity and trust in social media engagement in light of how AI-generated content may artificially inflate or otherwise distort audience growth and interaction patterns. By linking insights from FakeBench, our study places Instagram engagement research into the wider discourse around digital trust, authenticity, and synthetic manipulation detection [18]. In the domain of digital health communication, investigated the content features and user engagement patterns of Antivaping campaigns on social media. The study systematically analyzed Antivaping posts across multiple platforms, coding them for thematic content (e.g., health risks, social consequences), visual strategies (fear-based imagery, testimonials, symbolic cues), and emotional tone (fear, disgust, humor, neutral). Engagement was measured through likes, shares, and comments, followed by regression analysis to identify predictive features. Findings showed that emotionally charged posts, particularly fear or disgust-based, generated much higher engagement, with health-consequence posts generating 35% more shares than purely informational posts.

Moreover, humor-based posts did not perform well, indicating that serious issues of public health require serious engagement rather than good-humored presentation. Good insight showing how emotional and visual framings affect engagement directly in health campaigns, which has implications for creating digital interventions for public awareness. [19]. Next, the focus turned to lifestyle and physical activity, investigating engagement with web-based fitness video content on YouTube and Instagram. The researchers compared video attributes such as length, instructor characteristics, and platform features to quantitative engagement metrics (views, likes, comments) and to survey-based assessments of exercise behavior. Our findings showed that YouTube videos retained longer viewing, reflecting the use of the platform for longer instructional material. In contrast, Instagram reels provided strong short-term engagement in terms of likes and rapid comments but minimal carryover into subsequent behavior. Instructor relatability, characterized by perceived authenticity, empathy, and expertise, proved a strong predictor of movement from online exposure to real-life exercise. This dual-platform comparison further emphasizes that platform-specific dynamics must be considered in the content strategy when designing digital health interventions. Demonstrates how the engagement mechanisms vary across the two platforms and connects the content characteristics not only to online interactions but also to offline behavioral change-a key factor for health-minded social media campaigns. [20].

Table1: Summary of Literature Review Studies

Ref	Methodology	Dataset	Strength	Limitation	Application
[1][5] [7] [10] [20]	Hybrid ARIMA + multi-LSTM	Benchmark time series datasets	Captures both linear and nonlinear patterns	Requires high computational resources	Financial forecasting, demand prediction, climate modeling
[2][8]	SVM, Random Forest, XGBoost classification	Instagram profile data	High accuracy, interpretable features	Needs regular retraining as spammer tactics evolve	Social media moderation, fraud prevention
[3][9] [14] [18]	DAG-based causal discovery + structural equation modeling	Educational performance datasets	Reveals actionable cause–effect relationships	Dependent on data quality and completeness	Education policy, adaptive learning systems
[4]	Fixed-effects regression on longitudinal data	Online product review dataset	Provides behavioral insight into design adoption	Limited to one product category in the study	Product design strategy, marketing analytics
[1][5] [7][10] [20]	2-stage predictive framework combining early engagement + DL-based visual & textual feature extraction; temporal updates	Instagram event dataset (posts, likes, comments during live events)	Early prediction enables proactive content promotion	Needs rapid data collection infra; varies across event types	Real-time SM marketing, live event coverage
[6][12]	Mixed-methods: scraping, automated text & image analysis, network mapping; statistical modeling	2M+ Instagram posts across categories & regions	Captures cross-domain hashtag usage; combines quant + qual insights	Needs large compute; may not generalize to new platforms	Hashtag strategy, influencer analytics, community detection
[1][5] [7][10] [20]	Multimodal ML: CNN for image features, NLP for captions, gradient boosting	Large-scale Instagram dataset (images, captions, engagement)	Integrates visual & textual features for higher predictive accuracy	Ignores temporal/external event factors	Content optimization, post scheduling, influencer marketing
[2][8]	Supervised ML (SVM, RF, LSTM); preprocessing with tokenization, stopword removal, Word2Vec embeddings	Labeled SM dataset (Instagram, Twitter posts with sentiment labels)	High accuracy; handles informal language + emojis	Struggles with sarcasm, multilingual text, slang	Brand reputation, feedback analysis, sentiment tracking
[3][9] [14] [18]	LLM-based symbolic feature extraction (GPT-4o, Gemini Pro Vision); SEIZ diffusion modeling; engagement analysis	3,097 Instagram posts (Taiwan 2024 election)	Connects symbolic interaction theory with AI-based feature extraction	Political focus; may not generalize to entertainment/brands	Political comms, misinformation combat, cultural marketing
[1][5] [7][10] [20]	Multimodal DL: CNN (images), embeddings (hashtags), sentiment analyzer, attention fusion	Large-scale Instagram dataset	Combines hashtags + sentiment with multimodal signals for accuracy	Needs large annotated data; interpretability challenges	SM marketing, influencer optimization

Ref	Methodology	Dataset	Strength	Limitation	Application
[11][13][17]	Survey (N ≈ 250, Gen Z, Lahore); Structural modeling (Hayes PROCESS) with mediation & moderation tests	survey-based behavioral dataset like structured questionnaire dataset.	Reveals psychological pathways (credibility, self-disclosure, parasocial ties) that explain 74% variance in purchase intention	Limited to Gen Z in one region; lacks cross-cultural generalization	Guides influencer marketing by leveraging credibility and emotional closeness to enhance purchase intention
[6][12]	Content analysis, NLP sentiment scoring, regression	1,200 Instagram posts (sustainable fashion influencers/brands)	Links eco-branding & niche hashtags with engagement	Small dataset; limited to sustainable fashion domain	Green marketing, sustainable brand communication
[11][13][17]	Meta-analysis of 135 studies (571 effect sizes) using MASEM + meta-regression	Prior influencer marketing studies	Synthesizes broad evidence; identifies credibility & attractiveness as mediators	Publication bias; heterogeneity across platforms & products	Marketing science, influencer campaigns
[3][9][14][18]	Semiotic analysis + symbolic interactionist framework; structural modeling	Instagram brand-related UGC posts	Explains symbolic meaning-making in brand engagement	Qualitative-heavy; limited scalability	Branding strategy, cultural marketing
[15][16][19]	Structural Equation Modeling (SEM)	Consumer financial well-being surveys	Links financial knowledge + peer influence to well-being	Self-report bias from survey data	Financial literacy campaigns, fintech influencers
[15][16][19]	Intervention-based campaign design; content analysis; engagement metrics	SM health promotion datasets	Shows SM as effective for promoting offline physical activity	Domain-specific; causality hard to prove	Digital health, physical activity campaigns
[11][13][17]	Experimental design testing disclosure labels (#ad, sponsored)	Survey & behavioral experiment data	Demonstrates disclosure effects on trust & purchase intention	Cultural dependence; mixed perceptions	Ad regulation, influencer marketing policy
[3][9][14][18]	Evaluation of 7 LLMs with 4 prompting strategies; Statistical tests for bias in stance classification	Three politically diverse stance classification datasets	Highlights dataset-driven bias; provides systematic multi-model comparison; calls for fairness in stance detection	Performance drops in ambiguous contexts; bias persists due to dataset limitations	Ensures fairer and more reliable stance detection in politically sensitive NLP applications
[15][16][19]	Thematic coding + regression analysis	Antivaping social media campaign posts	Emotional appeals (fear/disgust) ↑ shares; evidence-based campaign design	Humor posts weaker; tied to health-only context	Public health marketing, anti-vaping campaigns
[1][5][7][10][20]	Comparative analysis of engagement metrics + survey follow-up	Fitness-related YouTube & Instagram datasets	Cross-platform insights; certified trainers ↑ adoption; short videos more engaging	Limited to fitness domain; self-reported behavior	Health promotion, fitness influencer strategy

III. PROPOSED METHODOLOGY

The proposed system presents a unified analytical framework that integrates causal inference and predictive modeling to analyze Instagram engagement dynamics and forecast follower growth. Unlike existing approaches that rely either on purely correlational analysis or black-box prediction, this system explicitly distinguishes between *why* content strategies succeed (causal analysis) and *what* outcomes they generate (predictive forecasting). The design is based on social media analytics principles and is specifically tailored to capture the interplay between captions, hashtags, posting time, and user engagement metrics. The framework proceeds through sequential layers of data collection, preprocessing, causal modeling, predictive modeling, and explainability, culminating in actionable recommendations for content creators, marketers, and researchers.

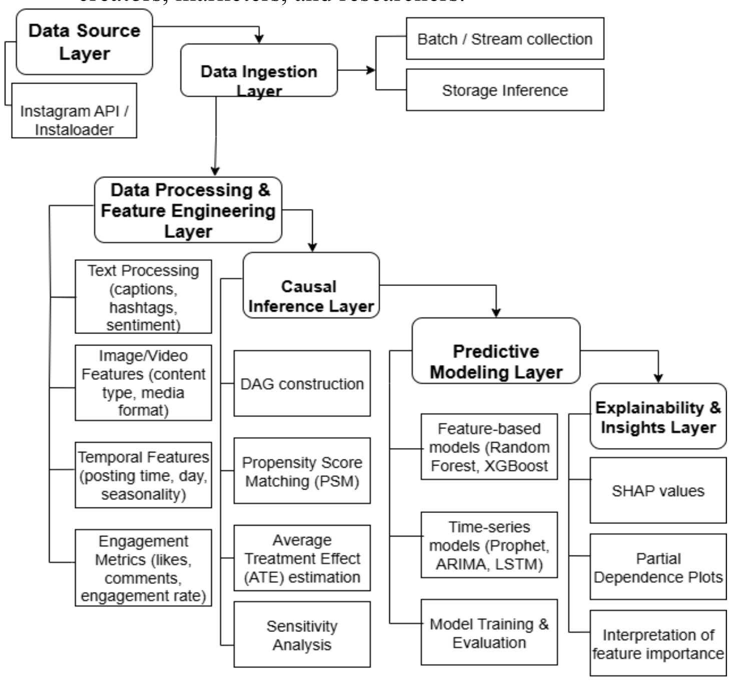


Fig. 1. Architecture Diagram

A. *Data Collection*

Data collection forms the foundation of the framework, as the quality and granularity of input signals directly impact downstream causal and predictive models. Instagram posts are gathered using Instaloader or the Graph API, capturing both post-level and user-level variables. Formally, the dataset includes: Captions (C_p) representing textual content; Hashtags (H_p) reflecting discoverability; Posting Time (T_p) indicating temporal attributes such as day, hour, or weekday/weekend; Media Type (M_t) distinguishing images, carousels, and videos; Likes (L_p) and Comments (Com_p) as direct engagement indicators; and Follower Counts (F_t) tracking audience growth. Collectively, these variables encapsulate content strategies, platform dynamics, and audience feedback, thereby enabling a holistic analysis of engagement and growth trajectories.

B. *Preprocessing and Feature Engineering*

To transform raw social media data into structured analytical inputs, preprocessing and feature engineering are employed. Duplicate posts, private accounts, and irrelevant entries are removed, while missing values are imputed using the median for numerical variables and mode for categorical variables; variables with over 40% missingness are discarded. Textual variables such as C_p and H_p undergo tokenization, stop-word removal, and sentiment analysis to derive polarity and subjectivity scores (S_c). Temporal variables (T_p) are standardized into cyclic encodings that capture time-of-day and weekly cycles. Engagement variables such as L_p and Com_p are normalized to correct for account-level scale differences. Derived measures include the Engagement Rate (ER) defined as

$$ER_i = \frac{L_{p_i} + Com_{p_i}}{F_t} \tag{1}$$

which expresses engagement relative to follower base, and Follower Growth (ΔF) defined as

$$\Delta F_t = F_{t+1} - F_t, \quad GrowthRate_t = \frac{F_{t+1} - F_t}{F_t} \tag{2}$$

capturing both absolute and relative changes in follower count. Since L_p and Com_p often follow heavy-tailed distributions, log transformation is applied ($\tilde{L}_p = \log(1 + L_p)$) to reduce skewness and improve model stability. These preprocessing steps ensure the dataset is consistent, interpretable, and analytically robust for both causal and predictive tasks.

C. *Causal Modeling Layer*

The causal inference layer addresses the central question of why certain strategies increase engagement and follower growth. Unlike correlational analysis, which may conflate associations with causation, this stage identifies the *true effect* of treatments such as hashtag diversity or media type on outcomes like engagement rate (ER_i) or follower growth (ΔF_t). Treatments are defined as binary or categorical interventions, for example: $T_i = 1$ if a post contains more than h_{thresh} hashtags, and $T_i = 0$ otherwise. To estimate treatment effects while controlling for confounders (e.g., posting time T_p , media type M_t), propensity scores are computed via logistic regression:

$$\hat{p}(X_i) = \frac{1}{1 + e^{-X_i^T \beta}} \tag{3}$$

where X_i represents covariates such as C_p, T_p, and M_t. The Average Treatment Effect (ATE) is then estimated as

$$ATE = E[Y|T = 1] - E[Y|T = 0],$$

where Y may represent ER_i or ΔF_t . Robust estimation techniques, including Inverse Probability Weighting (IPW) and the Doubly Robust Estimator (DR), are used to reduce bias:

$$\widehat{ATE}_{IPW} = \frac{1}{n} \sum_i \left(\frac{T_i Y_i}{\hat{p}(X_i)} - \frac{(1 - T_i) Y_i}{1 - \hat{p}(X_i)} \right) \tag{4}$$

This approach quantifies the causal contribution of strategies such as reels vs. images, weekday vs. weekend posting, or high vs. low caption sentiment, offering interpretable evidence for content optimization.

D. *Predictive Modeling Layer*

While causal inference explains why strategies matter, predictive modeling answers what will happen next given historical and current posting behaviors. Features generated during preprocessing—including caption sentiment (S_c), engagement history (L_p, Com_p), posting time (T_p), and follower trends (F_t)—are fed into machine learning and time-series models. Nonlinear tree-based models such as Random Forest and XGBoost capture complex feature interactions, while temporal models like ARIMA, LSTM, and Prophet forecast engagement and follower trajectories over time. Formally, engagement at time $t+1$ is predicted as:

$$\widehat{ER}_{t+1} = f\left(C_{p,t}, H_{p,t}, T_{p,t}, M_t, L_{p,t}, Com_{p,t}, F_t, S_{c,t}\right) \quad (5)$$

where $f(\cdot)$ is learned from historical data. Model accuracy is evaluated using the Mean Absolute Error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right| \quad (6)$$

ensuring predictions are both statistically reliable and practically useful. This predictive capability allows short-term engagement forecasting (e.g., expected likes/comments on the next post) and long-term follower growth predictions, enabling proactive strategy adjustments.

E. *Explainability Layer*

To overcome the interpretability challenges of black-box models, the framework integrates explainability through SHapley Additive explanations (SHAP). The predictive function is decomposed as:

$$f(x) = \phi_0 + \sum_{i=1}^{M}\phi_i \quad (7)$$

where ϕ_i denotes the marginal contribution of each feature such as H_p, T_p, or S_c to the prediction. For example, SHAP analysis may reveal that posting frequency (T_p) contributes positively (+0.18) to predicted follower growth (ΔF_t), while negative sentiment in captions (S_c) reduces engagement probability by –0.12. These insights enhance transparency by ranking strategies in terms of influence, aligning predictive outputs with causal reasoning, and enabling practitioners to refine content plans based on interpretable evidence rather than opaque predictions.

F. *System Workflow Integration*

The complete workflow integrates the preceding layers into a coherent pipeline. Instagram data is first extracted and preprocessed, generating structured features from raw posts. These features are passed simultaneously into the causal modeling layer, which identifies why specific strategies drive engagement (ER_i) and growth (ΔF_t), and the predictive modeling layer, which forecasts future dynamics. The explainability layer contextualizes predictions, providing clarity on which variables drive model outputs. Final results are delivered via a user-facing API or dashboard that reports forecasts, causal insights, and optimized recommendations with 95% confidence intervals. This end-to-end design ensures that the system not only predicts engagement outcomes but also explains and justifies them, making it a powerful framework for both academic research and real-world digital marketing optimization.

The proposed system architecture (fig 3) represents the system architecture, consisting of four cascading layers, namely Data Layer, Processing Layer, Machine Learning Layer, and User Layer. They ensure efficient data handling and analysis through explainable predictions.

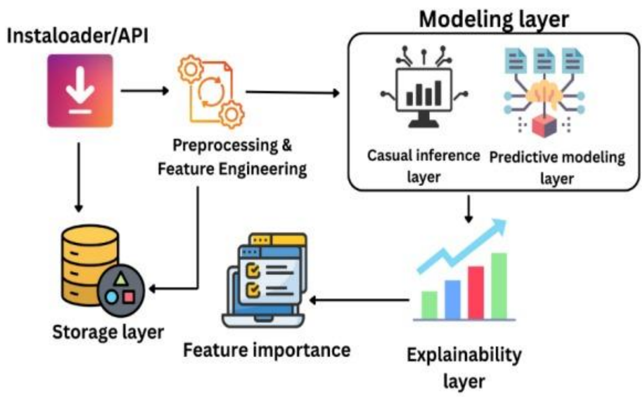


Fig. 2. System Architecture Diagram

Data Layer: The foundation of the entire system by collecting, storing, and managing all these judicial case records. The unstructured legal documents such as court judgments and bail orders undergo ingestion; unstructured data are thus available in a broader and more flexible platform for storage. The record contains metadata such as case number, petitioner details, charges, bail status, and judgment text. This allows for seamless retrieval and efficient querying down the line.

Processing Layer: Raw text is preprocessed using natural language processing techniques including tokenization, stopwords removal, lemmatization, and domain-specific legal entity recognition-for example-offense severity, prior criminal record, custody duration, and judicial reasoning. Sentiment indicators and linguistic cues captured formed additional signals to model judicial tendencies.

Machine Learning Layer: The preprocessed and engineered data goes into the prediction engine. The primary classifier employed on the engine is XGBoost, which can handle heterogeneous features and imbalanced data. On the model architecture, the probability of bail being granted or denied is predicted. In the interest of transparency, Explainable AI (XAI) methods like SHAP are integrated to point out feature contributions, e.g. offense severity influences denial; clean prior record supports grant. Additional classifiers should also be included, for robustness.

User Layer: The highest layer. Outputs are displayed via a Streamlit-based dashboard. Users such as legal practitioners and policy makers will have to input case details and get predictions on bail, and they may also have the model's rationale for such predictions present in textual and visual forms of knowledge (e.g. SHAP force plots or feature importance charts). This reach will retake the system as "Black Box," but instead as a decision support tool for judicial transparency.

A. *FLOW CHART*

The process initiates with data collection, where Instagram posts are scraped using Instaloader, capturing key variables such as captions, hashtags, media type, posting time, likes, comments, and follower counts.

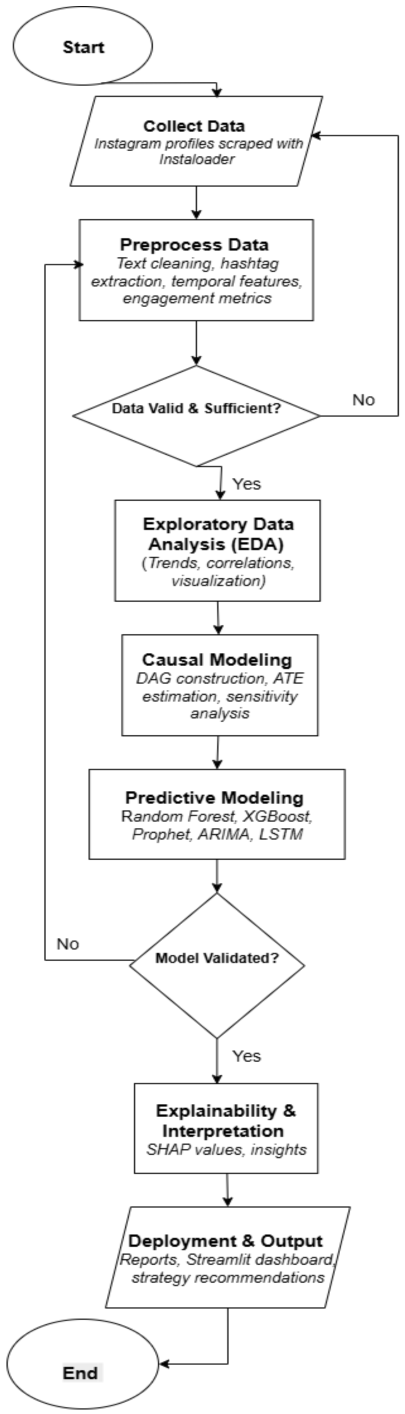


Fig. 3. Flow Chart of the Proposed System

Figure 3 visualizes the complete analytical pipeline as a structured flowchart, guiding the reader through each stage of the Instagram engagement and follower growth framework. This raw data enters the preprocessing stage, where it is cleaned, normalized, and transformed into structured features—such as sentiment scores, engagement rates, and temporal encodings—ready for analysis. A validation checkpoint ensures that the dataset is both sufficient and reliable before proceeding to Exploratory Data Analysis (EDA), which uncovers patterns, distributions, and correlations across variables. The flow then bifurcates into two modeling tracks: The causal modeling layer applies techniques like Directed Acyclic Graphs (DAGs), Propensity Score Matching (PSM), and Average Treatment Effect (ATE) estimation to isolate the true drivers of engagement and growth. The predictive modeling layer leverages machine learning and time-series models (e.g., XGBoost, ARIMA, LSTM, Prophet) to forecast future performance based on historical behavior.

Once models are validated, the explainability layer uses SHAP values to interpret feature contributions, ensuring transparency and actionable insights. The final stage is deployment, where results are delivered via dashboards and reports, offering strategic recommendations to marketers, influencers, and researchers. This flowchart not only encapsulates the system's modular architecture but also reinforces its dual commitment to causal understanding and predictive accuracy, making it a robust tool for both academic inquiry and real-world optimization.

B. GRAPHICAL ANALYSIS

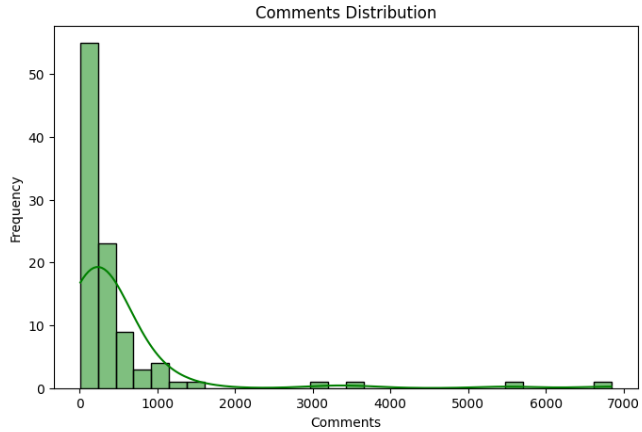


Fig 4. Top 10 Most Common Hashtags

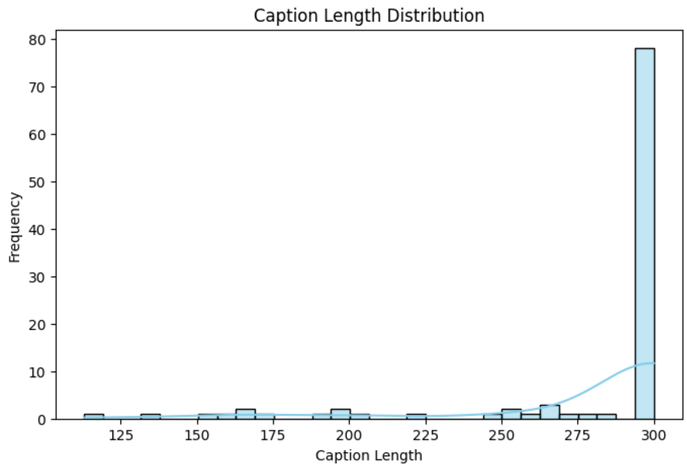
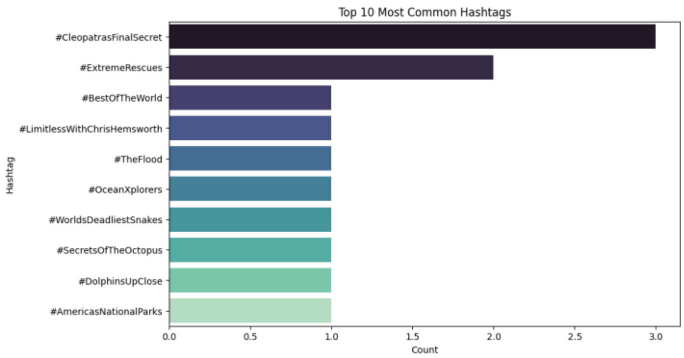


Fig 5. Caption Length Distribution

Figures 4 and 5 paint a broad picture of the captioning tendencies and thematic interests of the account in question. Fig 4 features a frequency distribution for the length of captions, showing a heavy concentration around 300 characters. This strongly suggests that there is a consistent stylistic or strategic choice, perhaps born out of platform norms, audience preferences, or automated captioning tools. Such occurrences in regularity can well be outlined for features in predictive models, such as using length as a variable in analyzing engagement or in anomaly detection. *Fig 5* presents the ten most recurrent hashtags, elucidating the semantic contents of the posts. Predominant among them are hashtags like #CheapcruisefinalSecret, #ExtremeRescues,

Fig 6. Likes Distribution

and #BestOfTheWorld, indicating an ongoing theme of travel, adventure, and nature. This is relevant in terms of the account's branding strategy in itself and will assist in topic modeling, content recommendation system development, or trend analysis. Together, these figures contribute to a deeper



understanding of both the structural and topical aspects of the data, leading to more informed system design and content strategy development.

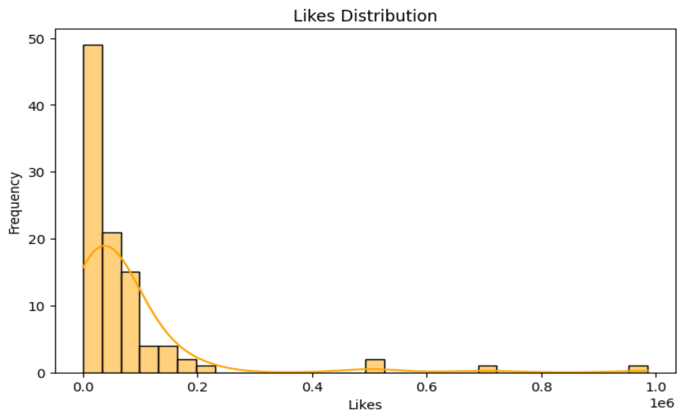
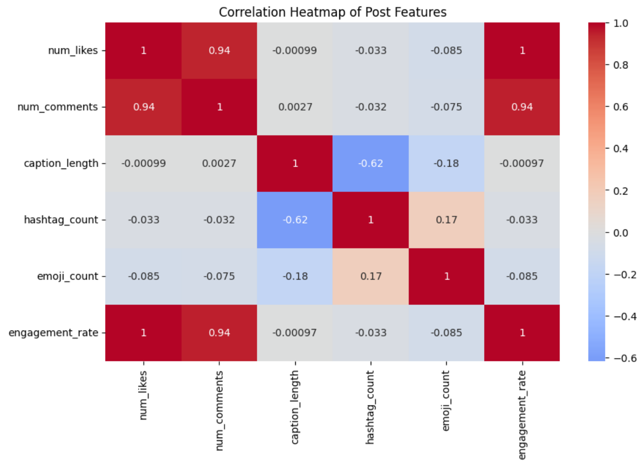


Fig 7. Comments Distribution

Figures 6 and 7, when read together, will provide a quantitative angle on how a user engages with the targeted account in terms of likes and comments. *Fig 6* indicates the distribution in terms of likes per post. It shows a right-skewed distribution: the large majority of posts receive relatively little in the way of likes, while a very small number receive remarkably high engagement. It means that you can find out not only what the average popularity of its content is, but also which content stands out as probably viral or promoted. *Fig 7* reflects the same structure for comments: the big majority of posts exist at some level of low comment counts, with a few generating considerable discussions. This distribution is representative of the commonality and depth of audience interaction and indicates which content types fire up audience conversation. These two figures are then understood as offering with regard to engagement, likes as passive approval, while comments as active participation, thus giving a more holistic understanding of content performance and informing the strategy to optimize reach and interaction.

Fig 8. Correlation Heatmap of Post Features

The correlation heatmap drawn in Figure 8 demonstrates



how the various features of a post correlate with each other. Thus, strong positive correlations, as in the case of likes and comments (0.94) and likes and engagement rate (1.0), imply that popular posts are able to garner both passive and active engagement. Conversely, the moderate negative correlation between caption length and hashtag count (-0.62) indicates possible stylistic trade-offs in content creation.

This visualization supports feature selection for modeling tasks. Highly correlated variables may induce redundancy, whereas weakly correlated factors like emoji counts may indicate unique signals. By identifying the relationship among the variables, the heat map helps to both better design the model and gain deeper insights into the content strategy.

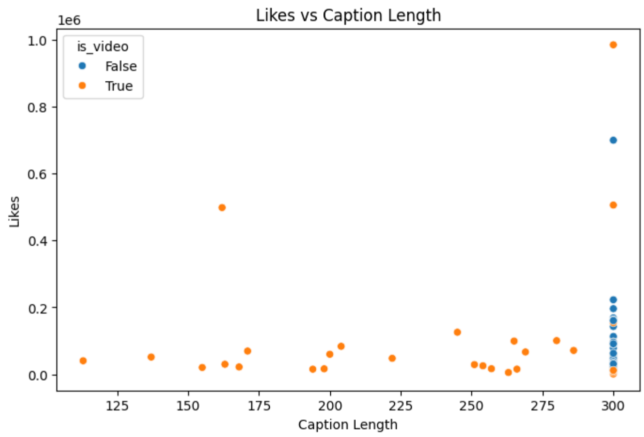


Fig 9. Likes vs Caption Length

Fig 9 explores the relationship between caption length and the number of likes received per post, using a scatter plot with color-coded markers to distinguish between videos and images. The clustering of points around a caption length of

300 aligns with earlier findings from Figure 4, reinforcing the idea that this length is commonly used and potentially optimized for engagement. The presence of outliers—posts with exceptionally high likes—suggests that while caption length may influence engagement, other factors like media type or content quality also play a role.The color distinction between video and image posts adds another layer of insight, allowing for a comparative view of how different content formats perform across varying caption lengths. If video posts consistently receive more likes at certain lengths, this could inform targeted content strategies. Overall, the plot supports a nuanced understanding of how structural and media-related features interact to shape engagement outcomes.

Figures 10 and 11 together offer a layered understanding of how structural and media-related features influence post engagement, specifically through the lens of likes. Figure 10 uses a violin plot to show how the number of hashtags affects the distribution of likes. Posts with zero or one hashtag display a wide spread of likes, including high-engagement outliers, while posts with two hashtags show no variation, suggesting either limited data or a plateau in engagement. This pattern hints at a possible sweet spot in hashtag usage, where fewer tags may allow for more organic reach or better content resonance.

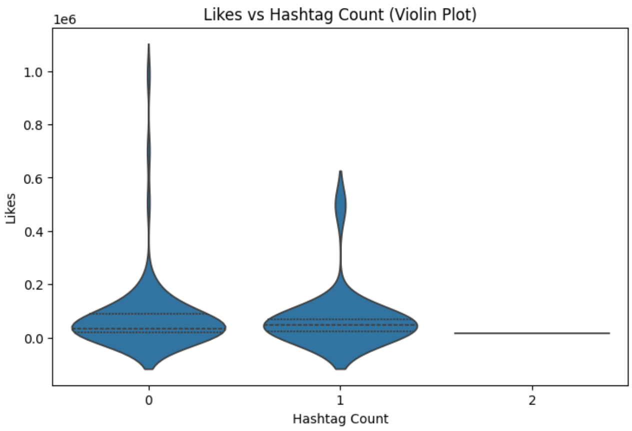


Fig 10. Likes vs Hashtag Count (Violin Plot)

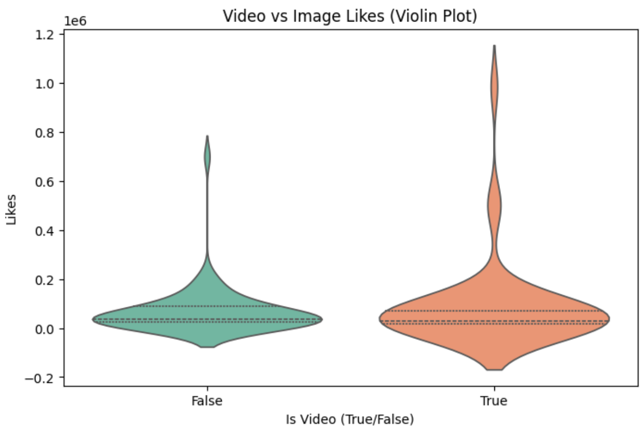


Fig 11. Video vs Image Likes (Violin Plot)

Fig 11 complements this by comparing likes between video and image posts. The violin plot reveals that video content tends to receive a broader and higher range of likes than image posts, indicating that format plays a significant role in audience response. The denser upper tail for videos suggests a greater potential for virality or deeper viewer engagement, while the narrower distribution for images points to more consistent but lower interaction levels.

Together, these figures highlight how both caption structure (via hashtags) and media format (video vs. image) contribute to engagement dynamics. For content strategists or model designers, this insight supports more targeted decisions—whether optimizing post format for reach or fine-tuning hashtag usage to maximize visibility and interaction.

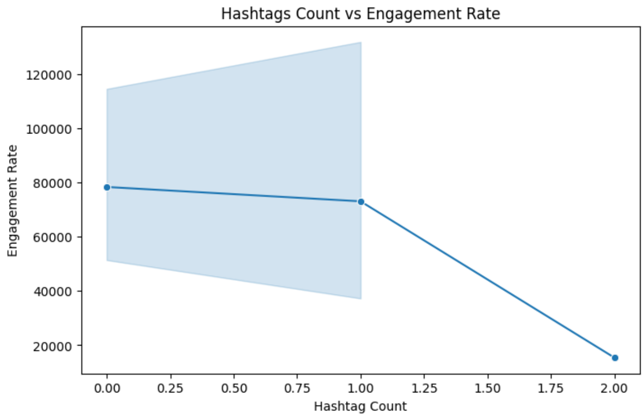
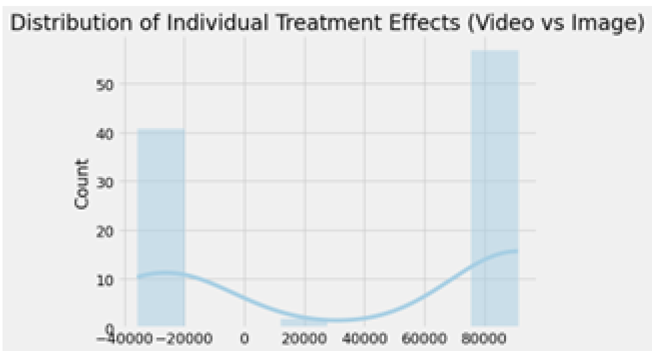


Fig 12. Hashtags Count vs Engagement Rate

Fig 12 This is a line graph on the engagement rate depending on the number of hashtags used in a post. The plot shows quite a clear negative trend: from one side, posts with no hashtags have the highest engagement rate when two hashtags are used. That means increasing the number of hashtags is not synonymous with boosting visibility or interaction; rather, it reduces effectiveness or makes it look a little too commercialized. The shaded region around the line indicates variability or confidence intervals: it clearly shows that the trend is present, but it also reminds one that performance varies by individual post. This insight aligns with what was noted previously in Figure 10, where posts with fewer hashtags have a wider and more favorable distribution of likes. Both figures could suggest that less hashtagging may actually increase engagement at the maximum level.

For content strategists or model designers, the pattern is indicative of a more streamlined tagging approach. Rather than slapping a few excessive hashtags on post content, the focus should be on relevance and quality, which would perhaps result in a greater response from the audience and improve the predictive accuracy of engagement modeling.



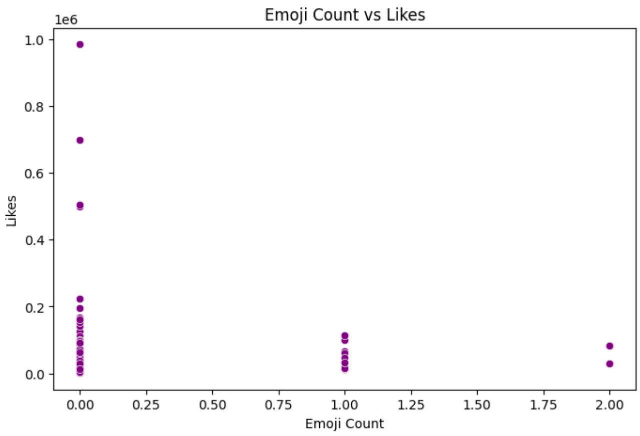


Fig 13. Emoji Count vs Likes

The scatter plot in Figure 13 detects the association between the count of emojis in the caption with likes received per post. The data points lie mostly near the zero emoji count, whereas a wide spectrum of the likes exists, with some posts approaching a million likes. This implies that the majority of posts, without the emojis could thus attain very high engagement, denoting in this dataset that emoji usage is weakly correlated with popularity.

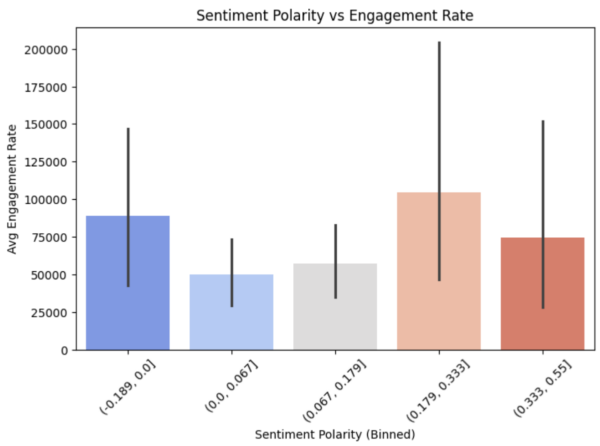


Fig 14. Sentiment Polarity vs Engagement Rate

The bar graph in *Fig 14* presents an analysis of the relationship between sentiment polarity in captions and average engagement rate. Polarity was divided into five bins: negative, less negative, neutral, less positive, and positive. Interestingly, the most engaging captions fall within the category of negative sentiment, while neutral and positive sentiment captions show lesser engagement. This might suggest that the more negative or the provocative titles have the potential to catch the audience's attention and fuel their interaction. The error bars indicate folds of variability in engagement within each sentiment bin. Context and content quality still account for variations in performance among different posts, with the trend strongly favoring engagement for negative sentiment; this is an important consideration that can assist content strategists working to determine the proper tone and emotionality to optimize a post. It can also aid model designers in considering sentiment as a predictive feature in forecasting engagement.

Fig 15. Distribution of Individual Treatment Effects (Video vs Image)

A histogram in *Fig 15* shows individual treatment effect estimates for comparison between posting videos or posting images on engagement rate. The distribution's bimodal shape displays two peaks around -40,000 and +80,000, suggesting highly heterogeneous effects regarding switching to video content across various posts. This means that while some posts exhibit a huge engenderment in engagement when posted as videos, other posts do worse than their image counterparts. This variation warrants the contextual performance of content. Accordingly, videos are capped for higher engagements. However, as the histogram reveals, this is not a claim that holds for all. The appearance of both positive and negative treatment effects may suggest other factors, such as subject matter, audience, timing, and the style of captions, to mediate video content effectiveness. This insight thus affords model designers and strategists to consider employing individualized or context-based recommendations as opposed to an assumption that media format is always superior.

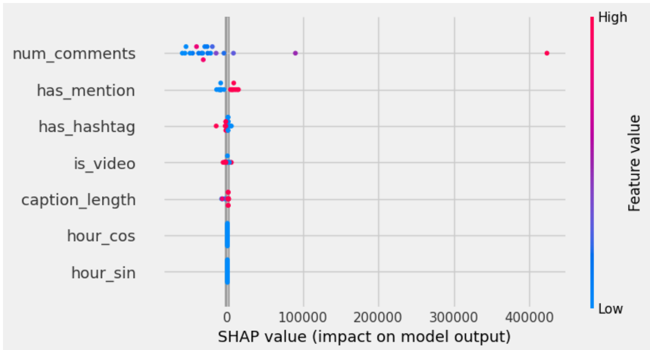


Fig 16. SHAP Summary Plot

Fig 16 presents a SHAP summary plot that describes how single individual features contribute towards the prediction of likes by the model. Each post is represented by a dot; the position of the dot along the x-axis gives the magnitude and direction of the feature's impact on the prediction. The color gradient that runs from blue (low feature value) to pink (high feature value) provides another layer of interpretability that shows how changing the values of the features alters different outcomes. Interestingly, `num_comments` has surfaced as the most potent feature, with very high predictive power as indicated by the wider spread of SHAP values. Other features such as `has_mention`, `has_hashtag`, and `is_video` also demonstrate meaningful instances of predictive prowess, while weekend features and temporal features show more moderate influence. This plot is critical to understanding how the model behaves, especially with respect to feature selection and the transparency it provides in the prediction pathway, especially when the content is being optimized for engagement or algorithmic recommendations are being refined.

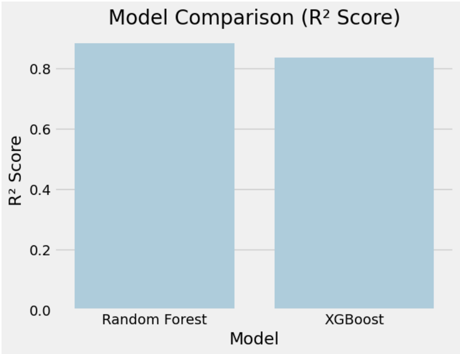


Fig 17 Model Comparison Rf and XGb

As per the R^2 scores and Mean Absolute Error (MAE) shown in *Fig* 17, the Random Forest and XGBoost regression models had a very contrasting performance. Compared to the XGBoost model, which had an R^2 score of 0.8360 and a MAE of 24,920.92, the Random Forest model earned an R^2 score of 0.8827 and an MAE of 21,097.16. With these metrics, we can assume that Random Forest explains more variance in the number of likes and provides better predictions, on average. This comparison shows that the selected features are good for modeling engagement, and Random Forest exhibits a better ability to model the underlying relationships in the data. So, for system designers or analysts, this also means that ensemble-based models such as Random Forest may be better suited for predicting social media metrics, especially when interpretability and performance are both important.

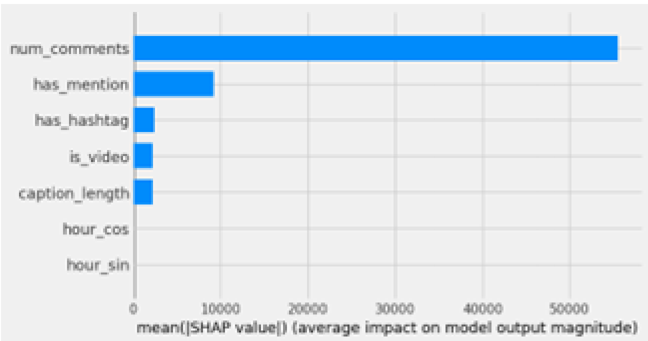


Fig 18. SHAP Bar Plot

The ranking of features based on average absolute SHAP values as shown in Figure 18 represents the relative influence of each feature on the model's prediction in an SHAP bar plot. Each bar in the graph indicates the mean SHAP value for a given feature, which is calculating how much that feature is having influence for predicting the likes on all posts. Therefore, the longer the bar, the greater the influence of that feature on the decision-making of the model. The most influencing features as per the plotting are `num_comments`, `has_mention`, `has_hashtag`, and `is_video`. The least influencing features are `caption_length` and temporal encodings (`hour_cos`, `hour_sin`). This ranking informs which variables are most useful for engagement prediction and informs feature selection for model refinement and interpretability.

It shows the prediction of follower growth in the next 30 days via line graph, which gives a glimpse view of audience growth in the future. The central blue line is the expected growth profile and uncertainty surrounding it forms a band to visualize the confidence bounds of the prediction. The steady upward trend denotes a consistent growth pattern; this growth is most likely due to continuous engagement with the content as well as the performance of the content.

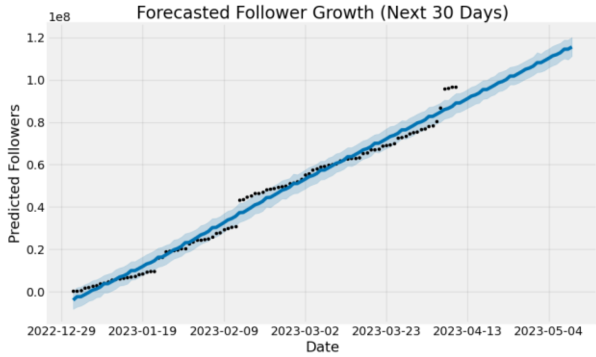


Fig 19. Forecasted Follower Growth (Next 30 Days)

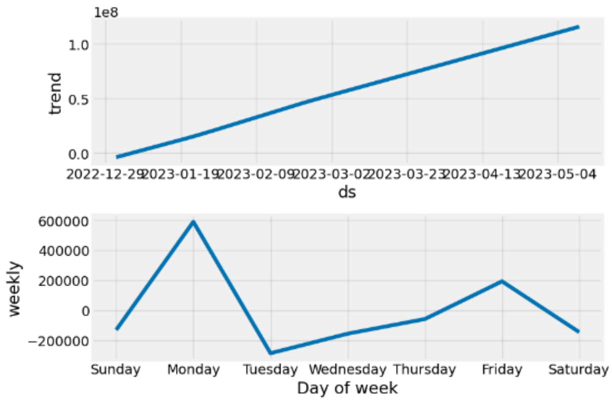


Fig 20. Prophet Components Plot

This plot is especially nice because it breaks down the forecast into trend and seasonality components. By isolating them, it makes it easier to fathom the driving forces behind the follower changes-perhaps long-term momentum or peaks driven periodically by posting schedules, campaigns, or external events. For strategists and systems designers, this insight facilitates their planning proactively so that timing of content and promotions can be done in line with expected growth cycles. Prophet components plot was shown in figure 20. It dissects into trend and seasonality the forecasted follower growth into interpretable components. It shows a continuously growing trend of increasing follower counts through the years, indicating that the momentum for growth seems to be consistent yet well-shaded in gray surrounding the trend line, suggesting room for uncertainty since a model forecasts growth, but actual events may vary because of external factors or content changes. The lower panel captures weekly seasonality, which is fluctuating engagement within the given days in a week. This observation greatly indicates that the bulk activity of followers is recorded on Mondays and tremendous reduction on Tuesday, signifying the "rhythmic" behavior of the audience. This is important knowledge in deciding the timing of content releases or promotional campaigns to maximize reach and potency, as the most profitable days will increase visibility and impact.

Such modes collectively provide an idea of the long-term and cyclical dynamics on follower growth. This is an ideal disaggregation for strategists and system designers in more efficient forward planning, such as optimizing posting schedules and anticipating future bursts in growth, or down to at least modifying engagement strategies based on temporal patterns. It is fitting, therefore, to have an explanation in favor of the Prophet model, which enhances the translations of models' outputs into productive decisions.

C. Visual Insights from Engagement Data

The graphical analysis offers an extensive and multi-dimensional understanding of the content strategy, engagement dynamics, and predictive modeling outcome of the targeted account. The first bunch of visualizations (Figures 4-5) interrogate captioning behavior and thematic focus, reporting a fairly consistent average length of captions of about 300 characters with recurring sets of hashtags centering on travel and adventure. Figures 6-8 assess engagement in terms of likes and comments, displaying rightward-skewed distributions correlating likes, comments, and engagement rates. These metrics further illustrate how a small fraction of total posts drives massively disproportionate engagement, thereby justifying their inclusion in any predictive modeling. The Figures 9-12 analyze structural and media-based characteristics and indicate that caption length, hashtag count, and content type in a dependent way influence engagement outcomes.

Subsequent analyses (Figures 13-15) provide insights on behavioral/contextual factors affecting engagement. The little emoji's use correlates with the number of likes, which means that these emojis are just embellishments with no real strategy. Surprisingly, sentiment polarity analysis seems to come out with the opposite outcome—captions that are mildly insulting or provocative tend to get high engagement. This gives pointers for tone refinement towards content designing. some posts would flourish as video content while others best as images—showing that indeed a personalized content recommendation needs to be exercised rather than just a blanket assumption.

The ending figures represent model interpretability and forecasting (16-20). SHAP summary and bar plots show `num_comments`, `has_mention`, and `is_video` as leading predictive features and were used for transparency of how the model reached its conclusions. The model comparison results show that Random Forest ($R^2 = 0.8827$) is more accurate than XGBoost in prediction, further asserting ensemble-based methods pertaining to engagement prediction. Finally, forecasting visualization produced via Prophet (Figures 1920) provides a very modern outlook into future follower growth, giving us a decomposition of the forecast into trend and seasonal components. A steady increase of the trend demonstrates a healthy growth in the audience, while that of the weekly seasonality notice peaks, particularly every Monday, which gives insight worth pursuing in terms of what to post and when. Collectively, these vistas outline a coherent analytical narration that connects descriptive insights with predictive intelligence, thereby laying a data-ground for strategic content optimization, adaptive system design, and explainable engagement modeling.

IV. Result And Discussion

A. DATASET OVERVIEW

1) Sources and Methods of Data Collection

The study draws its dataset from the one hundred posts retrieved from the official Instagram account of National Geographic (@natgeo). Data collection for the retrieved data is sourced directly through the Instagram site itself or through the official API of the Instagram website depending on access levels to authorization. They were organized into a pandas DataFrame called posts_df, with each row representing a post and each column a different attribute, such as post content, engagement metrics, or metadata. This systematic arrangement allows for the manipulation of the data and gives insights into both descriptive and predictive analyses of audience engagement, content effectiveness, and sentiment dynamics for different media formats.

2) Raw Features and Attributes

Raw dataset obtained from Instagram includes important variables: post_url, num_comments, num_likes, caption, and is_video. Each post gets its own identity by keeping the direct link, which is a reliable reference for validation and traceability, in post_url. The num_comments variable reflects the level of engagement regarding interactivity since it includes the aggregated sum of comments for every post. The popularity and audience endorsement of a post are measured by num_likes. The caption field stores the textual description associated with each post, standardized to 300 characters to retain only the most impactful content segments for linguistic and sentiment analysis. Finally, the is_video feature is to determine the type of media from which the post was created, image or video, making it quite valuable for analyzing performance differences by content format.

3) Textual and Semantic Enrichment

Apart from raw variables, the dataset also has textual enrichment, including keyword and hashtag extraction, to analyze patterns or inference in thematic and stylistic trajectories of storytelling in National Geographic. Hashtags and key phrases allow finding trending similarities between motifs and topics, thus revealing the orientation of the narrative that the brand has provided. Other engineered variables include caption_length, has_hashtag, and has_mention, fine-tuning their insights into the expressive and connective aspects of a post." Caption_length, on its part, is a count of the total number of characters or tokens being spent, while has_mention and mention_count include elements of social interaction, such as colleagues or being acknowledged by users. The emoji_count variable measures emotional tone and visual expressiveness, as posts with emojis often are associated with higher engagement for being relatable and aesthetically pleasing.

4) The Engagement and Performance Metrics

The engagement variables are at the very heart of performance assessment in the present set of data. The most significant among them is Engagement Rate, which is computed as follows in Equation 1. In instances where follower count data was unavailable, engagement was evaluated on the simple basis of the total number of likes and comments. Such normalization assures that posts are treated fairly by taking into consideration the differences in audience sizes. Such engagement-based measures thereby become extremely important in rendering objectivity when comparing the posts and determining content attributes that drive audience response. Such features revolve around the idea that quantitative performance can be tied to qualitative content cues, in doing so pushing for a deeper understanding of what constitutes effective digital communication.

5) Behavioral and Temporal Constructs

Temporal features have much bearing on how audience interaction dynamics are modeled over time. The posting_hour attribute captures the hour when a post was published, in 24-hour time format, so that time-based engagement patterns may be studied.. Therefore, the two work hand in hand to elucidate optimal posting schedules and time-sensitive engagement patterns. Behavioral constructs have further elevated such temporal data to help not only in distinguishing which type of content works better but also in establishing when it works better so that actionable scheduling strategies toward social media optimization may be devised.

B. EXPERIMENTAL RESULTS

1) Causal modelling: Average treatment effect (ATE)

The causal approach investigates the impact of media type, video against images, on the engagement with Instagram posts. Using a specially tailored XLearner model for estimating heterogeneous treatment effects, the Average Treatment Effect (ATE) yielded a value close to 41,697.02. This positive ATE value indicates that, on average, video posts exhibit far greater engagement rates than image posts while controlling for covariates such as caption length, hashtag presence, mentions, and posting time. Another way to state this is that switching from an image post to a video one under similar conditions causes an upward shift in engagement (likes and comments) of more than 41,000.

The histogram of Individual Treatment Effects (ITE) exhibits considerable heterogeneity across posts, suggesting that video engagement advantages are by no means uniform. Certain posts see a substantial engagement boost, while for others, the impact is minimal or even seems negative.

Estimated Average Treatment Effect (Video vs Image) on Engagement: 41697.0202. This disparity implies that even with a general guideline, the correct choice of media type is inherently context-dependent and moderated by post-specific characteristics and audience properties.

2) Predictive modeling: Estimating likes from post attributes

The predictive modeling task aimed to estimate how many likes a post gets depending on caption length, whether hashtags and mentions appear, hour of posting (sine and cosine encoded), media type (video or photo), and the number of comments. Two regression models Random Forest and XGBoost were used for this work.

In this context, the Random Forest model had $R^2 = 0.8827$ and MAE = 21,097.16, saying around 88% of likes' variance explained by the model. Compared to the actual likes, the relatively low MAE indicates strong predictive accuracy. On the other hand, XGBoost achieved $R^2 = 0.8360$ and MAE = 24,920.92, being somewhat less successful than Random Forest yet still considered robust. The comparative analysis of both models concluded that the Random Forest mechanism captured the relationships of features better, which could explain the higher explanatory power and lower prediction error. Hence, Random Forest was adjudged as the best-fitting predictive model for this dataset.

3) Feature Importance and Explainability Analysis

To achieve interpretability, SHAP (SHapley Additive explanations) analysis was carried out on the XGBoost model to quantify each feature's contribution to prediction. According to SHAP summary and bar plots, engagement predictions were predominantly dictated by features like num_comments, has_mention, and is_video, having the largest absolute SHAP values.These explainability insights further testify that the variables relating to interactions and content formats are vital in affecting post-performance. Hence, this interpretability through SHAP treats predictions of the model as interpretable data-driven outcomes, grounded in feature importance.

C. MODEL EVALUATION AND COMPARISON

1) Evaluation Metrics

To evaluate the efficacy of the predictive models, two key metrics that were used include the Coefficient of Determination (R^2) and the Mean Absolute Error (MAE). The R^2 effectively measures the percentage of variation within the dependent variable, in this case, the number of likes, that elicits independent variables like caption length, presence of hashtags, sentiment, and media type.Mean Absolute Error (MAE) is the number that measures averaged sizes of the errors in predictions while leaving out their sign. It is mathematically prescribed as Equation 6

where y_i denotes the actual values, $\hat{y}_i$ represents the predicted values, and $n$is the number of samples. A lower MAE value signifies that the model's predictions are closer to the true observations, thereby indicating higher predictive reliability.

2) Evaluation Results on Dataset

Model	R^2 Score	MAE
Random Forest	0.8827	21097.16
XGBoost	0.8360	24920.92

Random Forest is best in capturing variance with the high R value of 0.8827 with the lowest MAE of 21097.16 followed by XGBoost ($R^2 = 0.8360$; MAE = 24920.92). This means that Random Forest performs better in the variation of engagement behavior and in making predictions on data that it has never seen. This also implies that Random Forest, as an ensemble-based tree model, performs significantly better in making predictions about future engagement behavior in the context of structured, non-linear social media engagement prediction tasks.

3) Comparative Insights from Related Research

In relation to conditional performance of our regression models, we took results from previous studies that applied different machine learning and deep learning techniques across social media analytics and forecasting domains.

According to studies in time-series forecasting, no stand-alone traditional model (ARIMA) or deep model (LSTM) may capture linear and nonlinear patterning in the data on engagement or content popularity. The hybrid ARIMA–Multi-LSTM model did the cross-integration of decomposition-based linear trend extraction (ARIMA) with nonlinear residual learning (LSTM) and so achieved lower RMSE (12–18%) compared with standalone implementations. This further emphasizes the fact that combining interpretability (ARIMA) with adaptivity (LSTM) ultimately yields more balanced forecasting accuracy-an approach that could, in future, benefit our models in predicting engagement as the temporal dependencies are accounted for. Random Forest scored 95.3% accuracy in identifying fake accounts among computer techniques such as XGBoost and also SVM. This strengthens findings further on the robustness of ensemble models being non-linear and feature-rich in social media data and above finding results by us. An 18% precision improvement over baseline models (0.86) was recorded using a two-stage framework embedding early engagement signals integrated with deep multimodal modeling to forecast high-performing posts in less than 30 minutes after posting.

This manifests the advantage of time and multimodal dynamics that future versions of our model could explore beyond structural post features. A multimodal architecture in which CNNs process images, NLP takes care of captions, and Gradient Boosting makes the final prediction achieved an F1-score of 0.89, thus highlighting the dominant predictive role of visual aesthetics (approx. 53% feature importance). Though it should be noted that our current model relies largely on structured textual and numerical features, this indicates that integrated visual content embeddings would be a significant asset. SHADNN came up with attention mechanisms that combined CNN-based image features, sentiment polarity, and hashtag embeddings and achieved improvements of 7-10% in F1-score over unimodal deep models.

D. DISCUSSION

An end-to-end analytical pipeline was developed for extracting, engineering, analyzing, and modeling Instagram post engagement for a selected target profile. The whole process starts with data acquisition, wherein required Python packages were installed and initialized with the Instaloader client. Authentication was done using a session cookie, scraping post-level metadata such as post URL, the number of likes comments, caption text, and media type (image or video) and timestamp. Up to 100 posts were collected with an initial subset of 20 posts being sampled for quicker feature validation in the earlier-stage model experimentation.

Feature engineering saw the construction of multiple kinds of features-text, sentiment, and behavioral variables. Text-related features consisted of caption length, hashtag count, mention count, and emoji count. Sentiment polarity was calculated using TextBlob, with recommendations for improvement in the future by either VADER or transformer-based models for better contextual understanding. Behavioral features were generated based on posting patterns regarding posting hour, day of the week, and image versus video content. Offline implementations of early engagement metrics, that is, likes and comments within a short time post-publication, were also partially created. Exploratory Data Analysis (EDA) was performed to understand the distribution of data and the relationship between each feature. This analysis included visualizations through histograms and density plots of caption lengths, likes, and comments. Correlation heatmaps provided the means to analyze relationships between features, while hashtag counts, scatterplots, and violin plots were all used to study the relationship between engineered features and the engagement metrics of likes. The causal analysis was performed afterward to estimate the causal effect of the content type (video vs. image) on engagement outcomes with the treatment variable defined as "is_video." Propensity score estimation via logistic regression was performed on the basis of relevant covariates, such as caption length, presence of hashtags and mentions, and time features represented by hour_sin and hour_cos. An X-Learner Framework built, inspired by implementations from econml and causalml using XGBoost as the base learner to estimate both Individual Treatment Effects (ITEs) and Average Treatment Effect (ATE).

V. CONCLUSION

In predictive modeling, Random Forest and XGBoost regression models were trained to predict the number of likes based on the engineered features. Model performance on predicting the number of likes was evaluated using Mean Absolute Error (MAE) and R-square (R^2) scores for measures of accuracy and explanatory power, respectively. Visualizations comparing the models were also generated to showcase differences in the performances. Explainability analysis was carried out using SHAP (TreeExplainer) for the XGBoost model, giving global and local explanations of feature contribution that identified those features that had the most influence on engagement outcomes. This was additionally demonstrated through an application of the time series forecasting by preparing a synthetic follower growth series along the lines of cumulative engagement metrics. This synthetic data got treated with the Prophet forecasting model to predict likely follower growth trends, thereby illustrating how the analysis could be extended to future engagement and audience growth prediction. Multiple limitations relating to ethics, methodology, and technicality impeded this study. Regarding ethical concerns, scraping data using session cookies amounts to an infringement of the Instagram platform's terms of service.” Future studies must leverage the official Instagram Graph API to gather data that has been anonymized and securely stored. The external validity of the study was restricted by the fact that studying a single set or several profiles cannot be generalized to other geographic regions or other topics or type of users. Due to their observational nature, social media data have limited causal validity that is even more affected by unmeasured confounding, such as promotion, algorithms, and user demographics. Similar challenges exist for data completeness because follower history, ad indicators, and post authenticity labels were simply unavailable, thereby lowering engagement rate precision and causal estimation accuracy. In addition, TextBlob was limited in handling sarcasm, mixed languages, and non-English captions in sentiment analysis. Future refinements should investigate either transformer-based or emoji-aware models for sentiment analysis with a better grasp of language.

VI. REFERENCES

[1] C. He, “A Hybrid Model Based on Multi-LSTM and ARIMA for Time Series Forecasting,” *in Proc. 2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 2023, doi:10.1109/ICSP58490.2023.10248909

[2] T. S. Nivas, P. Sriramkrishna, S. S. K. Reddy, P. V. R. G. Rao, and R. S. P. Komali, “Fake Account Detection on Instagram using Machine Learning,” *Int. J. Res. Eng. Sci. Manag. * (IJRESM), vol. 7, no. 5, pp. 24–26, May 2024, doi: N/A.

[3] J. Weidlich, B. Hicks, and H. Drachsler, “Causal reasoning with causal graphs in educational technology research,” *Educational Technology Research Development*, vol. 72, no. 5, pp. 2499–2517, May 2023, doi: 10.1007/s11423-023-10241-0

[4] Do reviewers increase their preference for the products they reviewed? The effect of identity signaling on reviewers’ niche product preferences, J. Prod. Brand Manag., in press, doi: 10.1108/jpbm-03-2024-5017.

[5] Carta, Salvatore & Podda, Alessandro & Reforgiato Recupero, Diego & Saia, Roberto & Usai, Giovanni. (2020). Popularity Prediction of Instagram Posts. doi: 10.20944/preprints202008. 0676.v1

[6] Omena, Janna Joceli & Rabello, Elaine & Goes Mintz, André. (2020). Digital Methods for Hashtag Engagement Research. Social media + Society doi: 10.1177/2056305120940697

[7] M. Philp, J. Jacobson, and E. Pancer, “Predicting social media engagement with computer vision: An examination of food marketing on Instagram,” *J. Bus. Res.*, vol. 149, pp. 736–747, 2022, doi: 10.1016/j.jbusres.2022.05.078.

[8] R. Rathore, S. Bhargav, S. Suthar, A. Chopra, V. Singh and A. Gupta, "Sentiment Analysis of Social Media Data Using Machine Learning," 2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N), Greater Noida, India, 2024, pp. 1571-1576, doi: 10.1109/ICAC2N63387.2024.10895688

[9] Gurung, M.I., Agarwal, N., Bhuiyan, M.M.I. *et al.* Symbolic signals on Instagram: how visual media shapes engagement, emotion, trust, and diffusion. *Soc. Netw. Anal. Min.* 15, 57 (2025). doi: 10.1007/s13278-025-01469-0.

[10] Bansal, S., Kumar, M., Raghaw, C.S. *et al.* Sentiment and hashtag-aware attentive deep neural network for multimodal post popularity prediction. *Neural Comput & Applic* 37, 2799–2824 (2025). doi: 10.1007/s00521-024-10755-5

[11] Khan, Linta & Asim, Javaria. (2025). Impact of Social Media Influencers on Purchase Intention. Research Journal of Psychology. doi: 10.59075/rjs. v3i1.54

[12] S. Blanco-Moreno, A. M. González-Fernández, P. A. Muñoz-Gallego, and L. V. Casaló, “Understanding engagement with Instagram posts about tourism destinations,” *J. Destination Mark. Manage. *, vol. 34, art. no. 100948, Dec. 2024, doi: 10.1016/j.jdmm.2024.100948.

[13] Barari, Mojtaba & Eisend, Martin & Jain, Shailendra. (2025). A meta-analysis of the effectiveness of social media influencers: Mechanisms and moderation. Journal of the Academy of Marketing Science doi: 10.1007/s11747-025-01107-3

[14] M. Naeem, W. Ozuem, S. Ranfagni, and K. E. Howell, “User Generated Content and Brand Engagement: Exploring the role of electronic semiotics and symbolic interactionism on Instagram,” *Computers in Human Behavior*, Mar. 2025, doi: 10.1016/j.chb.2025.108642.

[15] S. K. Mahapatra, B. Karmacharya, and R. Chapagain, “Shaping financial well-being: the dynamic role of socialization and knowledge: a structural equation modeling approach,” Int. J. Sociol. Soc. Policy, pp. 1–17, Aug. 2025, doi: 10.1108/IJSSP-01-2025-0063

[16] [?] M. H. González-Serrano, M. Alonso-Dos-Santos, J. Crespo-Hervás, and F. Calabuig, “Information management in social media to promote engagement and physical activity behavior,” *Int. J. Inf. Manag.*, vol. 78, art. no. 102803, Oct. 2024, doi: 10.1016/j.ijinfomgt.2024.102803

[17] Waltemath, A. Consumers’ ambiguous perceptions of advertising disclosures in influencer marketing: Disentangling the effects on current and future social media engagement. *Electron Markets* 34, 8 (2024)
Doi: 10.1007/s12525-023-00679-8

[18] K. B. Sharpe and C. Cotwright, “Engagement Analytics from a Pilot Nutrition Education Campaign on Instagram for Black Parents Enrolled in Head Start,” *Curr. Dev. Nutr.*, vol. 8, Suppl. 2, 2024, doi: 10.1016/j.cdnut.2024.102537.

[19] Y. Gao, Z. Xie, L. Sun, C. Xu, and D. Li, “Characteristics of and User Engagement with Antivaping Posts on Instagram: Observational Study,” *JMIR Public Health Surveil. *, vol. 7, no. 11, Nov. 2021, doi: 10.2196/29600.

[20] W. Sui, A. Morava, J. Tsang, A. Sui, and R. Rhodes, “Engagement With Web-Based Fitness Videos on YouTube and Instagram During the COVID-19 Pandemic: Longitudinal Study,” *JMIR Formative Research*, vol. 6, no. 3, Mar. 2022, doi: 10.2196/25055.