Quest Journals
Journal of Research in Humanities and Social Science
Volume 13 ~ Issue 8 (August 2025) pp: 151-162
ISSN(Online):2321-9467
www.questjournals.org

Research Paper

Investigating Youth Purchase Intentions for Viral Products: A Machine Learning Approach Based on Social Media Influence

Evan Mei, Tianfeng Li
Correspondance Author: Evan Mei
Mountain Lakes High School, NJ

Abstract
Social media has significantly reshaped marketing strategies and consumer behavior over the past decade.This study investigates how social media engagement, advertising strategies, and economic behaviors impact the purchase intentions of students. By employing advanced machine learning techniques and integrating multiple models, the research analyzes these factors' influence on purchasing decisions. It tested eight machine learning models: Logistic Regression , Support Vector Classifier, Gaussian Naive Bayes, Random Forest, Extreme Gradient Boosting, Extra Trees, Gradient Boosting, and Adaptive Boosting..The Random Forest model demonstrated superior performance, achieving an accuracy of 71.2% and an ROC AUC of 79.1%. Feature selection and SHAP analysis were employed to identify critical drivers of purchase intentions. Findings reveal that social media engagement, advertising strategies, and economic behaviors significantly influence purchase decisions. The SHAP analysis highlights that the number of social media platforms used inversely affects viral product purchase intentions, challenging the notion that broader engagement necessarily leads to higher purchasing behavior.

*Received 12 Aug., 2025; **R**evised 22 Aug., 2025; **A**ccepted 24 Aug., 2025 © The author(s) 2025.*
P**ublished with open access at **www.questjournas.org

I. INTRODUCTION

The last decade has witnessed a transformative shift in marketing and consumer behavior due to the rise of social media. Platforms like Instagram, YouTube, TikTok, and others have not only changed how we communicate but have also become pivotal in shaping consumer behavior, particularly among younger audiences [1]. The growth of social media is remarkable, with the number of users escalating from 2.08 billion in 2015 to an estimated 4.26 billion by 2021, and projections nearing 6 billion by 2027 [2]. This vast and expanding user base has made social media a crucial tool for marketers, enabling them to connect with millions of potential customers rapidly [3].

Youth, including middle school students and college-aged individuals, represent a key demographic on social media. This group is highly active and significantly influences the virality of products and trends [4, 5]. According to a 2022 Pew Research Center report, 95% of U.S. adolescents aged 13 to 17 use YouTube, 67% engage with TikTok, and 62% are active on Instagram [6]. Despite their high engagement levels, these users display diverse reactions to advertisements and economic behaviors, impacting the effectiveness of targeted marketing strategies.

Understanding how social media engagement, advertising methods, and economic behaviors influence purchase intentions among students is crucial. This study aims to address this by utilizing advanced machine learning techniques to analyze these factors' impact on purchase intentions. The research seeks to answer key questions about the drivers of purchase decisions among students, offering actionable insights for marketers to enhance their strategies.

II. RELATED WORK

Related Work

Research into social media's impact on consumer behavior has surged, reflecting its evolving role in marketing and consumer engagement. Early studies laid the foundation by exploring how platforms like Facebook and Twitter facilitated peer-to-peer communication and word-of-mouth marketing, significantly influencing consumer decisions [7, 8].

With the rise of visual-centric platforms such as Instagram, YouTube, and TikTok, attention shifted to how visual content affects consumer engagement and purchase intentions. Research indicates that visual content, due to its immersive nature, enhances engagement more effectively than text-based content, leading to higher purchase intentions [9, 10]. Influencers on these platforms play a crucial role by enhancing trust and credibility, thus increasing purchase likelihood [11, 12].

The behaviors of young social media users have been the subject of extensive research. This demographic's high engagement with social media and their role in spreading viral content make them particularly susceptible to influencer marketing and brand promotions, which can drive impulsive buying [13, 14]. Studies have shown that young users perceive social media advertisements as more authentic and trustworthy when delivered by influencers or through user-generated content compared to traditional advertising [15].

Recent studies have examined specific elements of social media marketing that influence young adults' and adolescents' behavior. Key factors include the perceived informativeness, entertainment value, and credibility of social media content, which significantly impact consumer engagement and purchase intentions [16, 17]. Additionally, the concept of social proof, where individuals mimic the actions of their social network, has been shown to amplify the effects of social media marketing, especially in viral campaigns [18].

In specific contexts, such as the Lebanese market, viral marketing strategies have been shown to influence purchasing decisions through electronic word-of-mouth (EWOM) and consumer-generated content [19]. Similarly, scholars also explored how consumer interaction and shopping motivation, mediated by EWOM, brand awareness, and brand attitude, affect purchase intention. Their findings indicate that while EWOM directly impacts purchase intention, brand awareness and viral advertisements do not [20].

Machine learning models have been demonstrated significantly outperform traditional statistical methods in predicting purchase intentions, highlighting the potential of advanced analytics in understanding consumer behavior [21]. Researchers applied deep learning models to Twitter data, showing that LSTM Neural Networks effectively predict consumer purchase intentions based on social media interactions [22]. Additionally, the influence of online reviews as EWOM, revealing that review cues like star ratings and length impact perceived usefulness and purchase decisions [23].

Despite these insights, there remains a gap in understanding how various social media engagement metrics, advertising strategies, and economic behaviors collectively influence purchase intentions. Conventional statistical methods frequently fall short in capturing the intricate interactions between these factors.Our study aims to bridge this gap by employing machine learning techniques to analyze the combined effects of social media and economic behaviors on students' purchase intentions, benefiting both academic research and practical marketing strategies.

III. METHODOLOGY

The methodology for this study is designed to rigorously analyze the factors influencing purchase intentions among students using a machine learning approach. The research process is organized into several critical stages: data collection, preprocessing, model selection, feature selection, and impact analysis.

Data Collection

Data for this study was collected through a comprehensive questionnaire distributed to studentsaged from 11 to 22, across various educational levels, including middle school, high school, and college. The questionnaire was disseminated using a combination of snowball sampling and targeted social media posts, resulting in a total of 2,041 data points from July 2024.This study did not require approval by the ethics committee, because the participants voluntarily attended the survey and there was not sensitive personal information.

The questionnaire captured a wide range of variables, including demographic details, social media engagement, economic behaviors, and opinions on advertising strategies. Variables were derived from the survey questions, including "Gender", "Race", "Age", "Time spent", "People influenced", "Part time income", "Debit credit cards", "Paid by self", "Consuming sharing enjoyable", "Ads good info", "Ads up-to-date info", "Ads complete info", "Ads convenient info", "Ads trust", "Following increase trust", "Using increase interest", "Ads purchase reference", "Visit brand page", "Read brand post", "Use like option", "Comment post", "Like brands use ads", "Use products promoted", "Search product posts", "Purchase promotions", "Purchase recommended", and "#Social media".Specific questions were included to assess students' opinions on statements such as "Viral advertisements like challenges and videos are a good source of product information" and "I am more likely to purchase a product if I see it being used by influencers on social media."Detailed survey questions are provided in Appendix A: Survey Cover Letter and Questionnaire.

Our primary target variable, *viral product purchase intention*, is a binary variable based on the survey question regarding whether the participant has purchased products promoted through viral marketing. Since our objective is to predict the likelihood of purchasing viral products, other variables such as "Purchase Increase," "Purchase Promotions," and "Show Friends Promotions" were excluded from this analysis but may be considered in future research.

Data Preprocessing

Once the data was collected, it underwent a thorough preprocessing phase to ensure its quality and suitability for analysis. This phase included the elimination of blank responses and the imputation of irrelevant or incomplete responses. The cleaned data was then quantified, with categorical variables being encoded into numerical formats suitable for machine learning analysis.

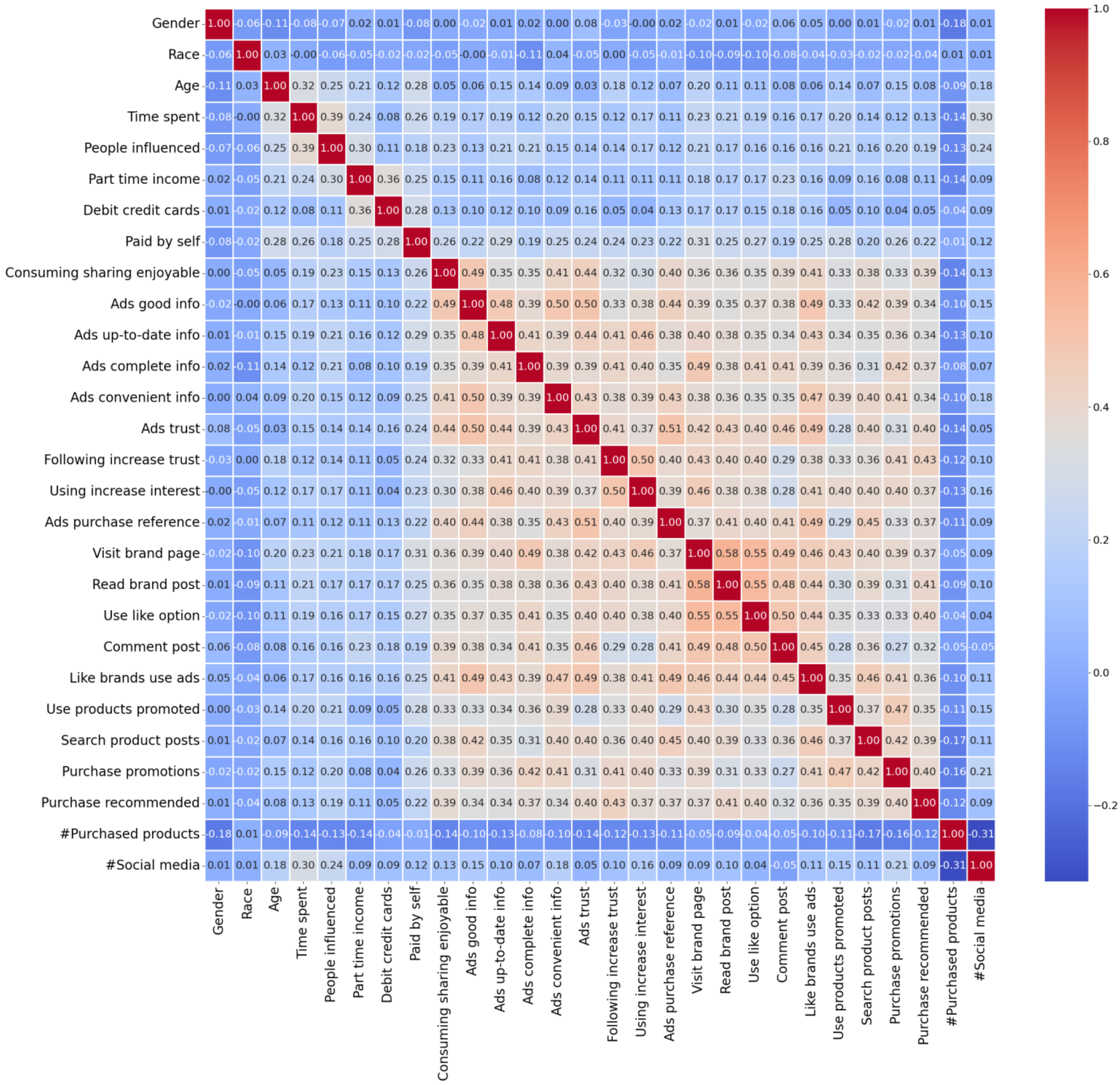

Figure 1. Feature Correlation Analysis.

Feature Selection

To mitigate overfitting and enhance model robustness, a thorough feature selection process was employed. The first step involved employing filter methods, including Pearson Correlation and Variance Inflation Factor (VIF), to assess the relevance of features.Pearson Correlation assessed the linear relationships between features and the target variable, retaining features with a correlation coefficient above 0.001, as depicted in Figure 1. Concurrently, VIF was utilized to detect and address multicollinearity, eliminating features with high redundancy. Next, wrapper methods such as Step Forward Selection, Backward Elimination, and Recursive Feature Elimination (RFE) were applied.Step Forward Selection incrementally added features based on their impact on model performance, while Backward Elimination employed Random Forest Regressors to iteratively remove features until the optimal subset was achieved. RFE further streamlined the process by systematically eliminating the least significant features.

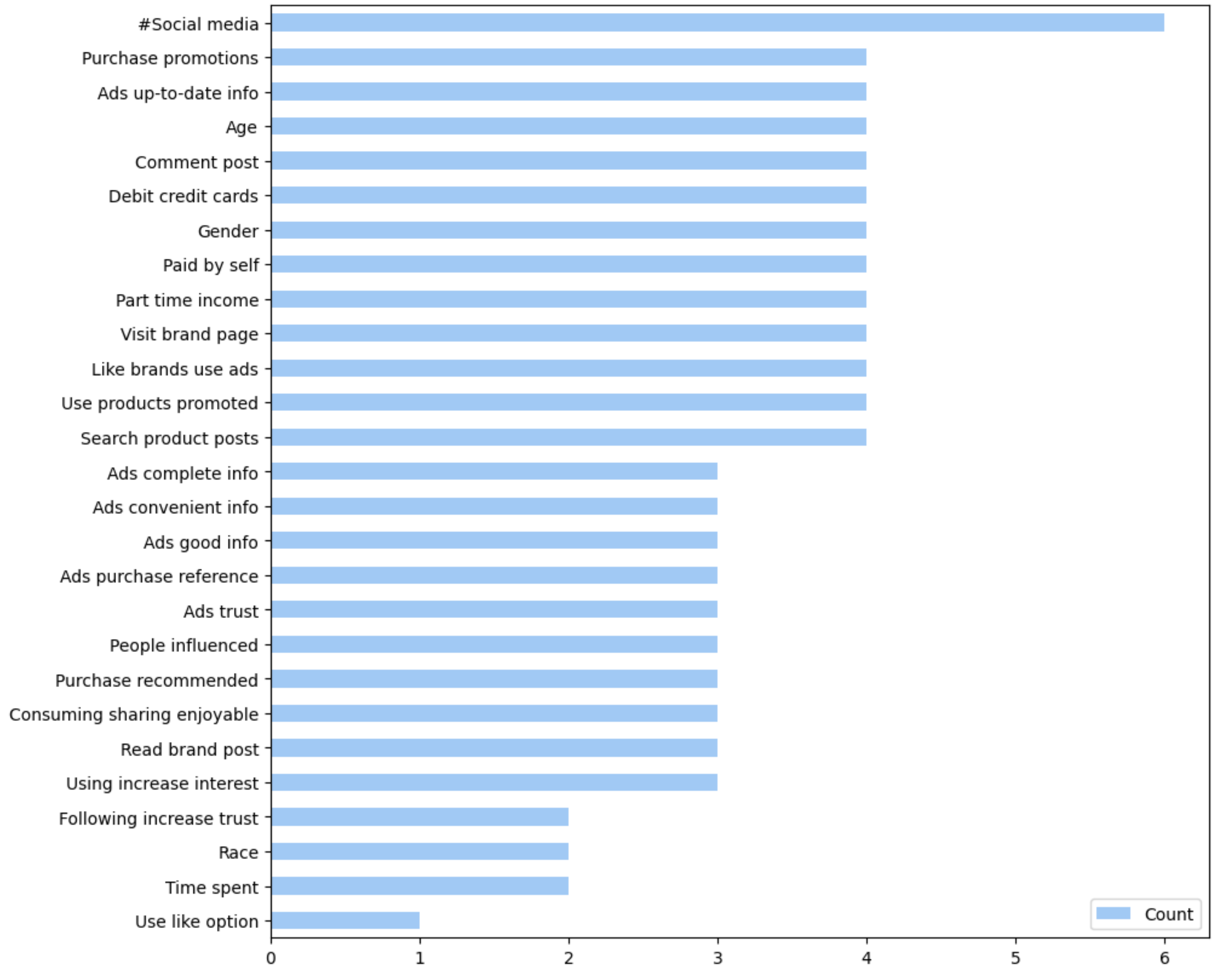

Figure 2. Feature Importance Distribution Based on Multiple Selection Techniques

In addition to these methods, embedded techniques such as LassoCV were applied to penalize features with negligible coefficients, thus focusing on those with higher predictive significance. Cross-validation was employed to identify the ideal number of features, ensuring that the final set was both robust and predictive.This multi-faceted approach resulted in a final set of 26 relevant features, excluding the "Use like option" due to its minimal effect on model accuracy. This meticulous feature selection process ensured that only the most impactful variables were included, thereby enhancing the overall performance of the model.Figure 2 provides a visual summary of feature relevance by displaying how frequently each feature was selected across the different methods used.

Table 1. Best Hyperparameters of the Algorithms

Algorithms	Best Hyperparameters
Logistic Regression	{‘C’: 0.5, ‘max_iter’: 1000, ‘random_state’: 0}
Support Vector Classify	{‘C’: 1, ‘kernel’: ‘rbf’, ‘random_state’: 0}
Gaussian Naive Bayes	{‘var_smoothing’: 1e-09}
Random Forest	{‘criterion’: ‘gini’, ‘n_estimators’: 100, ‘random_state’: 0}

Extreme Gradient Boosting	{‘eval_metric’: ‘error’, ‘learning_rate’: 0.05, ‘random_state’: 0}
Extra Trees	{‘criterion’: ‘entropy’, ‘n_estimators’: 150, ‘random_state’: 0}
Gradient Boosting	{‘learning_rate’: 0.1, ‘random_state’: 0}
Adaptive boosting	{‘learning_rate’: 0.05, ‘n_estimators’: 150, ‘random_state’: 0}

Model development

To tackle data imbalance and enhance prediction performance, we employ the Synthetic Minority Over-sampling Technique (SMOTE). We use eight different machine learning models: Logistic Regression, Support Vector Classifier, Gaussian Naive Bayes, Random Forest, Extreme Gradient Boosting, Extra Trees, Gradient Boosting, and Adaptive Boostingto develop predictive models.

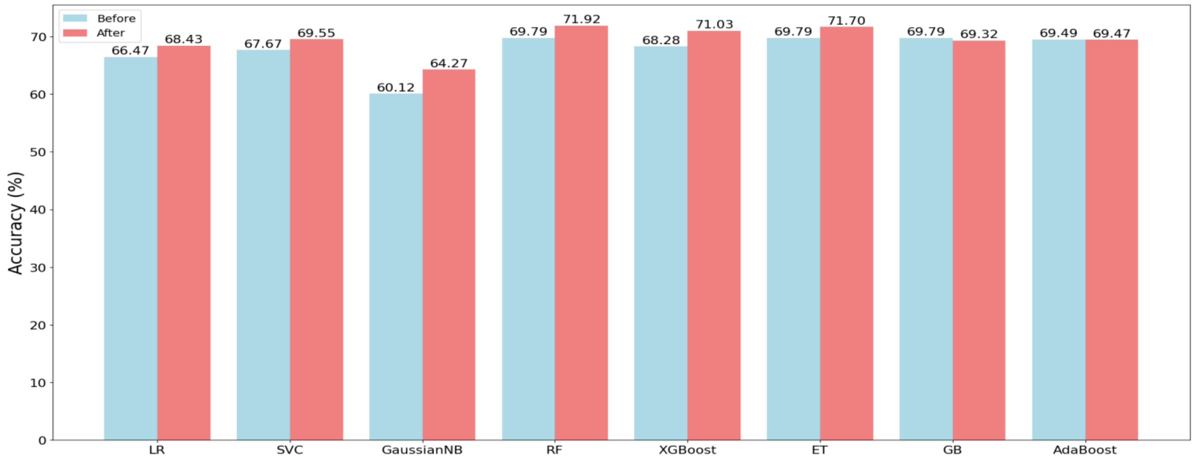

Figure 3. Accuracy of All the Boosting Algorithms Before and After Conducting Hyperparameter Tuning.

Grid Search is used for hyperparameter tuning, which involves exhaustively testing a predefined set of hyperparameter combinations to optimize model performance. Each combination is evaluated through tenfold cross-validation, where the dataset is split into ten subsets. The model is trained on nine of these subsets and tested on the remaining one. This process is repeated ten times, and the average performance across all iterations is used to determine the optimal hyperparameter configuration. This approach ensures reproducibility by setting a random seed and aims to determine the hyperparameters that maximize predictive accuracy and generalizability. The grid search process is systematically performed for each model to find the most effective hyperparameter settings, thereby optimizing their performance on the dataset. Table 1 provides a detailed overview of the specific hyperparameter settings for each classifier. The listed values for each parameter for the respective algorithms were found to be the best performers.

Figure 3 displays the testing accuracy of the eight machine learning algorithms. The figure presents a comparison of the accuracy of the considered algorithms before and after conducting hyperparameter tuning. Initially, Random Forest (RF), Extra Trees (ET), and Gradient Boosting (GB) each achieved the highest accuracy, reaching 69.79%.After hyperparameter tuning, Random Forest (RF) emerged as the top performer, with an improved accuracy of 71.92%, followed closely by Extra Trees (ET), which achieved an accuracy of 71.70%.

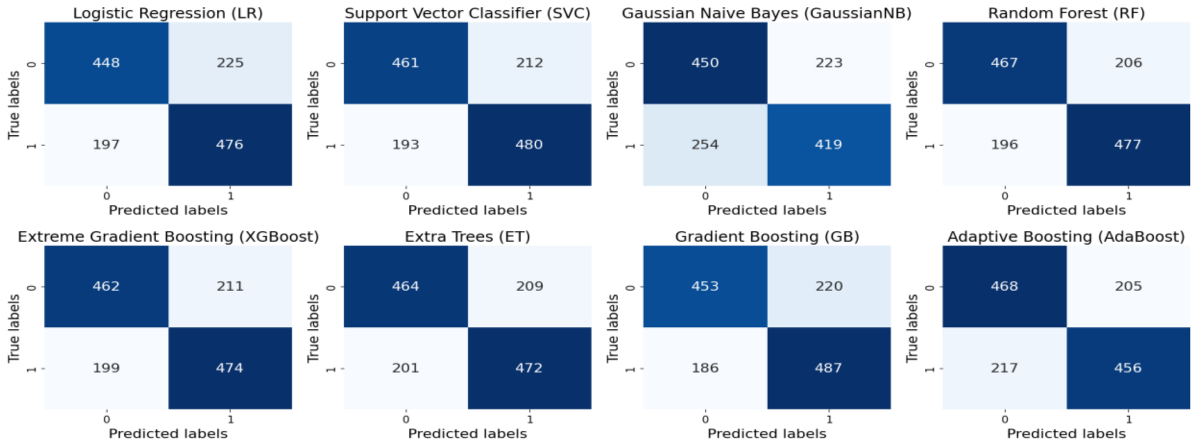

Figure 4. Performance Analysis with Confusion Matrix in 10-Fold Cross-Validation.

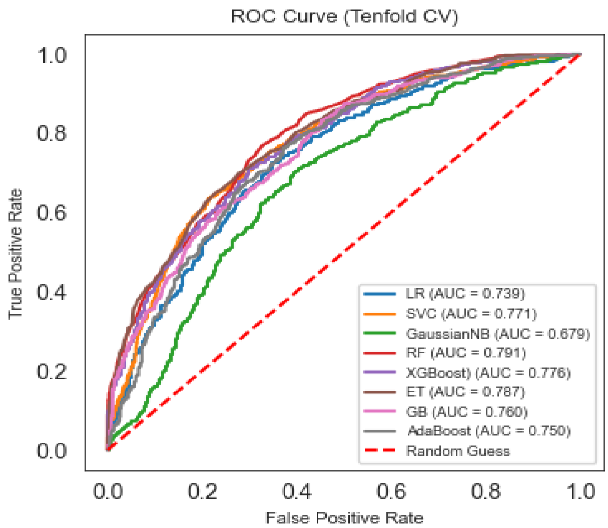

Figure 5. Performance Comparison of ROC curve

Model Evaluation

Figure 4 displays the confusion matrices for eight machine learning models—Logistic Regression, Support Vector Classifier, Gaussian Naive Bayes, Random Forest, Extreme Gradient Boosting, Extra Trees, Gradient Boosting, and Adaptive Boosting—derived from a ten-fold cross-validation process.This visualization reveals how well each model distinguishes between classes, with the Random Forest (RF) model showing the most favorable performance metrics.Figure 5 provides additional insight by showing the performance of these models through ROC curves, which compare their true positive rates against false positive rates. The ROC curve for Random Forest (RF) highlights its superior predictive performance, with the highest ROC AUC score of 0.791., indicating its effectiveness across different classification thresholds.

Table 2 offers a comprehensive quantitative summary of the models' performance metrics from the ten-fold cross-validation, including accuracy, precision, recall, F1 score, and ROC AUC for each model. The table confirms the findings from Figures 6 and 7, with Random Forest (RF) consistently outperforming the other models. Specifically, it achieves the highest accuracy (0.712, 95% CI, 0.687–0.738), precision (0.715, 95% CI, 0.689–0.741), recall (0.712, 95% CI, 0.687–0.738), F1 score (0.712, 95% CI, 0.687–0.737), and ROC AUC (0.791, 95% CI, 0.768–0.815).

Together, these figures and the table provide a comprehensive assessment of model performance, demonstrating the Random Forest (RF) model's superior capabilities across multiple evaluation metrics and visualizations.

Table 2. Classification Model Performance in 10-fold Cross-validation strategy.

Model	Accuracy	Precision	Recall	F1 Score	ROC AUC
Logistic Regression (LR)	0.683 (0.661, 0.705)	0.684 (0.663, 0.705)	0.683 (0.661, 0.705)	0.682 (0.660, 0.705)	0.738 (0.716, .760)
Support Vector Classifier (SVC)	0.695 (0.681, 0.709)	0.697 (0.684, 0.711)	0.695 (0.681, 0.709)	0.695 (0.681, 0.709)	0.767 (0.747, 0.786)
Gaussian Naive Bayes (GaussianNB)	0.644 (0.621, 0.667)	0.647 (0.624, 0.670)	0.644 (0.621, 0.667)	0.643 (0.619, 0.667)	0.679 (0.644, 0.714)
Random Forest (RF)	**0.712 (0.687, 0.738)**	**0.715 (0.689, 0.741)**	**0.712 (0.687, 0.738)**	**0.712 (0.687, 0.737)**	**0.791 (0.768, 0.815)**
Extreme Gradient Boosting (XGBoost)	0.700 (0.676, 0.723)	0.701 (0.677, 0.725)	0.700 (0.676, 0.723)	0.700 (0.676, 0.723)	0.776 (0.755, 0.798)
Extra Trees (ET)	0.708 (0.679, 0.737)	0.709 (0.680, 0.738)	0.708 (0.679, 0.737)	0.708 (0.679, 0.737)	0.787 (0.763, 0.811)
Gradient Boosting (GB)	0.674 (0.642, 0.707)	0.676 (0.643, 0.709)	0.674 (0.642, 0.707)	0.674 (0.642, 0.706)	0.760 (0.733, 0.786)
Adaptive Boosting (AdaBoost)	0.691 (0.670, 0.712)	0.693 (0.673, 0.714)	0.691 (0.670, 0.712)	0.691 (0.669, 0.712)	0.750 (0.727, 0.772)

Feature Attribution Analysis.

We can analyze their influence on the shield attitude control performance by calculating the SHAP value of the features and plotting figures that depict the feature contributions and feature dependence. Feature importance was determined using Shapley additive explanations (SHAP) values. SHAP helps explain machine learning model predictions by showing how much each feature contributes to the model's output. It calculates the average effect of a feature on predictions, comparing the model's performance with and without that feature[24].

(1) Impact analysis of single feature. In Figure 6(Panel A), we summarize our key findings for an easier global explanation of the impact of the 32 features on the Random Forest (RF) model and their association with viral product purchase intention. The horizontal length of each bar represents the extent of the feature's impact on the model. The features are sorted by decreasing mean relevance. The strongest predictors of viral product purchase intention were the number of social media used ("# Social media"), followed by being male, being female,

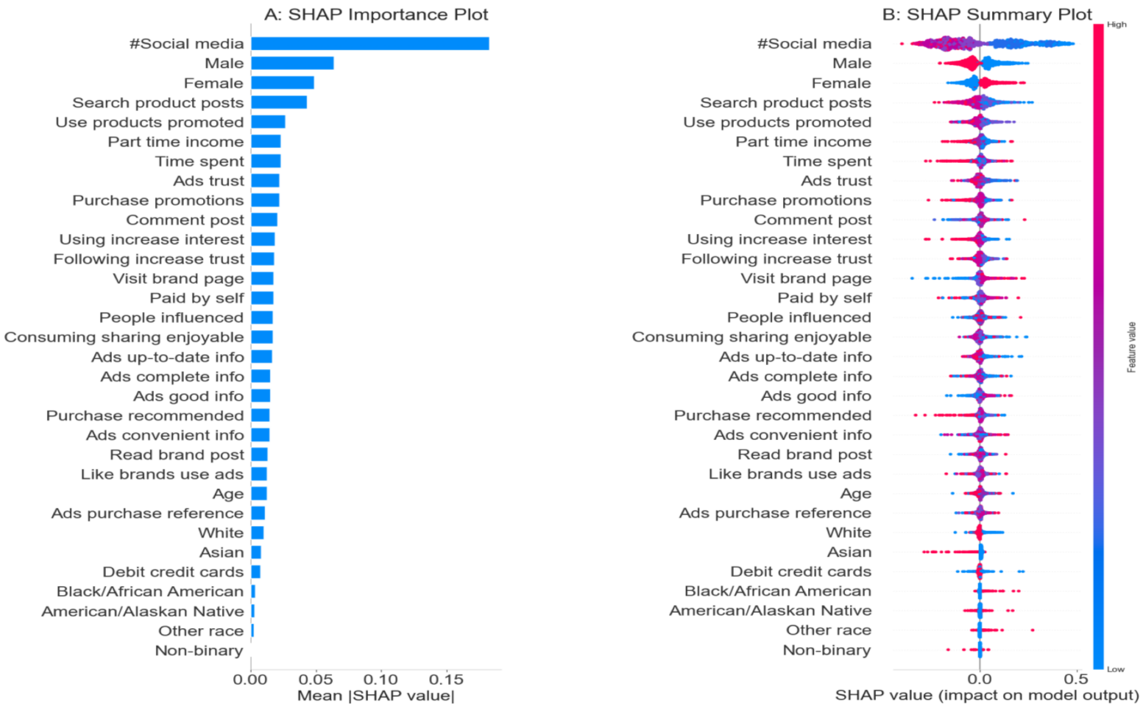

Figure 6. Summary of Feature Impact on Predictions

searching product posts, using products promoted, part time income, time spent on social media, advertisement trust, purchasing promotions, commenting on brand pages, increasing purchase interest via brand on social media ("Using increase interest"), increasing trust via following brands on social media ("Following increase trust"), visiting brand page, paid-by-self, number of nearby social media-driven buyers ("People influenced"), enjoyment of sharing viral brand marketing ("Consuming sharing enjoyable"), up-to-date product info provided by social media ("Ads up-to-date info"), complete product info provided social media ("Ads complete info"), social media as a good source of product information ("Ads good info"), purchasing recommendations, social media as a convenient source of product information ("Ads convenient info"), reading brand post, liking brands that use ads for communication ("Like brands use ads"), age, using ads as a purchase reference ("Ads purchase reference"), being White, Asian, Black/African American, American Indian/Alaskan Native, another race, non-binary, and owning debit/credit cards.

Figure 6(Panel B) displays a SHAP summary plot that illustrates feature importance, effect size, and the direction of association with viral product purchase intention, ordered by average contribution to the prediction. The plot on the x-axis shows SHAP values for each observation, where negative values indicate a decreased probability of viral product purchase intention, and positive values indicate an increased probability. Red dots represent higher feature values, while blue dots indicate lower values. Each point

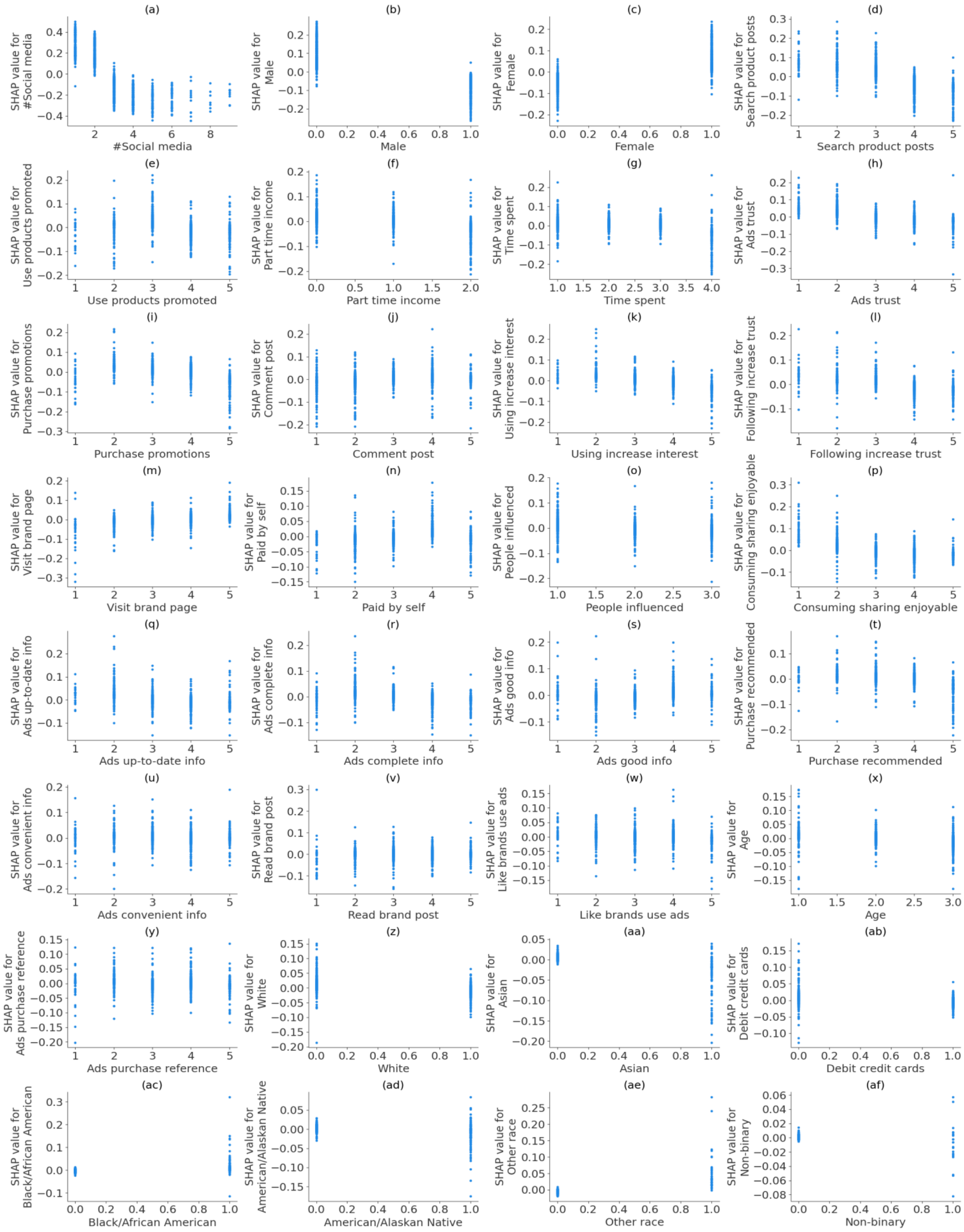

Figure 7. SHAP Single Feature Dependence Plots

reflects an individual observation, with its x-axis position showing the feature's influence on the model's output. The SHAP summary plot reveals that a larger number of social media platforms used (“# Social Media,”red dots) is associated with a lower incidence of viral product purchase intention (SHAP value < 0). Conversely, being male (red dots) is linked to a lower incidence of viral product purchase intention, while being female (red dots) is associated with a higher incidence (SHAP value > 0). Additionally, searching product posts, using promoted products, part-time income, and advertisement trust (all red dots) are associated with a lower incidence of viral product purchase intention (SHAP values < 0).

To further analyze feature traits, we created single feature partial dependence plots (Figure 7) to observe how feature contributions change with feature values. In these subplots, the x-axis represents the feature value, while the y-axis displays the SHAP value for the feature.Figure 7 (a) shows that a higher number of social media platforms used (“# Social Media”) negatively affects purchase intention. Figure 7 (b), (c), and (af) suggest that being male is associated with a lower incidence of viral product purchase intention, while being female is linked to a higher incidence. The connection between non-binary identity and the intention to purchase viral products is still uncertain.

Figure 7 (d), (e), (f), (g), (h), (i), (k), (l), (p), (q), (r), and (t) demonstrate that the features “Search product posts,” “Use products promoted,” “Part-time income,” “Time spent,” “Ads trust,” “Purchase promotions,” “Using increase interest,” “Following increase trust,” “Consuming sharing enjoyable,” “Ads up-to-date info,” “Ads complete info,” and “Purchase recommended” generally exhibit an inverse relationship with their contribution values.

Figure 7 (z), (aa), (ae), and (ad) indicate that being Asian is linked to a lower incidence of viral product purchase intention, while being of“Black/African American” or “Other race” is linked to a higher incidence. The relationships between being White or American/Alaska Native and viral product purchase intention are still unclear.

Figure 7 (m) shows a positive relationship between “Visit brand page” and its contribution to the probability of viral product purchase intention. Finally, Figure 7 (j), (n), (o), (s), (u), (v), (w), (x), (y), and (ab) reveal that the impact of “Comment post,” “Paid by self,” “People influenced,” “Ads good info,” “Ads convenient info,” “Read brand post,” “Like brands use ads,” “Age,” “Ads purchase reference,” and “Debit credit cards” on viral product purchase intention is not clearly defined.

(2) Dependence analysis of double features. In many cases, the impact of one feature on an outcome is dependent on another feature. To better understand these interactions, we present double feature partial dependence plots (Figure 8). These plots illustrate the relationship between pairs of features and their combined influence on the outcome, as well as the distribution characteristics of these interactions. In these plots, the x-axis represents the value of the primary feature, the y-axis shows the SHAP value for the same feature, and the color of each point indicates the value of the second feature.

The first set of figures demonstrates interactions that reduce the likelihood of viral product purchase intention. Figure 8 (a) shows that as "Visit brand page" increases, the effect of "Read brand post" on purchase intention diminishes. Similarly, Figure 8 (b) indicates that higher levels of "Using increase interest" weaken the impact of "Following increase trust" on purchase intention. Figure 8 (c) reveals that increased "Ads trust" reduces the influence of "Ads purchase reference," while Figure 8 (h) and (l) show that both "Ads purchase reference" and "Like brands use ads" lessen each other's influence and the impact of "Ads trust," respectively. Additionally, Figure 8 (t) and (y) highlight those higher levels of "Consuming sharing enjoyable" and "Ads up-to-date info" further diminish the influence of "Ads trust."

Other figures in this set also illustrate how increased interaction between certain features decreases the likelihood of a viral product purchase intention. For instance, Figure 8 (ac) shows that as "Search product posts" increases, the effect of "Purchase promotions" decreases. In Figure 8(ae), a higher "Ads purchase reference" weakens the impact of "Search product posts." Figure 8 (ah) and (aj) suggest that increased "Like brands use ads" reduces the influence of "Purchase promotions" and "Consuming sharing enjoyable" on purchase intention. Similarly, Figure 8 (n), (o), (s), (u), (w), (x), (aa), and (ab) all demonstrate that various feature interactions generally reduce the probability of a viral product purchase intention. These figures highlight the complex ways in which these variables affect consumer behavior.

Conversely, the second set of figures highlights interactions that increase the likelihood of viral product purchase intention. For example, Figure 8 (d) shows that as "Ads up-to-date info" increases, the effect of "Ads good info" on purchase intention also increases. Figure 8 (e) reveals that a higher "Visit brand page" amplifies the influence of "Comment post" on purchase intention. Similarly, Figure 8 (f) illustrates that an increase in "Ads trust" strengthens the impact of "Ads good info" on purchase intention Figure 8 (i) indicates that higher "Read brand post" enhances the effect of "Comment post" on purchase intention, while Figure 8 (j) shows that an increase in "Consuming sharing enjoyable" boosts the influence of "Ads good info."

Additional figures also support the positive interactions between features that can enhance the likelihood of a viral product purchase intention. Figure 8 (k) demonstrates that "Ads complete info" increases the effect of "Visit brand page" on purchase intention. Figure 8 (m) suggests that a higher level of "Like brands use ads" magnifies the impact of "Ads good info." In Figure (p), an increase in "Like brands use ads" amplifies the effect of "Comment post." Similarly, Figure 8 (q) shows that higher "Ads trust" heightens the impact of "Comment post" on purchase intention. Figure 8 (v) and (z) depict that an increase in "Using increase interest" and "Following increase trust," respectively, strengthens the effect of "Visit brand page" on purchase intention. Finally, Figure 8 (ad) reveals that higher "Ads purchase reference" enhances the influence of "Ads good info" on purchase intention.

Overall, these figures highlight that certain combinations of feature interactions can either negatively or positively affect the likelihood of a viral product purchase intention. Understanding these dynamics allows for more strategic alignment of features to either minimize or maximize consumer interest, depending on the desired outcome.

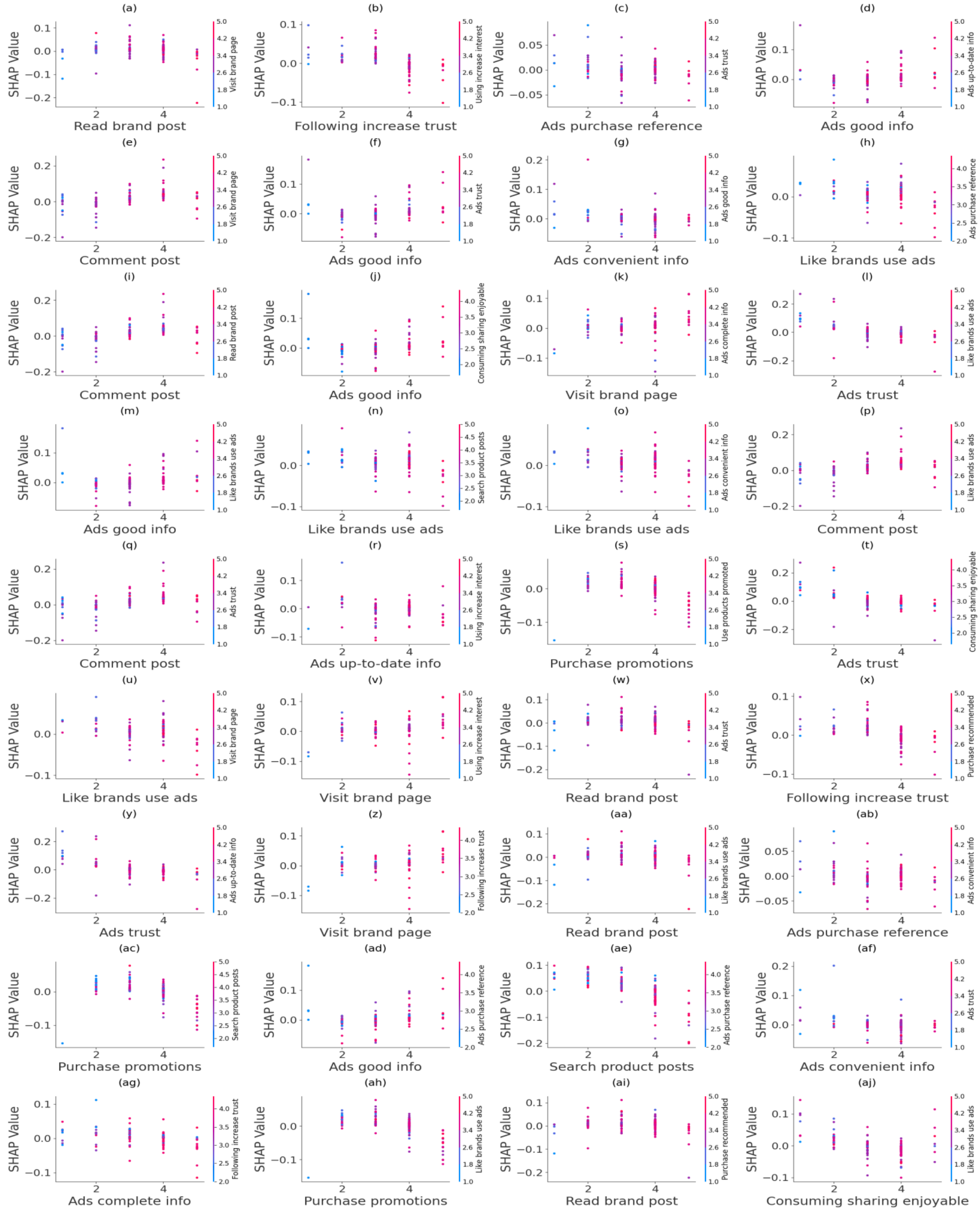

Figure 8. SHAP Double Feature Dependence Plots

IV. DISCUSSION

Interpretation of Findings

The results of this study demonstrate the significant influence of social media engagement, advertising strategies, and economic behaviors on the purchase intentions of students. The Random Forest (RF) model emerged as the most effective predictive tool, achieving high accuracy across multiple evaluation metrics. The SHAP analysis further highlighted the importance of specific features, such as the number of social media platforms used, gender, and advertisement trust, in shaping viral purchase intentions.

The negative association between the number of social media platforms used and viral product purchase intention suggests that overexposure to various platforms may dilute the effectiveness of viral marketing efforts. This finding contrasts with the traditional view that broader social media engagement directly correlates with higher purchasing behavior. Instead, it indicates a potential saturation effect, where excessive use of social media platforms might lead to consumer fatigue or skepticism.

The gender differences revealed by the SHAP analysis also warrant attention. The data indicates that female students are more likely to be swayed by social media marketing than male students. This aligns with previous research indicating that women are generally more responsive to social media advertising due to their higher engagement levels and affinity for social interaction online.

Comparison with Previous Studies

These findings are consistent with those of other studies that emphasize the role of social media in shaping consumer behavior, particularly among younger demographics. However, this study adds a nuanced understanding by using advanced machine learning techniques to quantify the impact of individual features on purchase intentions. For example, while earlier research has highlighted the significance of social media in marketing, this study specifically identifies how the number of platforms used can have a counterintuitive effect on purchase behavior.

Additionally, the study corroborates earlier findings on the influence of gender on social media marketing effectiveness, further supporting the idea that targeted marketing strategies should account for demographic differences to maximize impact.

Limitations and Future Research

This study highlights factors influencing students' viral purchase intentions on social media but has limitations. Firstly, the use of snowball sampling and targeted social media posts may introduce sampling bias, limiting the generalizability of findings. Future research should adopt more diverse and representative sampling methods to enhance the applicability of results.

Furthermore, using self-reported data may introduce inaccuracies due to social desirability and recall biases. Incorporating objective measures, like actual social media usage or purchase records, alongside self-reported data, could improve accuracy in future studies.

The study also doesn't account for the distinct effects of different social media platforms, such as Instagram, TikTok, and YouTube, each with unique user demographics and content formats. Future research should explore platform-specific effects to better understand how various social media environments influence purchase intentions.

Moreover, while advanced machine learning techniques were used, these models can be limited in interpretability. Future research might explore more interpretable models or hybrid approaches that balance accuracy with transparency, providing actionable insights for marketers and policymakers.

Addressing these limitations will deepen the understanding of social media's impact on student purchase intentions and offer valuable guidance for marketers and policymakers in this evolving landscape.

V. CONCLUSION

In conclusion, this study provides valuable insights into the factors driving purchase intentions among students, a group that is deeply engaged with social media and critical to the effectiveness of viral marketing campaigns. By employing advanced machine learning techniques, we have been able to identify key factors such as social media engagement, economic behaviors, and perceptions of advertising that significantly influence students' likelihood of purchasing products promoted through viral marketing.

The findings of this research emphasize the need for marketers to carefully consider these factors when designing campaigns aimed at younger audiences. Understanding how different variables interact and contribute to purchase intentions can help marketers craft more effective strategies that resonate with this critical demographic. Ultimately, this study contributes valuable insights into the complex relationship between social media engagement and consumer behavior, offering practical insights for marketers aiming to leverage social media to boost sales and enhance brand loyalty.

REFERENCES

[1]. Guan, J., Lau, Y. Y., Yang, H., & Ren, L. (2022). To buy or not to buy: how young consumers approach new smart products in the social media context. Young Consumers, 23(1), 90-111.

[2]. Jafar, Z., Quick, J.D., Larson, H.J., Venegas-Vera, V., Napoli, P., Musuka, G., Dzinamarira, T., Meena, K.S., Kanmani, T.R., &Rimányi, E. (2023). Social media for public health: Reaping the benefits, mitigating the harms. *Health promotion perspectives*, *13*(2), 105.

[3]. Tsimonis, G., & Dimitriadis, S. (2014). Brand strategies in social media. Marketing intelligence & planning, 32(3), 328-344.

[4]. Chung, A., Vieira, D., Donley, T., Tan, N., Jean-Louis, G., Gouley, K. K., & Seixas, A. (2021). Adolescent peer influence on eating behaviors via social media: scoping review. Journal of medical Internet research, 23(6), e19697.

[5]. Cash, S. B., Alhabash, S., Fretes, G., & Ma, M. (2022). Food, beverages and social media: trends and tools for economic research. In A Modern Guide to Food Economics (pp. 297-326). Edward Elgar Publishing.

[6]. Pew Research Center. (2022). Teens, Social Media and Technology 2022. https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/

[7]. Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business horizons, 53(1), 59-68.

[8]. Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. Business horizons, 52(4), 357-365.

[9]. De Vries, L., Gensler, S., & Leeflang, P. S. (2012). Popularity of brand posts on brand fan pages: An investigation of the effects of social media marketing. Journal of interactive marketing, 26(2), 83-91.

[10]. Voorveld, H. A., Araujo, T., Bernritter, S. F., Rietberg, E., &Vliegenthart, R. (2018). How advertising in offline media drives reach of and engagement with brands on Facebook. International Journal of Advertising, 37(5), 785-805.

[11]. Freberg, K., Graham, K., McGaughey, K., & Freberg, L. A. (2011). Who are the social media influencers? A study of public perceptions of personality. Public relations review, 37(1), 90-92.

[12]. Lou, C., & Yuan, S. (2019). Influencer marketing: How message value and credibility affect consumer trust of branded content on social media. Journal of interactive advertising, 19(1), 58-73.

[13]. Djafarova, E., & Rushworth, C. (2017). Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users. Computers in human behavior, 68, 1-7.

[14]. Casaló, L. V., Flavián, C., & Ibáñez-Sánchez, S. (2020). Influencers on Instagram: Antecedents and consequences of opinion leadership. Journal of business research, 117, 510-519.

[15]. Boerman, S. C., Kruikemeier, S., & Zuiderveen Borgesius, F. J. (2017). Online behavioral advertising: A literature review and research agenda. Journal of advertising, 46(3), 363-376.

[16]. Duffett, R. G. (2017). Influence of social media marketing communications on young consumers' attitudes. *Young Consumers*, *18*(1), 19-39.

[17]. Shareef, M. A., Mukerji, B., Dwivedi, Y. K., Rana, N. P., & Islam, R. (2019). Social media marketing: Comparative effect of advertisement sources. Journal of Retailing and Consumer Services, 46, 58-69.

[18]. Moe, W. W., &Schweidel, D. A. (2012). Online product opinions: Incidence, evaluation, and evolution. Marketing Science, 31(3), 372-386.

[19]. Hamade, L. (2021). Viral Marketing and Its Influence on Customer's Purchasing Decisions: A Quantitative Study in Lebanon. *Resilience and Economic Intelligence Through Digitalization and Big Data Analytics*, 331-340.

[20]. Kusumawardani, K., &Purniasari, P. (2021). The effect of online consumer interaction and shopping motivation on purchase intention. FIRM Journal of Management Studies, 6(2), 183-202.

[21]. Islam, T., Meade, N., Carson, R. T., Louviere, J. J., & Wang, J. (2022). The usefulness of socio-demographic variables in predicting purchase decisions: Evidence from machine learning procedures. *Journal of Business Research*, *151*, 324-338.

[22]. Imran, H., Zaidi, S. Z., Khan, N. S., & Uddin, N. (2022). Predicting Purchase Intention Using Twitter Data.

[23]. Majumder, M. G., Gupta, S. D., & Paul, J. (2022). Perceived usefulness of online customer reviews: A review mining approach using machine learning & exploratory data analysis. *Journal of Business Research*, *150*, 147-164.

[24]. Lundberg, S. M., Erion, G. G., & Lee, S. I. (2018). Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888*.