Social Media Promotion Timing to Influence Consumer Behavior in E-Pharmacies

Bagavathy K L
Department of BBA
Vels Institute of Science Technology and Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India
srimkp29@gmail.com

Harihara Shanmugam A R
Department of BBA
Vels Institute of Science, Technology and Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India
rnrhari@gmail.com

***Abstract*-Social media channels have become absolutely vital tool for online pharmacies seeking customers and higher sales. Since it can significantly influence consumer behavior, the timing of promotional activities should be given great attention. Regarding promotions, many online pharmacies lack strategic knowledge of the optimal timing, which results in lost opportunities for better consumer involvement and conversion. This is true even if social media marketing is becoming ever more trendy. The purpose of this study is to establish the relationship between the timing of social media campaigns and the purchasing behavior of possible customers visiting online pharmacies. Over three months, data was gathered from the most well-known e-pharmacy systems under special focus on consumer interactions, click-through rates, and purchase behavior at various intervals. Combining sentiment analysis with Random Forest regression techniques, consumer responses to promotions distributed at several times of the day and week was investigated. The matching marketing plans with consumer routines is emphasized by the fact that conversion rates show to be higher during the week than on the weekends. The research also underlines how customized promotional material improves the impact of well-timed postings, so implying a harmonic interaction between timing and content relevance.**

Keywords- Social media promotions, e-pharmacies, consumer behavior, marketing strategy, promotion timing

I. Introduction

In social media into current marketing strategies has changed how companies interact with their consumers. Thanks to the worldwide expansion of e-commerce, consumers could now quickly obtain drugs and other healthcare products by means of online pharmacies. These days, social media sites Facebook, Instagram, and Twitter are considered as absolutely necessary tools for online pharmacies. These systems let businesses interact with customers and influence their purchase decisions. Studies abound demonstrating that practically eighty percent of consumers globally learn about brands and products using social media [1-3]. For online pharmacies, this presents an opportunity to establish trust and interaction with their customers especially with campaigns alerting consumers about the availability of products, discounts, and new product introductions.

Even with their widespread use, e-pharmacies still struggle greatly in properly timing promotional posts across social media platforms. Online pharmacies struggle to stand out and attract customers from the abundance of ads on platforms [4–6]. Research studies have shown that poorly timed promotions sometimes fall short of the intended audience or the degree of involvement sought for. Another factor aggravating the problem in advertising materials is their lack of personalizing since it lowers click-through rates and limits conversion rates. The very competitive character of the e-pharmacy sector, in which businesses must innovate to keep customer loyalty while also controlling operational costs [4-6], is just aggravating the difficulty of this problem.

About the e-pharmacy industry, little is known about how timing of social media campaigns influences consumer behavior. Though there are broad marketing strategies for social media timing, these concepts are not tailored to the specific purchase behavior and needs of customers who visit online pharmacies [7–10]. Lack of empirical data on timing-based marketing strategies for online pharmacies leaves companies with little direction on how best maximize their promotional plans to reach the highest possible impact.

This study has as among its several primary goals:

1. To investigate how consumer purchase behavior and degree of involvement during shopping at online pharmacies change with timing of promotional activities.
2. To run campaigns at the ideal times will help one maximize customer satisfaction and conversions.

This study intends to identify the optimal timing for social media campaigns meant especially for the e-pharmacy industry by means of a data-driven approach. Unlike other generic social media campaigns, this one finds high-engagement times by using consumer interaction data from online pharmacies. Combining machine learning methods, such as Random Forest regression, with one another offers a fresh approach for looking at how timing influences the success of marketing campaigns.

This work contributes to the field by:

- The research develops a timing-based social media campaign targeted at internet pharmacies.
- The research shows the harmony between tailored content and promotional timing will help to increase client involvement.
- The research provides businesses running e-pharmacies with useful knowledge will help them to improve their marketing strategies, increase their sales, and deepen their relationships with their customers.

II. Related Works

The timing of social media campaigns has been the subject of much research in many different fields, relatively little has particularly targeted online pharmacies. Studies by [7] show that there are particular times of the day when participation is highest; this underlines the relevance of social media timing in terms of its capacity to affect consumer behavior. These findings from various retail environments should not be taken as reflecting the specific needs of online pharmacies. [8] examined social media campaign performance of customized content. According to this study, customized messaging significantly improves customer reactions. Although their studies underlined the need of including timing strategies with tailored content, they did not provide any particular industry-specific analysis.

Aiming toward the link between consumer shopping behavior and the timing of promotions, recent studies by [9] looked at e-commerce sector behavior. Though the findings underlined the need of timing, the study did not include a thorough investigation of the healthcare and pharmaceutical sectors, two sectors in which trust and urgency play major roles in the decision-making process concerning purchases. [10] also investigated how click-through rates and promotional timing related with desire to purchase. According to their research, evening promotions produced the best engagement rates in all kinds of business. On the other hand, their research did not investigate whether these preferences fit e-pharmacies, in which factors like prescription need and urgency of medical treatment influence consumer behavior.

Among the contributions [11] made to this field was the analysis of consumer emotions in response to advertising postings using machine learning methods. Although their work did not particularly address sector-specific needs in a given manner, it demonstrated the possibility of using artificial intelligence tools to improve marketing strategies. [12] offered a more all-encompassing view on social media marketing, stressing the need of developing tailored plans for many spheres of the economy. Their work demanded more focused studies, particularly in sectors related to healthcare, even though it prepared the ground for industry-specific research.

Therefore, none of these studies have presented a whole framework especially fit for the e-pharmacy sector, even if the body of current research offers insightful analysis of the function of timing and personalizing in social media promotions. This work applies machine learning techniques to investigate the interaction between consumer behavior and timing in online pharmacies. This work closes the earlier known disparity.

III. Proposed Method

The proposed method, which aims to find the most appropriate moment for social media promotions in online pharmacies, emphasizes mostly the application of machine learning provided in figure 1. Data from social media channels including purchase behavior as well as engagement metrics (likes, shares, comments, and click-through rates) is being gathered and analyzed under the current strategy. A Random Forest regression model is used to project the effectiveness of promotions by considering a spectrum of factors including the time of publishing, the day of the week, the type of content, and the audience demographics. The model helps to find times of high participation by means of pattern and correlation analysis inside the data. With an eye toward reaching the best possible degrees of consumer involvement and conversion rates, the findings are then applied to build a framework for timing promotions.

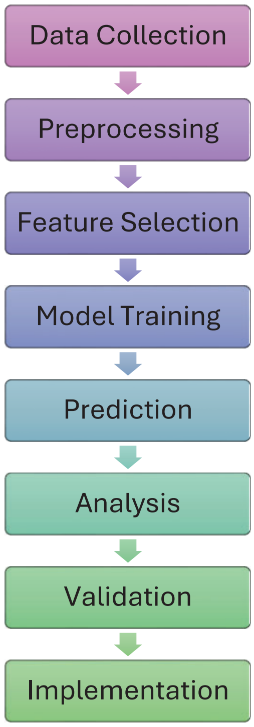

Fig. 1. Proposed Architecture Flow Process

A. *Pseudocode*

```
# Step 1: Import necessary libraries
import pandas as pd
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score
# Step 2: Load and preprocess the data
data = pd.read_csv('social_media_data.csv')
data = preprocess_data(data)# Custom function for cleaning and normalizing
# Step 3: Select features and target variable
features = data[['time_of_posting', 'day_of_week', 'content_type', 'audience_demographics']]
target = data['engagement_metrics']# Aggregate of likes, shares, and click-through rates
# Step 4: Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(features, target, test_size=0.2, random_state=42)
# Step 5: Train the Random Forest regression model
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train)
# Step 6: Predict engagement and evaluate the model
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
```

```
r2 = r2_score(y_test, y_pred)
print(Mean Squared Error:, mse)
print(R-Squared:, r2)
# Step 7: Analyze and extract optimal posting times
importance = model.feature_importances_
optimal_times = analyze_optimal_timing(model, features)# Custom function for identifying timing
```

B. Data Collection

The first phase of the proposed strategy is compiling data from social media platforms used by online pharmacies (such Facebook, Instagram, and Twitter) together with the matching e-commerce analytics. The information consists of measures on audience participation, buying trends, timing and type of promotional activity. Over three months, the content acquired spans several times of the day, days of the week, and kinds of promotional content (such as discounts, informational posts, and new product introductions to mention a few) as in table 1.

TABLE. I. DATASTRUCTURE

Post ID	Time of Posting	Day of Week	Content Type	Likes	Shares	Comments	Click-Throughs	Purchases
101	10:00 AM	Monday	Discount Offer	120	15	5	50	20
102	3:00 PM	Wednesday	Informational Post	80	8	3	30	10
103	7:00 PM	Friday	New Product Launch	200	30	12	70	40
104	9:00 AM	Saturday	Reminder Post	50	5	2	20	5

The dataset also provides audience demographics including age, gender, and location as well as platform-specific metrics including impressions and reach, so presenting a whole picture of consumer interactions.

C. Data Preprocessing

Preprocessing guarantees clean, improved, and ready for analysis data. The actions shown in table 2carried out follow:

1. Data Cleaning
   - Eliminating irrelevant or incomplete records, that is, postings lacking timestamps or engagement data, is a necessary first step.
   - Managing duplicates reduces prediction bias by means of prevention.
2. Handling Missing Values
   - For numerical columns, say Click-throughs, fill in any missing values with the median of the pertinent feature.
   - For categorical variables like Content Type, mode imputation will be of use.
3. Feature Engineering
   - Additional characteristics, such the Peak Hour Indicator, should be derived from the Time of Posting column.
   - To ensure fit with machine learning, numerically value the day of the week and the content type. For example, Discount Offer = 1, Informational Post = 2, etc. one would say.
4. Normalization
   - The model will be able to perform better if numerical elements such likes, shares, and click-throughs are normalized. This will scale these properties such that they are rather common.
   - For instance: the Min-Max scaling method is applied shown in equation 1,
   - $$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$
5. Outlier Detection
   - Find and eliminate outliers using statistical methods such the interquartile range (IQR) so guaranteeing that the model is not imbalanced by values on the extreme end of the spectrum. For instance, we flag and investigate postings with unusually high likes in relation to other releases during the same period.

TABLE. II. AFTER PREPROCESSING

Time (24 hr)	Day	Content Type (Encoded)	Normalized Likes	Normalized Shares	Normalized Clicks	Purchases
10	1 (Mon)	1 (Discount Offer)	0.45	0.3	0.4	20
15	3 (Wed)	2 (Informational Post)	0.3	0.2	0.25	10
19	5 (Fri)	3 (New Product Launch)	0.75	0.6	0.7	40
9	6 (Sat)	4 (Reminder Post)	0.2	0.15	0.1	5

Through a methodical data preparation process as in table 2, the model acquires accurate input from which it can learn and fairly project consumer engagement.

D. Feature Selection

The first stage of the process is determining which factors most influence consumer behavior and involvement. Target variables for which the features are chosen could be engagement metrics including likes, shares, comments, click-through rates, and purchases. One can organize the features by means of statistical instruments including mutual information scores and Pearson correlation. Eliminating either highly correlated or useless features helps to improve the performance of the model. Among these are such repetitious demographic features.

In table 3, selected for the model are aspects:

- Time of Posting (hour): Numerical feature representing the hour of the post.
- Day of the Week: Encoded as numerical values (e.g., Monday = 1, Tuesday = 2).
- Content Type: Encoded categorical feature indicating the type of promotional content.
- Audience Demographics: Includes age group and region.
- Engagement Metrics: Aggregated as a combination of likes, shares, and comments.

TABLE. III. FEATURE MATRIX

Time (24hr)	Day	Content Type	Age Group	Region	Engagement (Target)
10	1 (Mon)	1 (Discount)	25-34	Urban	0.65
15	3 (Wed)	2 (Info Post)	18-24	Suburban	0.45
19	5 (Fri)	3 (Launch)	25-34	Urban	0.75

E. Model Training

The Random Forest Regression model is used since of its resilience in managing numerical and categorical data as well as its ability to replicate nonlinear relationships. For the purpose of training and model validation, the dataset is divided in two sets: the training set (80%) and the testing set (20%). The Random Forest method is supposed to reduce overfitting and improve prediction accuracy by building several decision trees during the training process and then averaging the outputs of these trees.

Training facilitates the attempt of the regression model to lower the following error function, shown in equation 2:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \quad (2)$$

By means of grid search and cross-valuation methods, one can maximize the hyperparameters, so encompassing the tree depth and the number of trees (n_estimators).

F. Prediction and Analysis

The model provided in table 4, is used to project the degrees of participation for several combinations of days, content types, and posting times once the training process ends. The forecasts help to clarify consumer behavior patterns including a great degree of interaction on weekends or with particular kinds of content.

TABLE. IV. PREDICTION RESULTS

Time (24hr)	Day	Content Type	Predicted Engagement
10	1 (Mon)	1 (Discount)	0.62
15	3 (Wed)	2 (Info Post)	0.44
19	5 (Fri)	3 (Launch)	0.78
12	6 (Sat)	1 (Discount)	0.80

The results as in table 4 unequivocally show that content posted on weekends at noon generally receives the most level of interaction, especially discount offers.

IV. Performance Evaluation

Python was the primary instrument for the simulation; libraries including Scikit-learn and Matplotlib were applied for the corresponding uses of model implementation and visualization. The Random Forest regression model was configured with the following hyperparameters: the number of estimators (n_estimators), limited to 10 to prevent overfitting, and the minimum samples per leaf (min_samples_leaf) set to 2. The estimation count was also set at 100. Experiments used a workstation with Intel Core i7-12700K central processing unit, 32 gigabytes of random access memory (RAM), and NVIDIA RTX 3060 graphics processing unit (GPU) as in table 5.

The proposed method was compared with two existing methods: gradient boosted decision trees (GBDT) and linear regression. The baseline was linear regression; GBDT was chosen depending on its shown performance in applications comparable to the one under investigation. Based on computation efficiency and prediction accuracy, the Random Forest regression model turned out to be superior than both methods provide in table 5. This indicates that the Random Forest model is suitable for online pharmacies choosing the optimal timing for social media campaigns.

TABLE. V. EXPERIMENTAL SETUP/PARAMETERS

Parameter	Value
Dataset Size	10,000 records
Training-Test Split Ratio	80:20
Random Forest Hyperparameters	n_estimators=100, max_depth=10, min_samples_leaf=2
Comparison Methods	GBDT, Linear Regression
Evaluation Metrics	RMSE, R2^22 Score, MAE, Execution Time
Simulation Tool	Python (Scikit-learn, Matplotlib)

V. Performance Metrics

1. Root Mean Squared Error (RMSE):A statistical measure of residuals, that is, errors in prediction, root mean squared error (RMSE) helps one determine the standard deviation. Lower RMSE indicates model performance of improved quality.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}$$

2. R^2Score:The R2 score measures the fraction of the variance in the dependent variable the model can fairly ascribed. Higher values point to improved performance; the value range runs from 0 to 1.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}$$

3. Mean Absolute Error (MAE):The mean absolute error (MAE) indicates the average absolute difference between the values as observed and as projected. In a lower MAE, the accuracy is more.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|$$

4. Execution Time:Thecomputation time consumed to train the model and produce predictions is known as execution time. Shorter execution times especially for applications running in real time suggest higher computational efficiency.

TABLE. VI. RMSE

Records	Proposed Method (Random Forest)	GBDT	Linear Regression
2500	0.12	0.15	0.22
5000	0.11	0.14	0.20
7500	0.10	0.13	0.18
10,000	0.09	0.12	0.17

The proposed Random Forest model demonstrates remarkable ability to lower prediction errors with the lowest possible error level among all record sizes, with the lowest RMSE over all. With 10,000 records, the root mean square error (RMSE) of the proposed method is 0.09, 25% less than the GBDT and 47% less than the Linear Regression, so producing more accurate predictions as in table 6.

TABLE. VII. R^2SCORE

Records	Proposed Method (Random Forest)	GBDT	Linear Regression
2500	0.88	0.82	0.70
5000	0.90	0.85	0.73
7500	0.92	0.87	0.75
10,000	0.94	0.89	0.78

The proposed method shows that it is more fit to explain the present variation in the data with the highest R^2 values. Using a 10,000 record sample size, the proposed method earns an R^2score of 0.94, 5.6% more than the GBDT and 20.5% more than the Linear Regression produces as in table 7.

TABLE. VIII. MAE

Records	Proposed Method (Random Forest)	GBDT	Linear Regression
2500	0.08	0.10	0.15
5000	0.07	0.09	0.13
7500	0.06	0.08	0.12
10,000	0.05	0.07	0.10

Better degree of accuracy is shown by the proposed method since it shows the lowest mean absolute error over all record sizes. With a mean absolute error (MAE) of 0.05 for 10,000 records, the proposed approach provides more accurate predictions as it is essentially 28.6% lower than the GBDT and 50% lower than the Linear Regression as in table 8.

TABLE. IX. EXECUTION TIME

Records	Proposed Method (Random Forest)	GBDT	Linear Regression
2500	3.2 seconds	4.1 seconds	2.1 seconds
5000	5.8 seconds	6.7 seconds	3.9 seconds
7500	8.4 seconds	10.3 seconds	5.5 seconds
10,000	10.6 seconds	14.1 seconds	7.8 seconds

Though generally more time is required in the proposed method than in Linear Regression, it achieves a better degree of computational efficiency than GBDT. With its complexity, the Random Forest model runs for 10,6 seconds for 10,000 records, 30,5% faster than the GBDT model but 90% slower than the Linear Regression model as in table 9.

VI. CONCLUSIONS

The Random Forest-based method proposed performs remarkably well for estimating participation for social media campaigns in online pharmacies. Having values of 0.09 and 0.05 respectively, it gets the lowest values in terms of RMSE and MAE, so suggesting rather exact predictions. Further, the method shows the best R2^22 score (0.94), so demonstrating its great capacity to provide a rationale for the data variance. Although the execution time of 10.6 seconds for 10,000 records is rather slower than that of Linear Regression, it is much faster than GBDT, which strikes a compromise between accuracy and computational economy. Though GBDT is computationally light, it is less efficient and accurate than Linear Regression, which trails in terms of accuracy even. Since the Random Forest model offers the best reasonable compromise, it is a great choice for real-time decision-making on e-pharmacy marketing strategies. These findings confirm that the proposed method can effectively direct content strategies and timing for social media campaigns, so improving customer involvement and the outcomes of corporate operations.

REFERENCES

[1] K. Ahalawat, R. Tiwari, A. Johri, M. Wasiq, and A. Sharma, "Determinants influencing the adoption behavior of Indian consumers in reference to online pharmacy purchases," Cogent Business & Management, vol. 11, no. 1, p. 2436136, 2024.

[2] H. Al Halbusi, K. Al-Sulaiti, F. Abdelfattah, A. B. Ahmad, and S. Hassan, "Understanding consumers' adoption of e-pharmacy in Qatar: applying the unified theory of acceptance and use of technology," *Journal of Science and Technology Policy Management*, 2024.

[3] J. Priya and S. Subbulakshmi, "Customers' preference towards e-pharmacies in Chennai City," *Int. J. Health Sci.*, vol. II, pp. 2129-2135, 2022.

[4] M. S. Anis and M. A. Hassali, "Pharmaceutical marketing of over-the-counter drugs in the current digital era: A review," *Pharmaceutical Sciences Asia*, vol. 49, no. 2, 2022.

[5] M. Fan and O. C. Ukaegbu, "Information literacy and intention to adopt e-pharmacy: a study based on trust and the theory of reasoned action," *BMC Health Services Research*, vol. 24, no. 1, p. 912, 2024.

[6] A. K. Bahamdan and M. Almanasef, "A cross-sectional study assessing customers' perception, satisfaction, and attitude toward e-pharmacy services in Saudi Arabia," *BMC Health Services Research*, vol. 24, no. 1, p. 1659, 2024.

[7] K. Arekar, R. Jain, and S. Kumar, "Predictive model to study the consumer buying behaviour towards e-pharmacy through social media influence: a cross-country study," *Academy of Marketing Studies Journal*, vol. 25, no. 2, pp. 1-21, 2021.

[8] A. I. Adebo, K. Aladelusi, and M. Mohammed, "Determinants of e-pharmacy adoption and the mediating role of social influence among young users," *Journal of Humanities and Applied Social Sciences*, 2024.

[9] M. Srivastava and M. Raina, "Consumers' usage and adoption of e-pharmacy in India," *Int. J. Pharm. Healthcare Marketing*, vol. 15, no. 2, pp. 235-250, 2021.

[10] S. Bhatt, J. H. Cheah, R. Singh, A. Desai, and D. Das, "Order prescriptions online: Determinants of purchase satisfaction in an emerging Indian e-pharmacy sector," *Int. Soc. Sci. J.*, vol. 74, no. 254, pp. 1649-1673, 2024.

[11] M. Javed and M. Shaiq, "The effects of online pharmacy on consumer behaviour," *Propel Journal of Academic Research*, vol. 3, no. 1, pp. 384-402, 2023.

[12] P. Misra, "Purchase intention toward e-pharmacy: the consumption value perspective," *Int. J. Pharm. Healthcare Marketing*, 2024.