Journal of
Information Systems Engineering
and Business Intelligence

Vol.11, No.1, February 2025
Available online at: http://e-journal.unair.ac.id/index.php/JISEBI

Improving Café Reputation: Machine Learning Analytics for Predicting Customer Engagement on Google Maps

Siti Anisah [1)*], Meditya Wasesa [2)]

[1) 2)] *School of Business and Management, Institut Teknologi Bandung, Bandung, Indonesia*
Jl. Ganesha No.10, Bandung
[1)]siti_anisah@sbm-itb.ac.id, [2)]meditya.wasesa@itb.ac.id

Abstract

Background: Online reviews is a powerful tool in shaping customer decisions, as they significantly influence a business's reputation and the ability to attract new customer. Given the growing reliance on digital platforms, understanding engagement levels is crucial for business that want to enhance online presence. By analyzing these customer activities, business owners can leverage Machine Learning (ML) analytics to predict engagement on Google Maps reviews.
Objective: This study aimed to develop the most suitable ML model in order to predict customer engagement levels in café business on Google Maps, and determine the online review features that have the greatest impact on engagement. Additionally, the analysis aimed to provide actionable recommendations to help business owners improve online reputation and engagement strategies.
Method: A total of 5,626 online reviews data were collected using web scraping methods during the analysis. The data was then preprocessed by extracting major review features, calculating engagement levels, and addressing class imbalance with SMOTE method. In the study, K-Means clustering was used to segment engagement levels, while sentiment analysis through VADER Lexicon was applied to measure sentiment content. Various ML models were trained and validated using a 10-fold cross-validation method. Finally, Analysis was conducted using Spearman's correlation to identify relationships among features derived from the best-performing ML models.
Results: The result of the analysis showed that Random Forest model achieved the highest accuracy and PR AUC in predicting engagement levels. The four most influential factors were review length (16.23%), photos (15.57%), total rating (12.35%), and author review count (10.19%). Spearman's correlation analysis showed a positive relationship among review length, photos, and author review count, signifying the combined impact on engagement levels.
Conclusion: This study described the effectiveness of Random Forest model in predicting engagement levels in Google Maps reviews. Specifically, the model identified review length, photos, total rating, and author review count as the key factors influencing engagement. These results would provide valuable guidance for business owners that desire to improve customer engagement and online reputation. Building on this, future studies should explore larger datasets, integrate additional features, and examine how the engagement contribute to long-term customer retention.

Keywords: Online Reputation Management, Customer Engagement, Behavior, Machine Learning, Google Maps Review, Predictive Analytics

Article history: Received 24 September 2024, first decision 2 January 2025, accepted 17 March 2025, available online 28 March 2025

I. INTRODUCTION

The rapid expansion of online platforms is transforming the way business engage with customer, particularly for small enterprises. This is because consumer choices are increasingly influenced by online reviews. According to the report published by Statista in December 2022, this trend signified that the share of online reviews on Google grew from 67% in 2020 to 71% in 2021. This increase shows the growing impact of online reviews on customer perceptions and decision-making [1]. Given this trend, analyzing customer engagement on online review platforms is crucial for understanding how interactions and responses affect business reputation and customer relationships.

Online reviews serve multiple purposes, providing valuable data such as ratings, textual feedback, and visual content. On digital platforms, customer engagement includes direct interactions between customer and business that significantly influence online reputation and attract new customer [2], [3]. In a conventional manner, engagement has been associated with business-generated content, but this study adopts a broader perspective. Specifically, the study views engagement as a dynamic process in which both customer and business owners actively contribute to shaping

* Corresponding author

ISSN 2443-2555 (online) 2598-6333 (print)
doi: http://dx.doi.org/10.20473/jisebi.11.1.91-102

a shared narrative. This nuanced method provides a deeper understanding of how reviews and responses collectively influence customer engagement and business reputation.

Customer engagement is critical to create reliable and satisfying online reviews [4]. Customer engagement behavior on social commerce platforms is primarily controlled by social interactions, technological components including interactivity and system quality, as well as motivational factors comprising of hedonic and utilitarian incentives, and perceived value [5]. Following this discussion, major factors such as vividness, interactivity, content style, and temporal factors are essential to measure as well as increase engagement on social media platforms [6]. Interactivity is defined along six dimensions, namely communication direction, temporal flexibility, sense of place, degree of control, responsiveness, and perceived communication purpose [7], which impact user behaviors and attitudes [8]. Vividness in social engagement includes colorful language, elaborate examples, and imagery that improves memorability [9]. Content type signifies the properties of the types and amounts of content shared on social networks concerning user interaction or engagement [10]. At the same time, temporal factors indicate the timing of content publishing and how much it affects online engagements [6].

Due to the unstructured aspect of online reviews, sentiment analysis of a review is essential to obtain substantial knowledge, which can then be used as a predictor variable (content type section) to predict the level of engagement. Using VADER (Valence Aware Dictionary and sEntiment Reasoner), a binary sentiment analysis tool that uses a dictionary method and contains 7,518 unigrams, including punctuation, slang, acronyms, acronyms, and emoticons [11], the library measures meaning in text and generates a probability score for each sentence to classify it as positive, negative, or neutral.

Previous studies often relied on conventional statistical methods such as regression models, which while useful for identifying linear relationships, failed to capture the complex, non-linear patterns commonly found in text data, such as customer comments. Alternative methods to interpreting data started around the time of Machine Learning (ML). Advanced transformer-based architectures such as BERT showed the potential to better understand unstructured text data. However, these methods often required heavy computational power and extensive data input, posing a significant challenge for most industries, particularly Small and Medium-sized Businesses (SMBs). Recent success of large-scale models such as BERT came at a high cost, both in terms of specialized hardware and technical expertise for implementation, barriers that often-disqualified SMBs [12].

Understanding engagement prediction has become increasingly important in digital landscape of nowadays, where business analyze customer interactions on various platforms to extract valuable insights. Traditional methods, such as regression analysis [13] and hierarchical models [14], have been widely used to measure user engagement on Facebook, particularly in industries such as fashion and maritime. More recently, advanced algorithms, including transformer-based models such as BERT and hybrid methods that combine Structural Equation Modeling (SEM) with Artificial Neural Networks (ANN), have achieved traction [15]. Despite these methods offering improved predictive capabilities, the complexity often makes the models less interpretable for businesses pursuing clear, actionable insights.

Several studies have described the growing adoption of ML in engagement prediction, particularly through hybrid methods such as SEM-ANN and Partial Least Squares (PLS) combined with Necessary Condition Analysis (NCA). These methods have been applied to assess customer engagement in mobile applications and food delivery platforms [16]. Despite achieving high accuracy, the models require substantial computational resources and expertise, which may not be practical for small and medium-sized enterprises, such as cafes. Given this limitation, simpler algorithms such as Random Forests and Decision Trees remain valuable due to their interpretability and efficiency. For instance, a study [2] proved that Random Forests effectively model engagement levels in e-commerce on Twitter platform. The model assisted businesses to identify key features influencing engagement while maintaining relatively low computational costs. Similarly, traditional ML models such as Random Forest, K-Nearest Neighbors (KNN), and Decision Trees offer greater interpretability and scalability. This caused the models to be effective in analyzing datasets of appropriate dimensions, particularly when evaluating online reviews from digital platforms. In resource-constrained environments, the complexity of advanced models presents inherent challenges, indicating the need for studies that balance predictive power alongside practical usability.

Based on the above description, this study analyzed Google Maps reviews to predict engagement levels using ML models, including Random Forest, KNN, Decision Tree, Support Vector Machine (SVM), and Multinomial Logistic Regression. K-means clustering was applied to categorize engagement levels, while Spearman correlation analysis examined the relationship between important features. The findings provided actionable insights for business owners to increase engagement and manage online reputation, while balancing practical utility with theoretical depth.

Building on previous findings, interpretable ML methods were used to analyze and predict engagement levels based on online reviews. The analysis focused on three main objectives, including (1) to develop the most suitable ML model for predicting customer engagement levels in café businesses on Google Maps, (2) to identify the online review features that had the greatest impact on engagement, and (3) to provide actionable insights to help business owners improve online reputation and engagement strategies.

II. METHOD

A. Research Framework

The study framework (Fig. 1) for analyzing Google Maps review data started with the preparation of relevant datasets. The reviews were then processed using K-Means clustering to segment engagement levels and sentiment analysis with the VADER Lexicon. During model development, five machine learning algorithms Random Forest, KNN, SVM, Decision Tree, and Multinomial Logistic Regression were trained and evaluated using 10-fold cross-validation. The best-performing model was selected based on key metrics, followed by a feature importance analysis to assess each feature's contribution to engagement levels. Finally, Spearman correlation analysis was conducted on the top-ranked features to examine their relationships, offering deeper insights into their interactions.

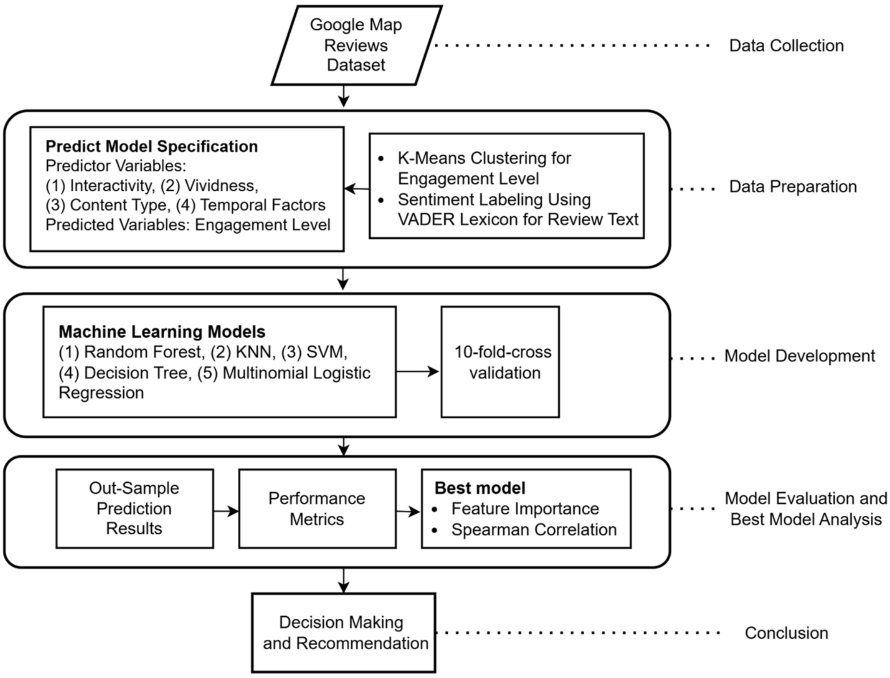

Fig. 1 Study Framework

B. Data Collection

This analysis used web scraping methods to extract customer reviews from Google Maps using Outscraper as the data extraction tool. The process targeted 21 cafés that were still operating in the Dago area of Bandung. Moreover, the location was selected because it was a strategic and well-known area in the city. The collected reviews spanned from January 1, 2023, to July 27, 2024, totaling 5,626 entries, and then processed using Python.

C. Data Preparation

At this stage, the variable features available on Google Maps Reviews were identified based on the online questionnaire used in the platform (Fig. 2). When customer wrote café reviews on Google Maps, a pop-up questionnaire appeared, allowing user to provide various details. These included a total rating which was mandatory for posting a review, ratings for food, service, and atmosphere, review text, uploaded photos and videos, meal service type, meal category, price per person, recommended dishes, as well as any additional information the reviewer wished to add. Two measurable variables were created to analyze review text, namely review length and review sentiment. Additionally, the number of reviews authored by each reviewer was included as a feature variable. These feature variables were examined as predictor variables to determine the impact on the predicted variable, and engagement level derived from grouping engagement rate count data. An example of a café review available on Google Maps was shown in Fig. 3.

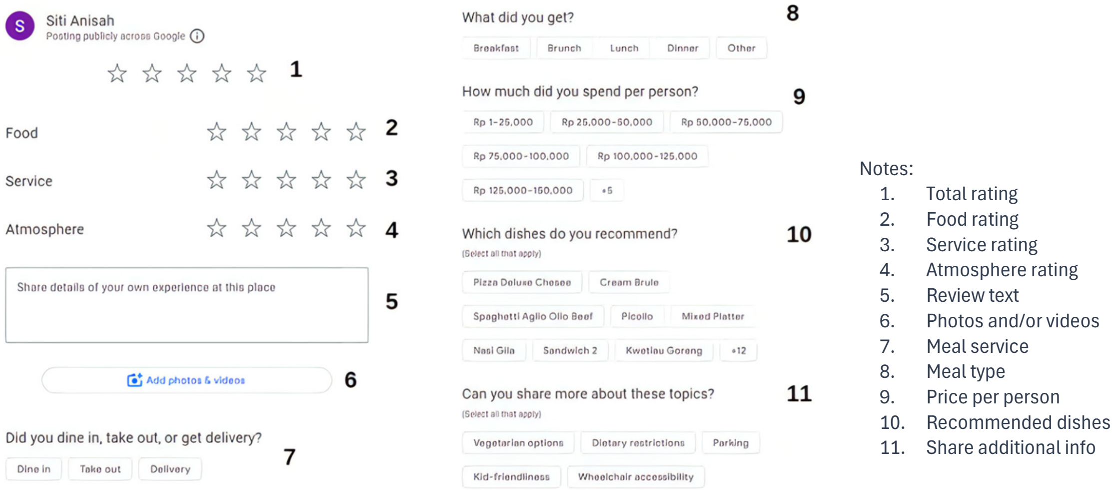

Fig. 2 Google Maps review online questionnaire

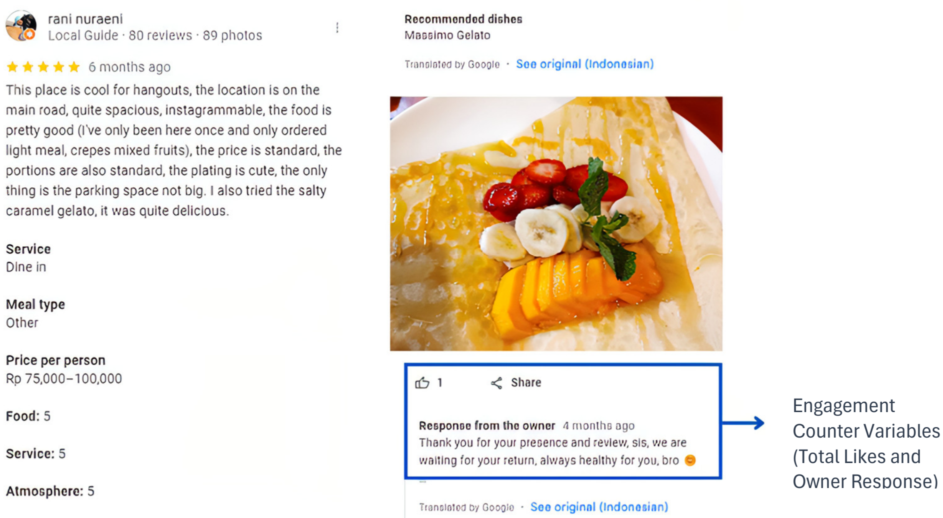

Fig. 3 Example of an online review of a café

Engagement rates for each café review on Google Maps were calculated by dividing the total number of engagements by the total number of reviews collected during the data collection period, then multiplying the result by 100 to express it as a percentage. The total engagement for each review was determined by summing the number of likes and responses from café owners. In a previous study [2], engagement levels for three e-commerce Twitter posts were categorized into four groups, Low, Good, High, and Very High. Using this as a benchmark, engagement rates of Google Maps reviews were classified into four distinct levels, namely Low, Moderate, High, and Very High. These classifications served as the predicted variables in the analysis. A more detailed explanation of the predictor variables was shown in Table 1.

To address class imbalance in the dataset, SMOTE (Synthetic Minority Oversampling Technique) was applied. This method generated synthetic samples to ensure the model had sufficient data for training on minority classes. Additionally, the dataset was split into training (80%) and testing (20%) subsets to enable a strong evaluation of ML models.

TABLE 1
PREDICTOR VARIABLES

Predictor Variables	Definition	Data Type	Data Code	n	%
a. Interactivity					
Recommended dishes	Mentioning a particular dish that the reviewer thinks other diners should try	Categorical	Yes	730	12.98%
			No	4,896	87.02%
Share additional info	Share additional review information such as the availability of vegetarian options, parking, wheelchair accessibility, and kid-friendliness	Categorical	Yes	308	5.47%
			No	5,318	94.53%
b. Vividness					
Photos	Use photos in the review	Categorical	Yes	1,442	25.63%
			No	4,184	74.37%
Author review count	Number of reviews that the review author has created	Numerical	-	5,626	100.00%
Total rating	Assessment of the total experience	Numerical	-	5,626	100.00%
Food rating	Assessment of food quality	Numerical	-	3,791	67.38%
Service rating	Assessment for service provided	Numerical	-	3,855	68.52%
Atmosphere rating	Assessment of the atmosphere of the place	Numerical	-	3,855	68.52%
c. Content Type					
Review length	Length of reviews written online	Numerical	-	5,626	100.00%
Review sentiment	Customer review sentiment	Categorical	Positive	2,563	45.56%
			Neutral	2,848	50.62%
			Negative	215	3.82%
Meal service	Fill in the type of food service	Categorical	Yes	282	5.01%
			No	5,344	94.99%
Meal type	Fill in the type of food that the reviewer ordered	Categorical	Yes	2,118	37.65%
			No	3,508	62.35%
Price per person	Fill in the amount of money spent per person	Categorical	Yes	2,064	36.69%
			No	3,562	63.31%
d. Temporal factor					
Day of review	Review during weekend or weekday	Categorical	Weekday	3,751	66.67%
			Weekend	1,875	33.33%

Notes: Data coding details for Recommended dishes, Share additional info, Photos, Meal service, Meal type, and Price per person variables: Yes =1, No = 0; Positive = 1, Neutral = 2, Negative = 3 for Review sentiment variable; Weekday = 1, Weekend = 0 for Day of review variable

D. Model Development

1) Random Forest

Random Forest efficiently categorized large volumes of data and applied the data to various classification and regression problems. This method was built on generating multiple classification trees to be used during the process. When these trees predicted an outcome, the model averaged the predictions and selected the value with the highest number of votes as the final result [17]. The output of Random Forest represented average or weighted result of all individual trees. Specifically, the ensemble score of Random Forest was calculated as follows [17]:

$$\hat{Y}_i = mode_{n=1\ldots N_{tress}} \hat{Y}_n \quad (1)$$

where:

- $\hat{Y}_i$ represented the ensemble prediction,
- $mode$ signified the most frequently occurring result among the trees,
- N_{tress} was the total number of trees in the forest,
- $\hat{Y}_n$ represented the prediction result from an individual tree.

2) K-Nearest Neighbors (KNN)

KNN was a classification algorithm that identified K closest neighbors of a given test data point based on a distance function. This algorithm classified a test sample by analyzing the classes of the nearest training samples. Concerning feature matrix X, KNN determined the conditional distribution of class label y and assigned the test sample to the class with the highest number of neighboring points among the selected K neighbors. The formula was expressed as follows [18]:

$$P(y = j | X = x) = \frac{1}{K} \sum_{i \in A_k} I(y^{(i)} = j) \quad (2)$$

where:

- $P(y = j|X = x)$ was the probability that a new data point x belonged to class j (in this case, j represented low, moderate, high, or very high).
- K represented the quantity of nearest neighbors that was considered in KNN.
- A_k was the set of KNN of data point x.
- $I(y^{(i)} = j)$ was an indicator function that took value 1 when the i neighbor had class j and 0 otherwise.

3) Support Vector Machine

SVM included classification and regression methods that relied on training data to make predictions for new or computed data [19]. The kernel function transformed the original input space into a higher-dimensional space through a nonlinear extension. Relating to the discussion, several types of kernel functions existed, including linear kernels, polynomial kernels, and radial basis function (RBF) kernels [19]. In this study, the radial basis function kernel was used, and the formula was expressed as follows [19]:

$$K(x_{i,}\ x) = \exp\left(-\gamma\ \|x - x'\|)^2\right) \quad (3)$$

4) Decision Tree

Decision tree provided an efficient method for generating classifiers from data. This method resembled a tree-like arrangement of nodes, each designed to make decisions concerning class membership or numerical target estimates. Each node represented a partition rule for a specific attribute, and new nodes were continuously created until a predefined termination criterion was met. In addition, the class label prediction was determined based on the majority of samples that reached a particular leaf node during construction of the tree [20]. To assess the effectiveness of a decision tree at each node, Gini index was commonly used, as it estimated the misclassification rate. The formula for calculating Gini index was expressed as follows [21]:

$$i(t) = 1 - \sum_{j=1}^{K} P_j(t)^2 \quad (4)$$

where:

- $i(t)$ represented the estimated probability of misclassification at node t,
- $P_j(t)$ signified the probability of class j at node t,
- K was the number of classes (in this case, K=4, corresponding to the classes, namely low, moderate, high, and very high).

5) Multinomial Logistic Regression

Logistic regression was a widely recognized multivariate statistical method used to predict the probability of an event occurring based on the relationship between a binary dependent variable and multiple independent variables [22]. When the predictor variables included more than two categories, multinomial logistic regression was applied. For a dependent variable YY with kk categories, the probability of each category (k=1,2,...,k) was determined using the following formula [23]:

$$P(Y = k) = \frac{e^{\beta_{0k}+\beta_{1k}X_1+\beta_{2k}X_2+\cdots+\beta_{pk}X_p}}{1+\sum_{j=1}^{k-1} e^{\beta_{0j}+\beta_{1j}X_1+\beta_{2j}X_2+\cdots+\beta_{pj}X_p}} \quad (5)$$

where P(Y=k) was the probability that the dependent variable Y belongs to category k. β_{ok}, β_{1k},…,β_{pk} were the coefficients of the model estimated by training, and X_1, X_2,…, X_p represented independent variables.

E. Model Evaluation

The performance of the prediction model was evaluated using accuracy, precision, recall, and F1-score, calculated using the following formulas [24]:

1) Accuracy

Accuracy quantified the proportion of correct predictions out of all predictions made by the model. The formula for calculating accuracy during the process was shown in Equation (6).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

2) Precision

Precision was used to evaluate the accuracy of a classifier and indicate how accurately the model made predictions. The formula to calculate precision in this study was signified in Equation (7).

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

3) Recall

Recall measured the ability of a model to make accurate predictions. The formula to calculate recall in the analysis was shown in Equation (8).

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

4) F1-score

Precision and recall alone might not have been sufficient to evaluate a model's performance comprehensively. To address this, F1-score was used as a performance metric that combined precision and recall into a single harmonic mean. A higher F1-score indicated better model performance by balancing both precision and recall. The formula to calculate F1-score was presented in Equation (9).

$$F1\ score = 2 \text{ x} \frac{Precision \text{ x } Recall}{Precision+Recall} \quad (9)$$

The abbreviations TP, TN, FP, and FN referred to a true positive, true negative, false positive, and false negative, respectively.

F. Spearman Correlation

Correlation analysis typically included computing a specific correlation statistic from a sample to assess the relationship between two variables. Spearman's rank correlation coefficient, a non-parametric statistic, was widely used to estimate relationships between variables, particularly in cases where data did not follow a normal distribution. This method was also applied to analyze correlations between major features of top-performing ML models [25]. The correlation coefficient (r_s) measured the strength of the correlation, producing a value in the range of $-1 < r_s < 1$. Higher absolute value of r_s, showed a stronger relationship between the two variables. During the process, a positive number correlated with higher values of one variable. Meanwhile, a negative value correlated with lower values of the other variables [26].

III. Results

Engagement rate data calculated earlier was analyzed using K-Means clustering to classify different engagement levels. In this study, engagement level was the predicted variable and was categorized into four distinct clusters with specific value ranges (Table 2). These clusters included Low Engagement (centroid: 0.027874), representing minimal interaction as well as the largest proportion of data points, and Moderate Engagement (centroid: 0.999646), indicating slightly higher interaction compared to the low engagement cluster. Other clusters comprised High Engagement (centroid: 3.029167), representing groups with significant interaction levels, and Very High Engagement (centroid: 6.886896), consisting of a small number of entities showing exceptionally high interaction. This classification facilitated an effective categorization of the data and enabled a clearer assessment of value distribution across the clusters. Additionally, Silhouette coefficient for K-Means clustering algorithm was calculated as 0.909856, signifying a high level of clustering accuracy.

TABLE 2
CLUSTER ENGAGEMENT LEVEL DATA RANGE

	Cluster	Lower Bound	Upper Bound	n	%
0	Low	0.000000	0.510204	5,305	94.29%
1	Moderate	0.531915	1.904762	260	4.62%
2	High	2.127660	4.761905	54	0.96%
3	Very High	5.172414	10.344828	7	0.12%

After clustering engagement level data, the analysis pre-processed the data first before labeling the sentiment of a review text using VADER. The sentiment of a review was then classified as positive, neutral, or negative, which was used as one of the predictor variables. Table 3 showed an overview of the sentiment distribution in the dataset.

TABLE 3
DISTRIBUTION OF SENTIMENT DATA LABELING

Sentiment	n	%
Positive	2,563	45.56%
Neutral	2,848	50.62%
Negative	215	3.82%

Based on the evaluation results for 10-fold cross-validation (Table 4), Random Forest showed the most consistent and highest average accuracy, achieving a score of 0.960273. Additionally, the model recorded the highest Precision-Recall Area Under the Curve (PR AUC) value of 0.984810, signifying its strong predictive performance, particularly in scenarios where accuracy and hit rate were critical. These results showed that Random Forest effectively balanced the trade-off between true positive rate and false positive rate, making it the most reliable model.

TABLE 4
RESULTS FOR 10-FOLD CROSS-VALIDATION

	Random Forest	KNN	SVM	Decision Tree	Multinomial Logistic Regression
Fold 1	0.962300	0.937323	0.844958	0.943921	0.594722
Fold 2	0.959943	0.934025	0.839303	0.943450	0.590481
Fold 3	0.956173	0.935910	0.848256	0.932611	0.601320
Fold 4	0.959001	0.939208	0.849670	0.932139	0.603676
Fold 5	0.962300	0.940151	0.841188	0.939680	0.596136
Fold 6	0.964656	0.937795	0.840716	0.940151	0.610273
Fold 7	0.959472	0.937323	0.837418	0.933553	0.581998
Fold 8	0.960415	0.930254	0.847314	0.935910	0.584354
Fold 9	0.959472	0.937323	0.852026	0.933082	0.595193
Fold 10	0.959001	0.929312	0.833176	0.936381	0.593308
Mean 10-fold cross-validation	0.960273	0.935862	0.843403	0.937088	0.595146
Mean PR AUC	0.984810	0.963372	0.897284	0.942262	0.662247
Model development time (seconds)	3.151151	0.038809	50.348120	0.163105	0.586107

KNN model ranked second in accuracy performance, achieving an average score of 0.935862 in 10-fold cross-validation. PR AUC of the model was also relatively high at 0.963372, closely reaching the performance of Random Forest. The results indicated that KNN was effective in making highly accurate predictions. KNN had the shortest development time, requiring only 0.038809 seconds, making it an appealing choice in terms of computational efficiency. Despite the speed of KNN, its predictive performance remained slightly lower compared to Random Forest. Meanwhile, SVM, Decision Tree, and Multinomial Logistic Regression models showed weaker performance compared to both Random Forest and KNN.

TABLE 5
PREDICTION ACCURACY OF MACHINE LEARNING MODELS

	Engagement Level				Macro-Average	Accuracy
	Low	Moderate	High	Very High		
Random Forest						
Precision	0.97	0.94	0.93	1.00	0.96	
Recall	0.98	0.92	0.93	1.00	0.96	
F1 Score	0.97	0.93	0.93	1.00	0.96	0.96
KNN						
Precision	0.93	0.94	0.86	1.00	0.93	
Recall	0.97	0.82	0.92	1.00	0.93	
F1 Score	0.95	0.87	0.89	1.00	0.93	0.93
SVM						
Precision	0.79	0.83	0.75	0.94	0.83	
Recall	0.86	0.68	0.76	1.00	0.83	
F1 Score	0.82	0.75	0.76	0.97	0.82	0.83
Decision Tree						
Precision	0.95	0.89	0.88	1.00	0.93	
Recall	0.96	0.88	0.88	1.00	0.93	
F1 Score	0.96	0.88	0.88	1.00	0.93	0.93
Multinomial Logistic Regression						
Precision	0.49	0.60	0.54	0.71	0.58	
Recall	0.47	0.61	0.45	0.82	0.59	
F1 Score	0.48	0.61	0.49	0.76	0.59	0.59

Table 5 showed that Random Forest model presented high precision, recall, and F1-score values across all engagement levels, achieving a near-perfect macro-average score. These results signified that Random Forest could accurately classify user into the correct engagement levels, effectively handling both majority and minority classes. Consequently, Multinomial Logistic Regression showed the weakest performance among all models. Low precision, recall, and F1-score values of the model, particularly at low engagement levels indicated difficulty in accurately classifying user with minimal engagement. As KNN and Decision Tree models performed well, accuracy remained lower compared to Random Forest. Therefore, Random Forest was the best model for predicting engagement levels in café reviews on Google Maps platform. The ability of the model to capture complex patterns in the data allowed it to generate highly accurate and reliable predictions.

TABLE 6
FEATURES IMPORTANCE ON RANDOM FOREST MODEL

Feature	Random Forest Importance (%)
Review length	16.226901
Photos	15.573961
Total rating	12.353983
Author review count	10.191231
Day of review	7.899680
Review sentiment	7.785192
Meal service	5.724406
Rating food	4.833545
Meal type	4.772723
Price per person	4.658235
Share additional info	4.277204
Rating service	2.166330
Rating atmosphere	1.999964
Recommended dishes	1.536645

Based on Table 6, the four most influential features in predicting engagement levels using Random Forest model were review length (16.23%), number of photos (15.57%), total rating (12.35%), and author reviews count (10.19%). These results showed that the features had a strong correlation with user activity on Google Maps review platform, particularly concerning café reviews.

TABLE 7
SPEARMAN CORRELATION MATRIX OF FEATURES

	review_length	photos	total_rating	author_review_count
review_length	1.00	0.42	-0.20	0.33
Photos	0.42	1.00	-0.01	0.36
total_rating	-0.20	-0.01	1.00	-0.05
author_review_count	0.33	0.36	-0.05	1.00

Spearman correlation matrix in Table 7 showed the relationships among the four most influential features, namely review length, number of photos, total rating, and author review count. The analysis signified a moderate positive correlation between review length, number of photos, and author review count, showing that user who wrote longer reviews were also more expected to include photos and leave reviews more frequently. However, no significant correlation was observed between total rating and other variables. This result indicated that the rating assigned by user was not influenced by factors such as review length, number of photos, or author review count.

Using the best-performing ML model, Random Forest, engagement level of the provided feature data was predicted. As shown in Table 8, the prediction results classified the data at the moderate engagement level.

IV. DISCUSSION

The results showed the effectiveness of ML models in predicting engagement levels for café reviews on Google Maps platform. The analysis signified that Random Forest outperformed KNN, SVM, Decision Tree, and Multinomial Logistic Regression in terms of accuracy as well as reliability. Relating to this discussion, the strength of Random Forest allowed it to handle non-linear data effectively, a capability that had been validated by studies in similar fields [2], [27], [28]. As KNN showed comparable performance with a significantly shorter processing time, its accuracy trade-off made the model less suitable for high-stakes prediction scenarios. This observation supported the results of [29], which also concluded that Random Forest was superior in both accuracy and reliability.

TABLE 8
SAMPLE TEST DATA FOR PREDICTION

Feature	Data Input
Review length	100
Photos	1
Total rating	5
Author review count	10
Day of review	1
Review sentiment	1
Meal service	1
Rating food	5
Meal type	0
Price per person	1
Share additional info	1
Rating service	4
Rating atmosphere	4
Recommended dishes	1

Review length, number of photos, total ratings, and author review count were the major factors influencing engagement levels. Longer reviews and photo uploads probably reflected greater customer engagement and emotional investment, as the factors required more effort to create. This showed that customer who provided more detailed reviews, through extended text or visual documentation, were more immersed in café experience and motivated to share personal thoughts. The results of [30] further supported this idea, indicating that images in online reviews improved perceived usefulness and also served as tangible reference points, reducing customer uncertainty. The confirmation made photos an essential element in capturing audience attention. Additionally, total ratings influenced interactions between reviewers and business owners. Higher ratings tended to attract more engagement, while lower ratings often prompted café owners to take corrective actions and strengthen customer relationships. High review count from an author contributed to trust and credibility, as frequent reviewers were perceived as more experienced and reliable. This encouraged greater engagement from both other users and business owners. Finally, this dynamic played a crucial role in shaping online reputation of café and improving meaningful interactions on the review platform.

Spearman correlation analysis showed a moderate positive correlation between review length, number of photos, and author review count. The result signified that user who wrote longer reviews were also more expected to include photos and post reviews frequently. However, no significant correlation was found between total rating and these three variables, indicating that the rating by user was not necessarily influenced by review length, number of photos, or author review count. These findings supported those of [31], which examined the relationship between extreme ratings and review helpfulness. The study found that longer reviews and those containing photos were perceived as more helpful, yet the reviews did not necessarily lead to higher ratings.

This study provided practical implications for cafés owners, proposing the owners should implement marketing strategies to encourage detailed, photo-rich reviews through initiatives such as loyalty programs, contests, and discounts. Additionally, maintaining a responsive online presence by responding to customer reviews improved positive feedback, increasing customer interaction and satisfaction. Cafés that adopted these strategies improved engagement, strengthened online reputation and also achieved a competitive edge in the industry.

This study was limited to Dago area in Bandung during the analysis, focusing on context-specific insights relevant to businesses operating in the region. However, the limitation also restricted the generalizability of the results. A broader method could help address the limitations of this study and provide more comprehensive insights. In addition to this view, ML analytics could be further explored to support sustainable SMEs in community-based tourism areas[32]. Cross-sector studies on online review behavior across different customer segments [33], [34]could offer valuable insights for hospitality industry (HoReCa) as well as policymakers and community developers. Furthermore, future studies could examine how customer engagement levels influence repeat visits, offering practical guidance for businesses pursuing to improve customer retention and long-term success.

V. CONCLUSIONS

In conclusion, this study showed the effectiveness of ML, particularly Random Forest model, in predicting engagement levels based on Google Maps reviews. The results identified review length, photos, total ratings, and author review as the most significant contributors to engagement, offering actionable insights for café owners. Moreover, businesses could refine strategies to increase engagement and strengthen online reputation by understanding what motivates customers. Cafe owners could also incentivize customers to leave longer, more detailed reviews by offering bonuses or loyalty points, which could be redeemed upon submitting an original review on Google

Maps. Similarly, discounts, reward points, or photo contests with prizes could urge users to upload images, further enriching the review content. Beyond these strategies, creating an aesthetically pleasing, clean, and visually engaging café environment could naturally inspire customers to capture as well as share experiences online. Additionally, café owners were encouraged to maintain an active online presence by responding to customer reviews in a timely and personalized manner. Engaging effectively with customer feedback could improve satisfaction and also increase stronger relationships, eventually advancing reputation as well as competitive advantage of café. Future studies should expand the geographical scope, including a wider range of industries and incorporating additional variables such as reviewer demographics as well as seasonal trends.

Author Contributions: *Siti Anisah*: Conceptualization, Methodology, Data Curation, Software, Writing - Original Draft. *Meditya Wasesa*: Supervision, Methodology, Writing - Review & Editing.

Funding: This research received no specific grant from any funding agency.

Conflicts of Interest: The authors declare no conflict of interest.

Data Availability: Data sources for this study can be accessed at https://bit.ly/googlereviews_data.

Informed Consent: There were no human subjects.

Animal Subjects: There were no animal subjects.

ORCID:
Siti Anisah: https://orcid.org/0009-0000-2468-3473
Meditya Wasesa: https://orcid.org/0000-0002-6051-4324

REFERENCES

[1] S. J. Dixon, "Share of consumer reviews posted on Google worldwide in 2020 and 2021," Statista, 2022. Accessed: Jan. 17, 2024. [Online]. Available: https://www.statista.com/statistics/1305930/consumer-reviews-posted-google/
[2] D. T. Andariesta and M. Wasesa, "Machine learning models to predict the engagement level of Twitter posts: Indonesian e-commerce case study," in *Procedia Computer Science*, Elsevier B.V., 2023, pp. 823–832. doi: 10.1016/j.procs.2023.10.588.
[3] G. Tandiawan and M. Wasesa, "Impact of Mixue's Halal Announcement on Company's Brand Reputation: a Naive Bayes Sentiment Analysis Approach," *Journal of Consumer Studies and Applied Marketing*, vol. 1, no. 1, pp. 74–80, Jul. 2023, doi: 10.58229/jcsam.v1i1.77.
[4] R. Thakur, "Customer engagement and online reviews," *Journal of Retailing and Consumer Services*, vol. 41, pp. 48–59, Mar. 2018, doi: 10.1016/j.jretconser.2017.11.002.
[5] A. H. Busalim, F. Ghabban, and A. R. C. Hussin, "Customer engagement behaviour on social commerce platforms: An empirical study," *Technol Soc*, vol. 64, Feb. 2021, doi: 10.1016/j.techsoc.2020.101437.
[6] I. Pletikosa Cvijikj and F. Michahelles, "Online engagement factors on Facebook brand pages," *Soc Netw Anal Min*, vol. 3, no. 4, pp. 843–861, Jan. 2013, doi: 10.1007/s13278-013-0098-8.
[7] E. J. Downes and S. J. McMillan, "Defining interactivity: A qualitative identification of key dimensions," *New Media Soc*, vol. 2, no. 2, pp. 157–179, 2000, doi: 10.1177/14614440022225751.
[8] S. S. Sundar, "Theorizing Interactivity's Effects," *The Information Society*, vol. 20, no. 5, pp. 385–389, Nov. 2004, doi: 10.1080/01972240490508072.
[9] K. P. Frey and A. H. Eagly, "Vividness can undermine the persuasiveness of messages.," *J Pers Soc Psychol*, vol. 65, no. 1, pp. 32–44, Jul. 1993, doi: 10.1037/0022-3514.65.1.32.
[10] M. Schreiner, T. Fischer, and R. Riedl, "Impact of content characteristics and emotion on behavioral engagement in social media: literature review and research agenda," *Electronic Commerce Research*, vol. 21, no. 2, pp. 329–345, Jun. 2021, doi: 10.1007/s10660-019-09353-8.
[11] P. Frangidis, K. Georgiou, and S. Papadopoulos, "Sentiment analysis on movie scripts and reviews: Utilizing sentiment scores in rating prediction," in *IFIP Advances in Information and Communication Technology*, Springer, 2020, pp. 430–438. doi: 10.1007/978-3-030-49161-1_36.
[12] D. Patel *et al.*, "Comparative Analysis of a Large Language Model and Machine Learning Method for Prediction of Hospitalization from Nurse Triage Notes: Implications for Machine Learning-based Resource Management," Aug. 10, 2023. doi: 10.1101/2023.08.07.23293699.
[13] D. C. Gkikas, K. Tzafilkou, P. K. Theodoridis, A. Garmpis, and M. C. Gkikas, "How do text characteristics impact user engagement in social media posts: Modeling content readability, length, and hashtags number in Facebook," *International Journal of Information Management Data Insights*, vol. 2, no. 1, Apr. 2022, doi: 10.1016/j.jjimei.2022.100067.
[14] L. Y. Koh, C. Ng, X. Wang, and K. F. Yuen, "Social media engagement in the maritime industry during the pandemic," *Technol Forecast Soc Change*, vol. 192, Jul. 2023, doi: 10.1016/j.techfore.2023.122553.
[15] S. K. Roy, G. Singh, S. Sadeque, P. Harrigan, and K. Coussement, "Customer engagement with digitalized interactive platforms in retailing," *J Bus Res*, vol. 164, Sep. 2023, doi: 10.1016/j.jbusres.2023.114001.
[16] A. Z. Abbasi, H. Qummar, S. Bashir, S. Aziz, and D. H. Ting, "Customer engagement in Saudi food delivery apps through social media marketing: Examining the antecedents and consequences using PLS-SEM and NCA," *Journal of Retailing and Consumer Services*, vol. 81, Nov. 2024, doi: 10.1016/j.jretconser.2024.104001.

[17] E. Izquierdo-Verdiguier and R. Zurita-Milla, "An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing," *International Journal of Applied Earth Observation and Geoinformation*, vol. 88, Jun. 2020, doi: 10.1016/j.jag.2020.102051.
[18] A. A. Hamidi, B. Robertson, and J. Ilow, "A new approach for ECG artifact detection using fine-KNN classification and wavelet scattering features in vital health applications," in *Procedia Computer Science*, Elsevier B.V., 2023, pp. 60–67. doi: 10.1016/j.procs.2023.09.011.
[19] A. Zamhuri Fuadi, Irsyad Nashirul Haq, and Edi Leksono, "Support Vector Machine to Predict Electricity Consumption in the Energy Management Laboratory," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 3, pp. 466–473, Jun. 2021, doi: 10.29207/resti.v5i3.2947.
[20] M. AminiMotlagh, H. S. Shahhoseini, and N. Fatehi, "A reliable sentiment analysis for classification of tweets in social networks," *Soc Netw Anal Min*, vol. 13, no. 1, Dec. 2023, doi: 10.1007/s13278-022-00998-2.
[21] X. E. Pantazi, D. Moshou, and D. Bochtis, *Intelligent Data Mining and Fusion Systems in Agriculture*. Elsevier, 2020. doi: 10.1016/C2017-0-01141-2.
[22] M. Nokkaew *et al.*, "Analyzing online public opinion on Thailand-China high-speed train and Laos-China railway mega-projects using advanced machine learning for sentiment analysis," *Soc Netw Anal Min*, vol. 14, no. 1, Dec. 2024, doi: 10.1007/s13278-023-01168-8.
[23] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. John Wiley & Sons, 2013. doi: 10.1002/9781118548387.
[24] Y. Qi and Z. Shabrina, "Sentiment analysis using Twitter data: a comparative application of lexicon- and machine-learning-based approach," *Soc Netw Anal Min*, vol. 13, no. 1, Dec. 2023, doi: 10.1007/s13278-023-01030-x.
[25] Y. Huang *et al.*, "Utilization of hyperspectral imaging for the analysis of aroma components of Soy Sauce-Aroma Type Baijiu," *Journal of Food Composition and Analysis*, vol. 134, Oct. 2024, doi: 10.1016/j.jfca.2024.106498.
[26] M. T. Puth, M. Neuhäuser, and G. D. Ruxton, "Effective use of Spearman's and Kendall's correlation coefficients forassociation between two measured traits," Apr. 01, 2015, *Academic Press*. doi: 10.1016/j.anbehav.2015.01.010.
[27] M. A. Afrianto and M. Wasesa, "The impact of tree-based machine learning models, length of training data, and quarantine search query on tourist arrival prediction's accuracy under COVID-19 in Indonesia," *Current Issues in Tourism*, vol. 25, no. 23, pp. 3854–3870, Dec. 2022, doi: 10.1080/13683500.2022.2085079.
[28] M. A. Afrianto and M. Wasesa, "Booking Prediction Models for Peer-to-peer Accommodation Listings using Logistics Regression, Decision Tree, K-Nearest Neighbor, and Random Forest Classifiers," *Journal of Information Systems Engineering and Business Intelligence*, vol. 6, no. 2, p. 123, Oct. 2020, doi: 10.20473/jisebi.6.2.123-132.
[29] A. D. Goenawan and S. Hartati, "The Comparison of K-Nearest Neighbors and Random Forest Algorithm to Recognize Indonesian Sign Language in a Real-Time," *Scientific Journal of Informatics*, vol. 11, no. 1, pp. 237–244, Feb. 2024, doi: 10.15294/sji.v11i1.48475.
[30] L. Su and D. H. Zhu, "A picture is worth a thousand words: Understanding the predictors of picture sharing in online consumer reviews," *Heliyon*, vol. 9, no. 12, p. e22789, Dec. 2023, doi: 10.1016/j.heliyon.2023.e22789.
[31] R. Filieri, E. Raguseo, and C. Vitari, "When are extreme ratings more helpful? Empirical evidence on the moderating effects of review characteristics and product type," *Human Behavior*, vol. 88, pp. 134–142, 2018, doi: 10.1016/j.chb.2018.05.042ï.
[32] M. Wasesa, "Business Analytics to Support Sustainable Community-Based Tourism: Insights from a Community Development Project in Pengudang Village, Bintan, Indonesia," 2024, pp. 117–130. doi: 10.1007/978-981-97-5219-5_7.
[33] R. Rahmadhan and M. Wasesa, "Segmentation using Customers Lifetime Value: Hybrid K-means Clustering and Analytic Hierarchy Process," *Journal of Information Systems Engineering and Business Intelligence*, vol. 8, no. 2, pp. 130–141, Oct. 2022, doi: 10.20473/jisebi.8.2.130-141.
[34] A. A. Nugraha and M. Wasesa, "Customer Segmentation and Preference Modeling of Indonesian Mobile Telecommunication Industry: A Data Mining Approach," in *2021 6th International Conference on Management in Emerging Markets (ICMEM)*, IEEE, Aug. 2021, pp. 1–6. doi: 10.1109/ICMEM53145.2021.9869383.

Publisher's Note: Publisher stays neutral regarding jurisdictional claims in published maps and institutional affiliations.