Purchase Intention and Sentiment Analysis on Twitter Related to Social Commerce

Muhammad Alviazra Virgananda[1], Indra Budi[2], Kamrozi[3], Ryan Randy Suryono[4]
Faculty of Computer Science Universitas Indonesia Jakarta, Indonesia[1, 3]
Faculty of Computer Science Universitas Indonesia Depok, Indonesia[2]
Faculty of Engineering and Computer Science Universitas Teknokrat Indonesia Bandar Lampung, Indonesia[4]

***Abstract*—Social commerce is a digital and efficient solution to transform existing commerce and address contemporary issues. TikTok Shop, a popular and trending social commerce platform, competes with established competitors like Facebook Marketplace and Instagram Shop. TikTok Shop offers benefits and incentives to attract users for both sales and product purchases. In this study, various algorithmic approaches such as Naïve Bayes, K-Nearest Neighbor, Support Vector Machine, Logistic Regression, Decision Tree, Random Forest, LGBM Boost, Ada Boost, and Voting Classifier are utilized to analyze and compare sentiments expressed on Twitter regarding Facebook, Instagram, and TikTok. The aim is to determine the methods with the best performance and identify the social commerce platform with the highest purchase intention and positive sentiment. The results indicate that TikTok has more positive sentiment than Facebook and Instagram at 93.07% with the best-performing classification model, Decision Tree. In conclusion, TikTok exhibits the highest positive sentiment percentage, indicating a greater number of positive reviews compared to Facebook and Instagram. According to the theory of evaluation scores for measuring model performance, values above 0.90 represent models with good performance.**

***Keywords*—Algorithm; machine learning; sentiment; social commerce**

I. Introduction

E-commerce has rapidly developed worldwide in the 2000s. This has led many traditional companies and stores to start opening online stores, which has brought about a change in the world of trade. Initially, business transactions were conducted in a traditional manner, with sellers and buyers interacting directly. However, with the emergence of e-commerce, these interactions shifted to an online platform. The increase in consumer trust and changes in online shopping behavior have supported the growth of e-commerce [1]. In the 2010s, there was a significant increase in social media users worldwide, marking the beginning of the emergence of social commerce. Social commerce combines e-commerce and social media, reflecting a shift in how businesses operate and interact with consumers. This shift has had a significant impact and positive benefits for both customers and sellers. Social commerce enables businesses to reach a wider and more distant target audience due to its ease of use [2], [3].

Social commerce is a hybrid of e-commerce and social media that enables trade transactions to take place on both e-commerce platforms and social media platforms like Facebook, Instagram, TikTok. With social commerce, consumers can purchase products or services from trusted sellers via social media platforms they use daily [4]. Sales on social commerce occur through social media features, such as links to online stores, business nuances, product reviewers, and instant shopping features. Additionally, social commerce serves as a marketing platform where businesses can promote their products and reach potential consumers through paid advertising or social listening. This type of social commerce has proven to be effective in boosting sales and user engagement [1], [5], [6].

There are many social commerce platforms circulating today such as Facebook marketplace, Instagram, and TikTok Shop. The Facebook marketplace is usually used for people to make transactions directly without a third person as an intermediary, while to make buying and selling transactions on Instagram features can be accessed through official accounts, or if we want to make sales on the Instagram market we have to register a special store first and if we want to make a purchase we have to visit the official store account of the product we are looking for on Instagram, in contrast to TikTok shop, which is a social e-commerce platform that is part of the TikTok feature and can be used for product sales [7], [8].

Currently, TikTok Shop is the main solution in the modern business world because the price it offers is relatively cheaper and benefits both sellers and buyers. Over time, TikTok Shop developed its innovations to allow users to make sales and purchases of products in the live videos they display. This feature was just launched in March 2021 and is currently available in several countries. The TikTok Shop offers a variety of products from various categories, such as fashion, beauty, electronics and more [9].

Purchase intention is based on people's interest and interest in an object. This interest led to purchases made by the public. N'da et al. 2023 [10] explores the direct and mediated effects of customers' perception of purchase budget (BGT) on purchase intention (PIT) through perceived quality (PPQ), perceived price (PPR) and perceived benefit (PB) in a cross-national context to understand the role of BGT in predicting customer purchase intention when selling smartphones online shopping through international platforms. While Jiang et al. 2023 [11] state that the market already has a significant size and the number is constantly increasing, the need to understand the factors affecting the purchase intention of consumers and explore the relationship between purchase intention and shopping behavior becomes more and more important and urgent. it can be concluded that purchase intention is very important as a parameter of the success of a market or business.

The aim of this research is to analyze the community's assessment of TikTok Shop, Instagram Shop, and Facebook Marketplace to determine which social commerce platform is profitable and receives the most positive sentiment for daily transactions such as buying and selling products. This research aims to increase trust, satisfaction, and purchase intentions of social commerce users. The study also explores the best-performing algorithms, the strengths and weaknesses of the methods used, and which social commerce platform generates the highest purchase intent and positive sentiment. Theoretically, this research contributes to a combination of algorithms that can be used and developed by other researches. Practically, this research contributes to social commerce developers (such as Facebook, Instagram and Tiktok) knowing their business position. So, the features in the application can be adjusted according to the community's response.

The chosen approach methods for this study include Naïve Bayes (NB), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), LGBM Boost (LGBM), Ada Boost (ADA), and Voting Classifier (VOT) algorithms. These algorithms have been frequently used in previous sentiment analysis studies due to their performance, accuracy, and ability to handle data effectively. Thus, they are deemed suitable for this research. Based on these papers [12]–[18], machine learning models are used in certain cases.

This paper is organized as follows: In Section II, we conducted several related works about sentiment analysis on social commerce. Then, in Section III explains the research methodology applied in this research. Section IV explains the results of the research. Last, the result of the research are explained in Section V.

II. Related Work

Facebook and Instagram have succeeded in attracting customer interest in buying or selling their products on those social commerce. The feature of buying and selling products in social commerce originally was not the main focus in their business. Since technological development was significantly improved, they take advantage of these features to increase interest from customers so they can maintain their loyalty. TikTok, being one of the social commerce participating in using the feature of buying and selling products as their business process TikTok provides various advantages that are quite advantageous to company owners. They will have more opportunities if they use the TikTok [9].

Nabiha et al. in 2021 [19] and Bayhaqy in 2018 [14] showed that sentiment analysis related to social commerce is only carried out with several single algorithm classification methods, such as Naïve Bayes, SVM, Decision Tree and KNN, while Lestari in 2022 [20] only uses a single algorithm classification method, namely KNN and the application of N-Gram to the method, and also focuses on only one platform, namely the TikTok Shop. Nabila et al. in 2021 [19] did not used model evaluation measurements such as Precision, Recall and F1-Score, and the accuracy value of the classification model tends to be quite low, namely below 0.75. Those research [14], [19], [20] only focused on social commerce or one of the social commerce platforms without comparing one platform to another, while this study uses nine classification methods consisting of single algorithms and ensemble algorithms which might improve the evaluation results for each method for each sentiment data related to three social commerce platforms, namely Facebook, Instagram and TikTok to determine which platform has more positive sentiment.

Botchway R et al. in 2022 [21] used binary particle swarm optimization to improve the accuracy of their models. The optimization method used can produce different accuracy values for each classification method. Botchway R and friends succeeded in increasing the accuracy value of the Naïve Bayes method by 11.6%, SVM by 8.43% and KNN by 0.91%. Meanwhile, in research Kamrozi et al. in 2023 [22] does not use a special method to increase the evaluation value of the method used, namely the Lexicon Method. Therefore, this study uses hyper parameter tuning to increase the evaluation value of the method used and produce different evaluation values for each method.

Das et al. 2023 [13] used machine learning approaches to stop hateful activities from happening, such as Logistic Regression, Gaussian Naïve Bayes, K-Nearest Neighbor, Decision Tree, Random Forest and Support Vector Machine on detection of hate speech from Twitter. Support Vector Machine, Decision Tree and Random Forest outperformed all the other models, achieving state-of-art 95.5%, 96.2% and 98.2% accuracy respectively at finding the hidden meaning inside the large number of comments and therefore determining whether there is any hateful event is going on or not. However, this research needs to use more models in order to be able to compare and test how well the other models are, especially ensemble algorithms.

III. Material and Methods

In this chapter, the methodology used to predict intent detection and sentiment analysis is explained in relation to the user experience when using TikTok. First, people's Twitter remarks will be extracted. Second, the pre-processing stage will eliminate inconsistent and incomplete data. Third, a feature selection approach for identifying discriminating phrases for training and classification will be deployed. Fourth, nine machine learning approaches using NB, KNN, SVM, LR, DT, RF, LGBM, ADA and VOT will be used to classify sentiments into two categories, positive and negative. Finally, an evaluation will be conducted by calculating the performance value for each method using metrics and then comparing which method has the best performance. Fig. 1 illustrates the research methodology used.

A. Data Collection

Data collection or information extraction was conducted on social media platform which frequently used by the public, namely Twitter. The information was gathered via Twitter's Application Programming Interface (API). Data of Twitter may be used to uncover themes and items being discussed based on certain keywords, to evaluate sentiment on specific businesses, and to obtain opinion on the latest products and services. For the intended data extraction needs, users may define needed attributes such as usernames, keywords, locations, name of place and others. [9], [23], [24].

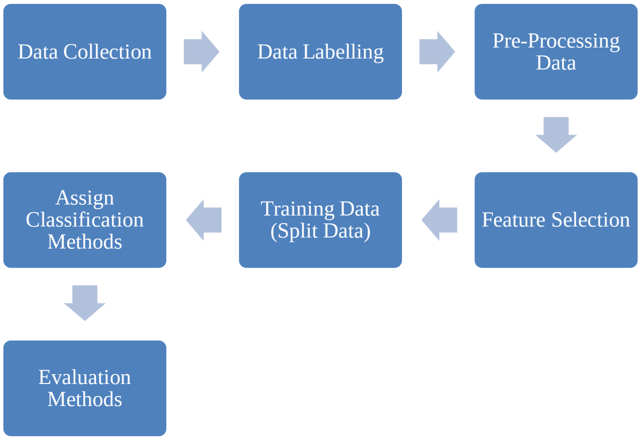

Fig. 1. Research methodology.

Data collection was carried out by collecting several community's tweet on Twitter regarding their ratings and experiences when using social commerce. The data collection was carried out with crawling data method using the Python programming language and Jupyter Notebook. Therefore, the dataset was developed independently and was not obtained from other paper references [Dataset is available in Supporting Materials section, Page 7].

In this section, the sentiment data obtained is divided into two categories, namely, positive and negative, for each social commerce platform: Facebook, Instagram, and TikTok. Fig. 2, 3, and 4 display the sentiment data collected from Twitter.

username	tweet	sentiment
@ihwanulfahri	ini orang marketplace facebook pada bego apa gimana sih ya, males banget nanggepin nya	negative
@fajrimiftahul	Jualan di marketplace facebook adalah jalan melatih kesabaran karena kebanyakan isinya orang ga ada pikiran	negative
@syndik	kalo soal jeans marketplace facebook narik customer nya ngebalap shopee. selalu begini kalo tiap celana jeans cowo	positive
@PoetStories	Thanks God laki gw ga insecure bininya beli HP baru dia cuma pake HP beli marketplace facebook.	positive
@LegaCahya	Marketplace facebook... Ditulis lokasi jogja, giliran ditanya jogjanya mana eh jawabnya dikirim dari jepara nanti kena biaya ongkir hzzz	negative
@Fikrinaufal77	Marketplace facebook napa aneh banget sih isinya? Ga bisa dibuang ya?	negative
@bledekkkk	Tim lebih suka mantengin marketplace facebook daripada shopee	positive
@queenduyza	ternyata marketplace facebook mantab wkwk	positive

Fig. 2. Example of Facebook's sentiment.

username	tweet	sentiment
@qeelaqeel	instagram shop ternyata lebih indah daripada syopi	positive
@cenamol	apasi instagram shop ini gaguna asli	negative
@tttokebi	barang di instagram shop emg mahal apa gmana	negative
@yifeiraa	instagram shop ilang ya? bagus deh	negative
@emilmarioo	kalo lu ngerasa ga berguna coba liat instagram shop	negative
@keknagojes	Instagram shop kalau boleh. Jangan post content hq banyak sangat. Asyik gambar gambar gambar. Twist gambar sendiri edit	negative
@badluckryaan	siapa juga mau gunain fitur instagram shop. ew	negative
@pepenjere	tab instagram shop yg ngegantiin tab activity ganggu bgt dah	negative
@nrpanthera	Kirain twitter story adalah fitur terabsurd semedsos. Instagram shop lebih kocak lagi anjing wkwkwkk	negative
@earlpeachy	Udah tanggal tua, nih Instagram pake nambahin fitur Instagram	positive
@keenandwataa	kalo lu ngerasa ga berguna coba cek instagram shop	negative

Fig. 3. Example of Instagram's sentiment.

username	tweet	sentiment
@emiss_	dulu demen banget ke mall sekarang klo ke mall kok ngantuk ya mikir juga ngapain ke mall udah nemu toko baju untuk anak suami dan aku yg cocok di tiktokshop	positive
@beubixing	gak lagi deh beli di tiktokshop	negative
@pifces	gua tagih belanja di tiktok shop anjir karna lebih murah dan harga beda jauh ama shopi cepet bat pula nyampenya	positive
@luvanillaa	dari shopee mulai alih ke lazada tiktokshop dari gojekgrab mulai alih ke maxim semua demi selisih yang cukup lumayan	positive
@bbykksl	buka tiktok shop ga bagus soal ak sedang hemat bisa jajan tp ak hrus puasa dlu	positive
@caffeeined	tiktok shop gajelas banget anjir gue beli pas live udah co liat fs kok mana gabisa dibatalin mau ganti yg sale padahal	negative
@ayudisafitri	ongkir tiktokshop sekarang mahal bgt saat say goodbye	negative
@Kacoweh	tiktok live shop ngeracunin banget anjirrr	positive

Fig. 4. Example of TikTok's sentiment.

TABLE I. HEADCOUNT OF SENTIMENT DATA

Categories	Facebook	Instagram	TikTok
Positive	931	485	6772
Negative	584	188	504
Percentage of Positive (%)	60.65	72,06	93,07
Percentage of Negative (%)	39,35	7,94	6,93
Total	1535	673	7276

Based on Table I, it is known that TikTok has the highest number of positive sentiments, amounting to 6772 sentiments, which accounts for 93.07% out of 7276 sentiments. It is followed by Instagram, which has 485 positive sentiments, representing the second-highest percentage of 72.06% out of 673 data. Facebook, on the other hand, has the lowest number of positive sentiments, with 931 sentiments, which corresponds to the lowest percentage of 60.65% out of 1535 sentiments. This indicates that TikTok generates a greater number of positive sentiments from the community compared to Facebook and Instagram.

B. Pre-processing Data

The data obtained from Twitter is in the form of text, as it contains sentiments or tweets from the public regarding a product or service. However, such data or information usually contains noise, which can make data analysis more challenging. Therefore, data pre-processing is carried out to remove unwanted words in tweets. All tweets are processed through four stages of pre-processing, which are as follows: Tokenization; Stopword Removal; Stemming; POS Tagging; and Bag of Words [18], [21], [23]–[25].

In this section, the results of data pre-processing for each dataset of Facebook, Instagram, and TikTok, regarding their positive and negative sentiments, are obtained. Table II represents the outcomes of a collection of words along with their frequency of occurrence in each social commerce platform.

Based on Table II, it is known that the sentiment data for each social commerce platform contains different words and their frequencies. This is done to observe which words are frequently discussed by the community regarding those social commerce platforms on Twitter. Additionally, it is also done to examine the correlation between the conducted research and the collected sentiment data, whether there is any connection or not. Based on the performed data pre-processing, it is found that the sentiment data related to Facebook, Instagram, and TikTok have the same data alignment, which involves discussions on purchases, sales, and users' utilization of those social commerce platforms.

TABLE II. STEM WORD AND FREQUENCY

Social Commerce	Stem Word (Frequency)
Facebook	('facebook', 1564), ('jual', 767), ('beli', 707), ('marketplace', 589), ('barang', 282), ('orang', 217), ('kalo', 156), ('harga', 149), ('jualan', 128), ('grup', 98), ('liat', 97), ('belanja', 66), ('murah', 64), ('langsung', 60), ('nawar', 57), ('rumah', 56), ('twitter', 52), ('shopee', 51), ('suka', 51), ('bikin', 51), ('pake', 50), ('iklan', 49), ('sampe', 49), ('buka', 48), ('cari', 46), ('harganya', 45), ('motor', 44), ('akun', 44), ('forum', 44), ('bener', 43), ('foto', 41), ('kena', 41)
Instagram	('instagram', 705), ('beli', 429), ('jual', 179), ('shop', 154), ('kalo', 80), ('liat', 73), ('orang', 64), ('iklan', 57), ('akun', 48), ('belanja', 46), ('baju', 45), ('barang', 42), ('tiket', 37), ('jualan', 35), ('twitter', 35), ('rumah', 33), ('bikin', 31), ('temen', 29), ('pake', 26), ('suka', 25), ('gimana', 24), ('kena', 24), ('harga', 24), ('muncul', 24)
TikTok	('tiktok', 5537), ('shop', 4964), ('tiktokshop', 2352), ('beli', 1415), ('shopee', 1004), ('live', 626), ('belanja', 605), ('barang', 577), ('jual', 540), ('murah', 401), ('racun', 389), ('bahaya', 319), ('kalo', 289), ('harga', 267), ('duit', 264), ('ongkir', 233), ('buka', 232), ('liat', 229), ('baju', 211), ('pake', 207), ('order', 202), ('diskon', 197), ('checkout', 193), ('kirim', 176), ('kena', 170), ('akun', 162)

C. Feature Selection Method

After the data is pre-processed, feature selection is performed as an important step in sentiment analysis. By selecting a subset of the important attributes to be included in the model's creation, this stage determines the most predictive traits. This method offers the benefit of lowering the data's high dimensionality and removing redundant, noisy, and unneeded content. Furthermore, this strategy can contribute on the development of a quick and accurate sentiment categorization. In this study, several factors influence feature selection, including data consistency, data amount, and the need to find the most effective feature selection approach [16], [20], [21], [23]–[28].

D. Split Data

Split data is data that has been partitioned into two or more subsets. A two-part split is commonly used to analyze or test the data and train the model. Data splitting is an important aspect of data science, particularly for building data-driven models. This strategy improves the accuracy of data models and data-driven processes such as machine learning [16], [19].

To reduce overfitting, data splitting is often employed in machine learning. In this scenario, a machine learning model fits its training data too well and fails to consistently fit further data. The initial data in a machine learning model is often separated into three or four categories. The training set, the development set, and the testing set are the three most popular sets [23]:

1) The training set is the collection of data used to train the model. The model should keep an eye on and learn from the training data, and any of its parameters should be improved.

2) The testing set is the piece of data examined in the final model and compared to the preceding data sets. The testing set is used to evaluate the final mode and algorithm.

Data should be separated so that large amounts of training data may be included in data sets. Data may be split 80-20 or 70-30 between training and testing, for example. The exact ratio varies depending on the data, but for small data sets, a 70-20-10 split for training, development, and testing works best [28].

The data in this study is separated into training and testing data. The data split is carried out using an 80% training data size and a 20% testing data size of the total data in the dataset. Table III illustrates the number of training and testing data from the three social commerce platforms for each sentiment data.

TABLE III. SPLIT DATA

Categories	Facebook	Instagram	TikTok
Training Data	1213	538	5821
Testing Data	303	135	1455
Total	1535	673	7276

Based on Table III, it is known that each dataset related to the sentiments of Facebook, Instagram, and TikTok has different numbers of training and testing data. This is due to the varying percentages of training and testing data sizes. The larger the percentage of training data, the higher the likelihood of accurate predictions on the testing data.

E. Evaluation of Classification Methods

At this stage, the performance evaluation of nine classification methods is carried out using standard classification performance metrics. Furthermore, four outcomes are possible at this point: true positive, false positive, true negative, and false negative. It is a true positive if the document label is positive and is classified as such. It is referred to as a false negative if it is classified as negative. A true negative is a negative document label that is classified as such. If it is labeled positive, it is considered a false positive [16]–[21], [23]–[28].

The accuracy measure is used to assess the accuracy of the likelihood of taking [16], [17], [19]–[21], [23], [24], [28]. The precision metric is the proportion of predicted classes that are the actual classes [16], [18], [20], [23]–[28]. The recall metric is the proportion of actual classes that are predicted as a class [17], [18], [20], [23]–[28]. The F1-Score (F) is used to assess model performance [18], [23]–[28]. Here are the formulas for each metric.

$$Accuracy\ (A) = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (1)$$

$$Precision\ (P) = \frac{TP}{TP+FP} \times 100 \quad (2)$$

$$Recall\ (R) = \frac{TP}{TP+FN} \times 100 \quad (3)$$

$$F1 - Score\ (F) = 2 * \frac{P*R}{P+R} \qquad (4)$$

The Accuracy, Precision, Recall, and F1-Score values varied from 0 to 1, with 1 being 100% similar and 0 representing 100% different. While TP, TN, FP, and FN represent the number of relevant identified features, relevant non-identified features, irrelevant identified features, and irrelevant non-identified features. Then, we applied 10-fold cross validation for evaluation.

F. Hyper Parameter Tuning

A Machine Learning model is a mathematical model with several parameters that must be taught from data. By training a model using existing data, we may fit the model parameters. [28], [29].

Another type of parameter, known as hyper parameters, cannot be taught using the normal training procedure. They are usually resolved before the training method begins. These parameters describe important model characteristics like as complexity and learning speed. The following are some examples of model hyper parameters [21], [26]:

1) The penalty in Logistic Regression Classifier i.e. L1 or L2 regularization.

2) The learning rate for training a neural network.

3) The C and sigma hyper parameters for support vector machines.

4) The k in k-nearest neighbors.

Models can have a large number of hyperparameters, and choosing the best combination of parameters can be thought of as a search problem. In this work, GridSearchCV and RandomizedSearchCV were used for hyper parameter optimization. This technique for GridSearch CV seeks the optimal collection of hyper parameters from a grid of hyper parameter values. While RandomizedSearchCV addresses the shortcomings of GridSearchCV by going through only a limited number of hyper parameter choices. It travels randomly within the grid to discover the optimal set of hyper parameters while also reducing wasteful processing.

IV. Results and Discussion

A. Results

In this section, evaluation is conducted on the classification methods used to measure their performance on sentiment data from Facebook, Instagram, and TikTok on Twitter. This evaluation includes accuracy, precision, recall, and F1-score. Table IV presents the evaluation scores for each method.

Based on Table IV, the evaluation scores for each classification model used on sentiment data from Facebook, Instagram, and TikTok are known. The evaluation scores range from 0 to 1, where a score closer to 1 indicates better performance, while a score closer to 0 indicates poorer performance.

The Random Forest Classifier has the highest assessment ratings for Facebook, with accuracy, precision, recall, and F1-score values of 0.80, 0.83, 0.75, and 0.77, respectively. The Logistic Regression classification model has the highest assessment ratings for Instagram, with accuracy, precision, recall, and F1-score values of 0.84, 0.86, 0.72, and 0.76, respectively. The Decision Tree classification model has the highest assessment ratings on TikTok, with accuracy, precision, recall, and F1-score values of 0.94, 0.74, 0.77, and 0.76, respectively.

However, for TikTok, there are no evaluation scores for the Random Forest and Voting Classifier models, as they have complex algorithms that require extensive memory capabilities to be executed on TikTok sentiment data. Therefore, during execution, the evaluation scores did not appear.

After evaluation value from the classification method had calculated, hyper parameter tuning was performed and it obtains the results. Table V presents the outcomes of utilizing hyper parameter tuning for multiple classification models on the sentiment data from each of the three social commerce platforms.

TABLE IV. EVALUATION OF CLASSIFICATION METHODS

Facebook				
Classifier	**Accuracy**	**Precision**	**Recall**	**F1-Score**
NB	0.518152	0.482661	0.291230	0.299336
KNN	0.729373	0.744252	0.657933	0.662345
SVM	0.768977	0.799486	0.704118	0.715866
LR	0.772277	0.758883	0.741819	0.748026
DT	0.749175	0.733304	0.714262	0.720534
RF	0.808581	0.831439	0.757690	0.773671
LGBM	0.752475	0.740581	0.711341	0.719420
ADA	0.716172	0.698963	0.665927	0.672177
VOT	0.798680	0.812544	0.749836	0.764106
Instagram				
Classifier	**Accuracy**	**Precision**	**Recall**	**F1-Score**
NB	0.725926	0.670817	0.698232	0.678178
KNN	0.725926	0.647857	0.645202	0.646472
SVM	0.785185	0.886719	0.597222	0.598914
LR	0.844444	0.865472	0.726010	0.760254
DT	0.800000	0.744652	0.731061	0.737184
RF	0.829630	0.834846	0.707071	0.737421
LGBM	0.800000	0.784895	0.660354	0.683071
ADA	0.807407	0.755309	0.736111	0.744467
VOT	0.837037	0.829445	0.729798	0.758458
TikTok				
Classifier	**Accuracy**	**Precision**	**Recall**	**F1-Score**
NB	0.936770	0.768966	0.515395	0.514254
KNN	0.923711	0.602292	0.543484	0.556365
SVM	0.944330	0.936527	0.569525	0.607047
LR	0.948454	0.949216	0.601783	0.654734
DT	0.939519	0.747713	0.772334	0.759298
LGBM	0.940893	0.970304	0.537634	0.554698
ADA	0.945704	0.859073	0.605324	0.653915

TABLE V. HYPER PARAMETER TUNING

Facebook				
Classifier	**Par 1**	**Par 2**	**Par 3**	**Accuracy**
NB	Var_smoothing: 0.0152	-	-	0.757273
SVM	C: 10	Gamma: 0.01	Kernel: rbf	0.781670
LR	C: 11,28	Penalty: 12	Solver: liblinear	0.790893
DT	Max_depth: 41	-	-	0.749986
LGBM	Learning_rate: 0.1	n_estimators: 50	-	0.725577
ADA	n_estimators: 50	-	-	0.702490
Instagram				
Classifier	**Accuracy**	**Precision**	**Recall**	**Accuracy**
NB	Var_smoothing: 0.2848	-	-	0.757943
SVM	C: 10	Gamma: 0.1	Kernel: rbf	0.795013
LR	C: 78,47	Penalty: 12	Solver: sag	0.805406
DT	Max_depth: 37	-	-	0.792128
LGBM	Learning_rate: 0.1	n_estimators: 50	-	0.,771232
ADA	n_estimators: 50	-	-	0.781691
TikTok				
Classifier	**Accuracy**	**Precision**	**Recall**	**Accuracy**
NB	Var_smoothing: 0.4328	-	-	0.933204
LR	C: 4,28	Penalty: 12	Solver: newton-cg	0.947223
LGBM	Learning_rate: 0.1	n_estimators: 150	-	0.934029

Based on Table V, the latest accuracy values are known after conducting hyper parameter tuning. Hyper parameter tuning was performed only on a subset of classification models due to memory limitations for highly complex classification models such as Random Forest, Voting Classifier, and others. Table VI compares the accuracy value before and after hyper parameter tuning was performed.

Based on Table VI, it is known that after conducting hyper parameter tuning, certain classification models experienced a decrease in accuracy scores. However, there are also several classification models that showed an improvement in accuracy scores. This can be influenced by the capacity and memory capabilities of the device used to execute the program. Upon re-execution, the obtained scores are also expected to differ. The values generated in this study represent the maximum iterations conducted to obtain the best possible scores. For example, Naive Bayes which has an accuracy value of 0.51 before using hyper parameter tuning, and 0.75 after using hyper parameter tuning. It can be said that hyper parameter tuning can increase and optimize the performance evaluation value of the resulting model. Mendes et al. 2023 [30] and Martineau et al. 2023 [31] explained that the use of hyper parameter tuning can optimize the performance evaluation value of the model so that it can increase the chances of success.

TABLE VI. BEFORE AND AFTER HYPER PARAMETER TUNING

Facebook		
Classifier	**Before**	**After**
NB	0.518152	0.757273
SVM	0.768977	0.781670
LR	0.772277	0.790893
DT	0.749175	0.749986
LGBM	0.752475	0.725577
ADA	0.716172	0.702490
Instagram		
Classifier	**Before**	**After**
NB	0.725926	0.757943
SVM	0.785185	0.795013
LR	0.844444	0.805406
DT	0.800000	0.792128
LGBM	0.800000	0.,771232
ADA	0.807407	0.781691
TikTok		
Classifier	**Before**	**After**
NB	0.936770	0.933204
LR	0.948454	0.947223
LGBM	0.940893	0.934029

B. Discussion

Based on the obtained results, it is evident that each sentiment dataset related to Facebook, Instagram, and TikTok has different percentages of positive and negative sentiments. Among the three social commerce platforms, TikTok has the highest percentage of positive sentiments, amounting to 93.07%, using the best classification model, which is Decision Tree with an accuracy score of 0.94. After evaluating the performance of each classification method on each social commerce platform, the following findings were obtained:

1) Facebook had a positive sentiment percentage of 60.65%, with the best classification model being Random Forest Classifier, achieving accuracy, precision, recall, and F1-score values of 0.80; 0.83; 0.75; and 0.77 respectively.

2) Instagram had a positive sentiment percentage of 72,06%, with the best classification model being Logistic Regression, achieving accuracy, precision, recall, and F1-score values of 0.84; 0.86; 0.72; and 0.76 respectively.

3) TikTok had a positive sentiment percentage of 93.07%, with the best classification model being Decision Tree, achieving accuracy, precision, recall, and F1-score values of 0.94, 0.74, 0.77, and 0.76, respectively.

This indicates that TikTok receives more positive reviews from the community compared to Facebook and Instagram, which cannot be known in research [14], [19]–[22].

Additionally, the classification model used for TikTok sentiment data also exhibits significantly high and consistent accuracy scores, always exceeding 0.90. According to the theory of evaluation scores for measuring model performance, values above 0.90 are considered to represent models with excellent performance.

With the results obtained that Logistic Regression, Decision Tree and Random Forest can produce high model performance evaluation values, it can be proven in the paper Das et al. 2023 [13], Imran et al. 2022 [12], Gulati et al. 2022 [15] which state that the Logistic Regression, Decision Tree and Random Forest models have high performance values and can be the best some case.

C. Limitations

There are several limitations on this study such as:

- We had only collect sentiment data from Facebook, Instagram and TikTok.
- We only measure the evaluation of methods through nine classification methods; therefore we do not use the other methods since the device of researcher is not compatible to execute the other methods.
- We only compared nine classification methods and prioritized the accuracy of evaluation from each method.

V. Conclusion

This section will explain the conclusion of the conducted research. Based on the research conducted, multiple machine learning classification models were tested on sentiment data related to Facebook, Instagram, and TikTok to determine which social commerce platform had the highest number of positive reviews.

Based on these results, it can be concluded that the classification models have different evaluation scores depending on the data used. The best classification model for Facebook is Random Forest Classifier, for Instagram is Logistic Regression, and for TikTok is Decision Tree. However, for classification models with low evaluation scores, using hyper parameter tuning may improve their performance.

In conclusion, TikTok exhibits the highest positive sentiment percentage, indicating a greater number of positive reviews compared to Facebook and Instagram. According to the theory of evaluation scores for measuring model performance, values above 0.90 represent models with excellent performance. Notably, the classification model used for TikTok sentiment data consistently achieves accuracy scores above 0.90. With this research, hopefully can help people choose the best social commerce to use and social commerce developers can increase their application and business value in order to increase public interest.

Acknowledgment

This study was supported by Hibah Publikasi Terindeks International (PUTI) Pascasarjana Tahun Anggaran 2023-2024 No. NKB-021/UN2.RST/HKP.05.00/2023.

Supporting Materials

[Link 1 – Facebook] Facebook - IJACSA.xlsx

[Link 2 – Instagram] Instagram - IJACSA.xlsx

[Link 3 – TikTok] TikTok - IJACSA.xlsx

References

[1] L. Zhoua, P. Zhang, and H.-D. Zimmermann, "Social Commerce Research: An Integrated View," Electron Commer Res Appl, vol. 12, pp. 61–68, Apr. 2013, doi: 10.1016/j.elerap.2013.02.003.

[2] X. Yin, H. Wang, Q. Xia, and Q. Gu, "How Social Interaction Affects Purchase Intention in Social Commerce: A Cultural Perspective," Sustainability, vol. 11, p. 2423, Apr. 2019, doi: 10.3390/su11082423.

[3] C. Gan and W. Wang, "The influence of perceived value on purchase intention in social commerce context," Internet Research, vol. 27, no. 4, pp. 772–785, Jan. 2017, doi: 10.1108/IntR-06-2016-0164.

[4] E. Turban, J. Strauss, and L. Lai, Social Commerce: Marketing, Technology and Management, 1st ed., vol. 21. Springer Cham, 2015. doi: https://doi.org/10.1007/978-3-319-17028-2.

[5] T.-P. Liang and E. Turban, "Introduction to the Special Issue Social Commerce: A Research Framework for Social Commerce," International Journal of Electronic Commerce, vol. 16, pp. 5–13, Dec. 2011, doi: 10.2307/23106391.

[6] J. Mou, M. Benyoucef, and J. Kim, "Benefits, risks and social factors in consumer acceptance of social commerce: A meta-analytic approach," in 26th Americas Conference on Information Systems, AMCIS 2020, 2020.

[7] D. Piranda, D. Sinaga, and E. Putri, "ONLINE MARKETING STRATEGY IN FACEBOOK MARKETPLACE AS A DIGITAL MARKETING TOOL," JOURNAL OF HUMANITIES, SOCIAL SCIENCES AND BUSINESS (JHSSB), vol. 1, pp. 1–8, Mar. 2022, doi: 10.55047/jhssb.v1i2.123.

[8] J. Che, C. Cheung, and D. Thadani, Consumer Purchase Decision in Instagram Stores:The Role of Consumer Trust. 2017. doi: 10.24251/HICSS.2017.004.

[9] C. I. Ratnapuri, M. Karmagatri, D. Kurnianingrum, I. D. Utama, and A. Darisman, "USERS OPINION MINING OF TIKTOK SHOP SOCIAL MEDIA COMMERCE TO FIND BUSINESS OPPORTUNITIES FOR SMALL BUSINESSES," J Theor Appl Inf Technol, vol. 101, no. 1, pp. 214–222, 2023.

[10] K. N'da, J. Ge, S. Ji-Fan Ren, and J. Wang, "Perception of the purchase budget (BGT) and purchase intention in smartphone selling industry: A cross-country analysis," PLoS One, vol. 18, no. 7, pp. e0279575-, Jul. 2023, [Online]. Available: https://doi.org/10.1371/journal.pone.0279575

[11] Q. Jiang, Y. Li, H. Wang, and D. Xie, "Analysis the Influential Factors of Consumers' Purchase Intention in Online Shopping," Lecture Notes in Education Psychology and Public Media, vol. 6, pp. 319–328, Jul. 2023, doi: 10.54254/2753-7048/6/20220352.

[12] B. Imran, Zaeniah, Sriasih, S. Erniwati, and S. Edu, "DATA MINING USING A SUPPORT VECTOR MACHINE, DECISION TREE, LOGISTIC REGRESSION AND RANDOM FOREST FOR PNEUMONIA PREDICTION AND CLASSIFICATION," vol. 10, pp. 792–802, Jul. 2022.

[13] S. Das, K. Bhattacharyya, and S. Sarkar, "Performance Analysis of Logistic Regression, Naive Bayes, KNN, Decision Tree, Random Forest and SVM on Hate Speech Detection from Twitter," vol. 7, pp. 24–28, Jul. 2023, doi: 10.47001/IRJIET/2023.703004.

[14] A. Bayhaqy, S. Sfenrianto, K. Nainggolan, and E. R. Kaburuan, "Sentiment Analysis about E-Commerce from Tweets Using Decision Tree, K-Nearest Neighbor, and Naïve Bayes," in 2018 International Conference on Orange Technologies (ICOT), 2018, pp. 1–6. doi: 10.1109/ICOT.2018.8705796.

[15] K. Gulati, S. Kumar, R. Boddu, K. Sarvakar, D. Sharma, and M. Nomani, "Comparative analysis of machine learning-based classification models using sentiment classification of tweets related to COVID-19 pandemic," Mater Today Proc, vol. 51, pp. 38–41, Feb. 2022, doi: 10.1016/j.matpr.2021.04.364.

[16] L. Mandloi and R. Patel, "Twitter Sentiments Analysis Using Machine Learninig Methods," in 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1–5. doi: 10.1109/INCET49848.2020.9154183.

[17] R. Othman, Y. Abdelsadek, K. Chelghoum, I. Kacem, and R. Faiz, "Improving Sentiment Analysis in Twitter Using Sentiment Specific Word Embeddings," in 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2019, pp. 854–858. doi: 10.1109/IDAACS.2019.8924403.

[18] J. Ho, D. Ondusko, B. Roy, and D. F. Hsu, "Sentiment Analysis on Tweets Using Machine Learning and Combinatorial Fusion," in 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), 2019, pp. 1066–1071. doi: 10.1109/DASC/PiCom/CBDCom/CyberSciTech.2019.00191.

[19] A. Nabiha, S. Mutalib, and A. M. A. Malik, "Sentiment Analysis for Informal Malay Text in Social Commerce," in 2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS), 2021, pp. 1–6. doi: 10.1109/AiDAS53897.2021.9574436.

[20] R. D. A. Lestari, B. S. Rintyarna, and Moh. Dasuki, "Application Of N-Gram On K-Nearest Neighbor Algorithm To Sentiment Analysis Of TikTok Shop Shopping Features," Journal of Institute of Computer Science (IOCS), vol. 6, no. 3, Oct. 2022.

[21] R. K. Botchway, V. Yadav, Z. O. Kominkova, and R. Senkerik, "Text-based feature selection using binary particle swarm optimization for sentiment analysis," in International Conference on Electrical, Computer, and Energy Technologies, ICECET 2022, 2022. doi: 10.1109/ICECET55527.2022.9872823.

[22] Kamrozi, A. N. Hidayanto, K. Y. P.M., Muh. A. Virgananda, and R. R. Suryono, "Sentiment Analysis of Cryptocurrency Trading Platform Service Quality on Playstore Data: A Case of Indodax," Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 2023.

[23] G. A. A. J. Alkubaisi, S. S. Kamaruddin, and H. Husni, "Conceptual framework for stock market classification model using sentiment analysis on twitter based on Hybrid Naïve Bayes Classifiers," International Journal of Engineering and Technology(UAE), vol. 7, no. 2, pp. 57–61, 2018, doi: 10.14419/ijet.v7i2.14.11156.

[24] M. I. Eshak, R. Ahmad, and A. Sarlan, "A preliminary study on hybrid sentiment model for customer purchase intention analysis in socialcommerce," in 2017 IEEE Conference on Big Data and Analytics (ICBDA), 2017, pp. 61–66. doi: 10.1109/ICBDAA.2017.8284108.

[25] F. Iqbal et al., "A Hybrid Framework for Sentiment Analysis Using Genetic Algorithm Based Feature Reduction," IEEE Access, vol. 7, pp. 14637–14652, 2019, doi: 10.1109/ACCESS.2019.2892852.

[26] S. N. Almuayqil, M. Humayun, N. Z. Jhanjhi, M. F. Almufareh, and N. A. Khan, "Enhancing Sentiment Analysis via Random Majority Under-Sampling with Reduced Time Complexity for Classifying Tweet Reviews," Electronics (Switzerland), vol. 11, no. 21, 2022, doi: 10.3390/electronics11213624.

[27] S. N. Almuayqil, M. Humayun, N. Z. Jhanjhi, M. F. Almufareh, and D. Javed, "Framework for Improved Sentiment Analysis via Random Minority Oversampling for User Tweet Review Classification," Electronics (Switzerland), vol. 11, no. 19, 2022, doi: 10.3390/electronics11193058.

[28] R. Obiedat et al., "Sentiment Analysis of Customers' Reviews Using a Hybrid Evolutionary SVM-Based Approach in an Imbalanced Data Distribution," IEEE Access, vol. 10, pp. 22260–22273, 2022, doi: 10.1109/ACCESS.2022.3149482.

[29] S. Urologin and S. Thomas, "3D visualization of sentiment measures and sentiment classification using combined classifier for customer product reviews," International Journal of Advanced Computer Science and Applications, vol. 9, no. 5, pp. 60–68, 2018, doi: 10.14569/IJACSA.2018.090508.

[30] P. Mendes, P. Romano, and D. Garlan, "Hyper-parameter Tuning for Adversarially Robust Models." Jul. 2023.

[31] T. Martineau, S. He, R. Vaidyanathan, and H. Tan, "Hyper-parameter tuning and feature extraction for asynchronous action detection from sub-thalamic nucleus local field potentials," Front Hum Neurosci, vol. 17, 2023, doi: 10.3389/fnhum.2023.1111590.