IJASCE : Int. J. of Adv. Sci. Comp. and Eng, 6(2) - August 2024 52-56

International Journal of
Advanced Science Computing and Engineering

Journal Homepage www.ijasce.org/index.php/ijasce

Predicting Customer Sentiment in Social Media Interactions: Analyzing Amazon Help Twitter Conversations Using Machine Learning

Md Arif [a], Mehedi Hasan [b,*], Sarder Abdulla Al Shiam [b], Md Parvez Ahmed [c], Mazharul Islam Tusher [d], Md Zikar Hossan [e], Aftab Uddin [f], Suniti Devi [a], Md Habibur Rahman [e], Md Zinnat Ali Biswas [g], Touhid Imam [h]

[a] Department of Management Science and Quantitative Methods, Gannon University, United States
[b] Department of Management -Business Analytics, St Francis College, United States
[c] Master of Science in Information Technology, Washington University of Science and Technology, United States
[d] Department Of Computer Science, Monroe College, New Rochelle, New York, United States
[e] Department of Business Administration, International American University, Los Angeles, California, United States
[f] Fox School of Business & Management, Temple University, United States
[g] MA in Education, University of South Wales, Wales, United Kingdom
[h] Department of Computer Science, University of South Dakota, Vermillion, South Dakota, United States
Corresponding author: [*]mhasan5@sfc.edu

***Abstract*— Social media platforms, particularly Twitter, have become essential sources of data for various applications, including marketing and customer service. This study focuses on analyzing customer interactions with Amazon's official support account, "@AmazonHelp," to understand and predict changes in customer sentiment during these interactions. Using the Twitter API, we extracted English-language tweets mentioning "@AmazonHelp," pre-processed the data, and categorized conversations to facilitate analysis. The primary objectives were to classify changes in customer sentiment and predict the overall sentiment change based on initial sentiment. We conducted experiments using multiple machines learning algorithms, including K-nearest neighbor, Naive Bayes, Artificial Neural Network, Bayes Net, Support Vector Machine, Logistic Regression, and Bagging with RepTree. Our dataset comprised over 6,500 conversations, filtered to include those with four or more tweets. Results indicated that K-nearest neighbor and Support Vector Machine offered the best balance between accuracy and F-measure, while Bagging with RepTree achieved the highest accuracy but had a lower F-measure. This study demonstrates the potential of integrating sentiment analysis and machine learning to effectively predict customer sentiment in social networks, providing valuable insights for improving customer engagement strategies.**

***Keywords*—K-nearest neighbor; naive bayes; artificial neural network; bayes net; support vector machine; logistic regression.**

Manuscript received 15 Jan. 2024; revised 29 Mar. 2024; accepted 14 Apr. 2024. Date of publication 31 Aug. 2024.

I. INTRODUCTION

Social media platforms like Twitter have become invaluable sources of data for a wide range of applications, particularly in marketing and social studies [1], [2] Companies can leverage this data to enhance customer engagement by embedding deep links in their tweets, which display call-to-action buttons for direct messaging. This study focuses on analyzing customer interactions with Amazon's official support account, "@AmazonHelp," to understand and predict changes in customer sentiment during these interactions. By utilizing the Twitter API, we extracted English-language tweets mentioning "@AmazonHelp" and pre-processed the data through extraction and text cleaning steps. Conversations were simplified into a two-level structure to facilitate analysis. The text cleaning process involved removing URLs, usernames, hashtags, emoticons, special characters, and separators, followed by lemmatization and grammatical tagging.

The primary goals of this research are twofold: to classify changes in customer sentiment throughout their interactions with retail service providers and to predict the overall sentiment change of a conversation based on its initial sentiment [3]-[9]. To achieve these goals, we conducted experiments using various machine learning algorithms to determine the most effective methods for classifying and predicting sentiment changes [10]-[15]. The dataset, collected using the Twitter API, comprised over 6,500 conversations, filtered to include only those with four or more tweets to ensure meaningful analysis. The final dataset, balanced for polarity change, was split into training and test sets.

We evaluated the performance of several machine learning algorithms, including K-nearest neighbor, Naive Bayes, Artificial Neural Network, Bayes Net, Support Vector Machine, Logistic Regression, and Bagging with RepTree. The algorithms were assessed based on their accuracy and F-measure scores to determine their suitability for predicting customer sentiment in social networks. The results indicate that K-nearest neighbor and Support Vector Machine provide the best balance between accuracy and F-measure, while Bagging with RepTree achieves the highest accuracy but has a lower F-measure [16]. In contrast, the Artificial Neural Network underperformed significantly, highlighting the importance of selecting an appropriate algorithm based on the specific application requirements. This study underscores the potential of integrating sentiment analysis and machine learning to predict customer perception in social networks effectively [17], [18].

In this stage, unprocessed data from source users and conversations was turned into features for tasks involving prediction and classification. Tweet features (text, author, time, hashtags, media, URLs, retweet count, and favorite count), source user features (author's verified status, followers count, and followers count), and conversational features (conversation length, total duration, number of external users, and number of replies/comments from external users) are among the attributes that are extracted from the Twitter API for each conversation. According to earlier research, these characteristics—which can be summed up as follows—were chosen for their possible influence on the emotional transitions that occur between clients and customer service representatives [19]. Notably, understanding the variables that change a customer's negative sentiment into a positive sentiment depends on the length of the conversation.

Twitter has become a valuable source of data for various domains, including marketing and social studies Companies can enhance customer engagement by adding deep links to their tweets, which display a call-to-action button for sending direct messages. In this context, we focused on tweets directed at Amazon's official customer support account, "@AmazonHelp" [18] to analyze conversations and the change in customer sentiment throughout these interactions. Using the Twitter API, we extracted only English tweets mentioning "@AmazonHelp." The pre-processing involved two steps: extracting entire conversations and cleaning the text. Conversations were simplified into a two-level structure, with the source tweet at the first level and all subsequent replies at the second level. Source tweets were identified as those not replying to any other tweets. Once the source tweets were collected, the rest of the conversation was retrieved, including tweets from the original customer, AmazonHelp, and other participants. Text cleaning involved stripping URLs, usernames, hashtags, and emoticons, normalizing the tweets by removing special characters and separators, and applying lemmatization and grammatical tagging. This approach aimed to uncover significant features influencing changes in customer sentiment during these interactions.

Saragih and Girsang *et al* [20] investigated customer engagement by analyzing comments on Facebook and Twitter for Indonesian online transportation companies: Gojek, Grab, and Uber. They utilized API services to collect data from public pages and accounts, categorizing comments into positive, negative, and neutral sentiments using an Indonesian sentiment word library. A scoring system was applied to count comments in each category, and results were compared with the companies' number of followers to examine the correlation between engagement and feedback. Al-Otaibi *et al* [21] created a Tweet advisor system to assess customer satisfaction through sentiment analysis using a Support Vector Machine (SVM) algorithm. Their system analyzed Twitter accounts for activity and engagement and allowed searches for keywords, hashtags, or mentions. They found that the unigram method combined with SVM achieved 87% accuracy, surpassing the bigram method. Fitri *et al.* [18] measured public satisfaction with a telecommunications operator's data services in Indonesia using Twitter sentiment analysis. Employing a Naive Bayes Classifier (NBC), they classified tweets from the AmazonHelp Twitter account over four months, achieving 99% accuracy.

II. Materials and Method

This paper focuses on two primary goals. The first goal is to classify changes in customer sentiment throughout their interactions with retail service providers. The second goal aims to predict how the sentiment of a conversation will change based on its initial sentiment, allowing for the prediction of the conversation's sentiment at its conclusion. This section discusses the application of machine learning techniques to classify and forecast changes in customer sentiment polarity.

The study conducted three primary experiments, each with distinct objectives. The first experiment utilized all feature groups outlined applying various machine learning algorithms to identify the most effective one for the problem at hand. The second experiment evaluated the importance of each feature group in predicting changes in conversation polarity. The third experiment aimed to predict polarity changes throughout the conversation, using a model trained and tested with only a subset of features available at the beginning of the conversation.

A. Dataset collection and processing

Conversations were sourced from the Twitter API by searching for English tweets mentioning @AmazonHelp and retrieving subsequent tweets. Conversations varied in length, and only those with four or more tweets were included, resulting in 6,538 conversations collected from January 1, 2021, to December 31, 2022. Brief exchanges with only two tweets were excluded. To understand customer perceptions, sentiment analysis tools were applied to label the first and last tweets of each conversation as positive, negative, or neutral, reflecting the customer's sentiment towards the product or service.

For this study, the Stanford API [19], [22] was used to detect customer attitudes in conversations by labeling the initial and final customer tweets. This tool leverages a

Recursive Neural Network (RNN) model [23], which analyzes polarity by considering the structure of the sentences. Utilizing an RNN model helps preserve the sentence's meaning and semantics, improving the accuracy of sentiment labeling throughout the conversation. The conversation labeling process comprises two main steps, each aimed at discerning the sentiment expressed by the customer. These steps were elucidated in detail to provide clarity on how the sentiment analysis tool effectively captures customers' [24] perceptions and sentiments throughout their interactions.

B. Dataset Training and Testing

Preparing the training and testing sets was the second step in our experimental setup. The dataset of conversations initially showed a class imbalance, with about 50.94% of the conversations displaying no change in polarity. This disparity was primarily the consequence of interactions in which clients only asked for advice rather than complaining. In order to rectify this, we eliminated discussions that started out neutrally and showed no signs of polarity shifting. For the first tweet, a sentiment label of 0 denoted a neutral start, and a CPC value of zero suggested no polarity shift.

Following this filtering process, we achieved a more balanced dataset suitable for classification. The final dataset comprised 5,000 conversations, totaling 40,000 tweets. Among these, 1,200 conversations showed a negative change, 1,900 showed a positive change, and 900 exhibited no change in polarity. The conversations and their features were then divided into training and testing sets in an 80%-20% split, resulting in 3,500 instances in the training set and 1,500 instances in the test set. We ensured that the class proportions remained consistent in both sets.

III. Result and Discussion

An experiment was carried out to evaluate the accuracy of various classification algorithms in order to compare different classifiers and validate the proposed framework. The accuracy and F-measure results of the evaluation of both classical and ensemble classifiers are summarized in Table 1. For this, the Weka tool was employed, which is well-known for its wide variety of machine-learning algorithms for data mining. For classification, all features were combined into a feature vector. Table 1 lists the classifiers that were tested: Support Vector Machine (SMO), Naive Bayes (Naive Bayes), Artificial Neural Network (Multilayer-Perceptron), K-nearest neighbor (IBk), Bayes Net (Bayes Net), and Logistic Regression (Logistic) [25], [26]. In addition, the comparison included the ensemble classifier Bagging [27].

Based on the result, we have a set of machine learning algorithms with their respective accuracy and F-measure percentages for a task, presumably predicting customer perception in social networks by integrating sentiment analysis. Here's the figure 1 comparison of the different machine-learning results. For predicting customer perception in social networks using sentiment analysis and machine learning, the choice of algorithm can significantly affect performance. The K-nearest neighbor (k = 7) and Support Vector [27] Machine (complexity = 2.0) offer the best balance between accuracy and F-measureBagging with a RepTree classifier, although highly accurate, may not be as reliable when considering both precision and recall, as indicated by its lower F-measure. On the other hand, an Artificial Neural Network with a learning rate of 0.1 seems to underperform significantly in this context. Selecting the appropriate algorithm will depend on whether accuracy or the balance between precision and recall (as reflected by the F-measure) is more critical for the specific application.

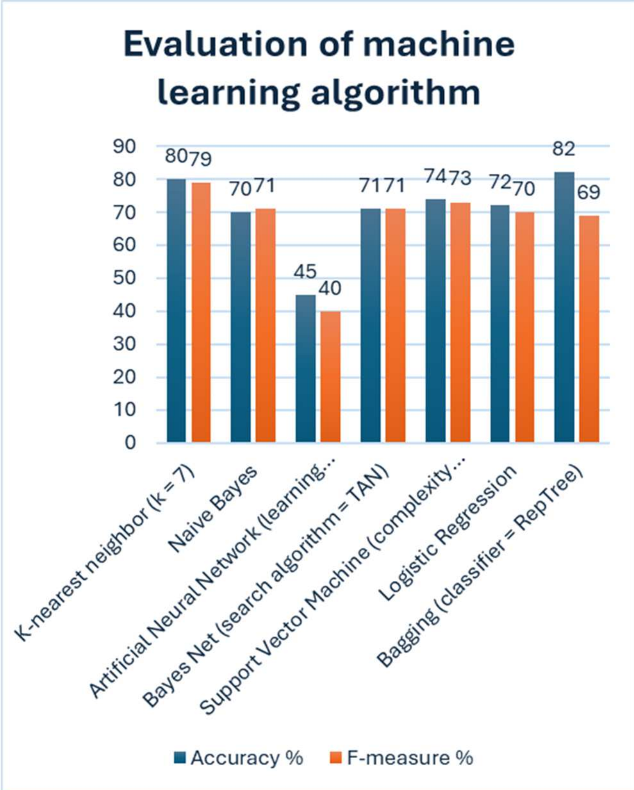

Fig. 1 Evaluation of different machine learning algorithms

TABLE I
Evaluation of Different Machine Learning Algorithm

ML Algorithm	**Accuracy %**	**F-measure %**
K-nearest neighbor (*k* = *7*)	80	79
Naive Bayes	70	71
Artificial Neural Network (*learning rate* = *0.1*)	45	40
Bayes Net *(search algorithm = TAN)*	71	71
Support Vector Machine (*complexity* = *2.0*)	74	73
Logistic Regression	72	70
Bagging *(classifier = RepTree)*	82	69

Predicting customer perception in social networks by integrating sentiment analysis with machine learning, different algorithms demonstrate varying levels of effectiveness. The Bagging classifier using RepTree achieves the highest accuracy at 82%, but its lower F-measure of 69% suggests it may struggle with precision or recall balance. The K-nearest neighbor (k = 7) offers a strong balance with an accuracy of 80% and a high F-measure of 79%, making it a reliable choice for balanced performance. Support Vector Machine (complexity = 2.0) also performs well with 74% accuracy and a 73% F-measure. Naive Bayes and Bayes Net

(TAN) provide moderate and balanced results, both around 70-71% for both metrics. Logistic Regression, while having a slight discrepancy, still performs reasonably with 72% accuracy and a 70% F-measure. In contrast, the Artificial Neural Network (learning rate = 0.1) significantly underperforms with the lowest scores of 45% accuracy and a 40% F-measure, indicating it may not be suitable for this task. Overall, K-nearest neighbor and Support Vector Machine are the most balanced and effective choices, while Bagging offers high accuracy but at the cost of lower F-measure.

IV. Conclusion

In conclusion, this study underscores the significance of integrating sentiment analysis with machine learning to predict customer perception in social networks effectively. Through the analysis of customer interactions with Amazon's official support account, "@AmazonHelp," we aimed to classify changes in customer sentiment and predict overall sentiment changes during these interactions.

Our experiments evaluated various machine learning algorithms, including K-nearest neighbor, Naive Bayes, Artificial Neural Network, Bayes Net, Support Vector Machine, Logistic Regression, and Bagging with RepTree. Results indicated that K-nearest neighbor and Support Vector Machine offered the best balance between accuracy and F-measure, while Bagging with RepTree achieved the highest accuracy but had a lower F-measure. Conversely, the Artificial Neural Network underperformed significantly in this context.

The choice of algorithm significantly impacts performance, with considerations of accuracy and the balance between precision and recall (as reflected by the F-measure) being crucial for the specific application. K-nearest neighbor and Support Vector Machine emerged as reliable choices for balanced performance, whereas Bagging offered high accuracy but at the cost of lower F-measure [28].

Moving forward, understanding customer sentiment dynamics in social media interactions can offer valuable insights for enhancing customer engagement strategies. Further research could explore the integration of additional features or fine-tuning algorithms to improve predictive accuracy and better capture nuanced changes in customer sentiment throughout interactions. Additionally, exploring real-time applications of sentiment analysis in social media customer service could further enhance its utility for businesses in effectively addressing customer needs and concerns.

References

[1] A. K. Dwivedi, "Performance evaluation of different machine learning techniques for prediction of heart disease," *Neural Comput. Appl.*, vol. 29, no. 10, pp. 685–693, Sep. 2016, doi: 10.1007/s00521-016-2604-1.

[2] R. Aswani, A. K. Kar, P. V. Ilavarasan, and Y. K. Dwivedi, "Search engine marketing is not all gold: Insights from Twitter and Seoclerks," *Int. J. Inf. Manage.*, vol. 81, pp. 107–116, 2018, doi:10.1016/j.ijinfomgt.2017.07.005.

[3] S. Park, J. Cho, K. Park, and H. Shin, "Customer sentiment analysis with more sensibility," *Eng. Appl. Artif. Intell.*, vol. 104, p. 104356, 2021, doi: 10.1016/j.engappai.2021.104356.

[4] D. Gräbner, M. Zanker, G. Fliedl, and M. Fuchs, "Classification of customer reviews based on sentiment analysis," in *Inf. Commun. Technol. Tour. 2012*, Vienna, Austria: Springer, 2012, pp. 460–470, doi: 10.1007/978-3-7091-1142-0_40.

[5] S. Yi and X. Liu, "Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review," *Complex Intell. Syst.*, vol. 6, no. 3, pp. 621–634, 2020, doi:10.1007/s40747-020-00155-2.

[6] O. Grljević and Z. Bošnjak, "Sentiment analysis of customer data," *Strategic Manage. – Int. J. Strategic Manage. Decision Support Syst. Strategic Manage.*, vol. 23, no. 3, 2018, doi:10.5937/StraMan1803038G.

[7] F. Sudirjo, K. Diantoro, J. A. Al-Gasawneh, H. K. Azzaakiyyah, and A. M. A. Ausat, "Application of ChatGPT in improving customer sentiment analysis for businesses," *J. Teknol. Syst. Inf. Bisnis*, vol. 5, no. 3, pp. 283–288, 2023, doi: 10.47233/jteksis.v5i3.871.

[8] V. Sharma and T. Manocha, "Comparative analysis of online fashion retailers using customer sentiment analysis on Twitter," in *Proc. Int. Conf. Innov. Comput. Commun. (ICICC)*, Feb. 2022, doi:10.2139/ssrn.4361107.

[9] S. Jumabek, K. Tuwelbay, A. Allabay, and B. Geldibayev, "Current trends in the formation of mathematical terminology on an online platform," *Int. J. Adv. Sci. Comput. Eng.*, vol. 6, no. 1, pp. 7–12, 2024, doi: 10.62527/ijasce.6.1.188.

[10] Z. Singla, S. Randhawa, and S. Jain, "Sentiment analysis of customer product reviews using machine learning," in *Proc. Int. Conf. Intell. Comput. Control (I2C2)*, Jun. 2017, pp. 1–5, doi:10.1109/I2C2.2017.8321910.

[11] M. S. Hossain and M. F. Rahman, "Customer sentiment analysis and prediction of insurance products' reviews using machine learning approaches," *FIIB Bus. Rev.*, vol. 12, no. 4, pp. 386–402, 2023, doi:10.1177/23197145221115793.

[12] Z. A. Dieksona et al., "Sentiment analysis for customer review: Case study of Traveloka," *Procedia Comput. Sci.*, vol. 216, pp. 682–690, 2023, doi: 10.1016/j.procs.2022.12.184.

[13] A. R. Prananda and I. Thalib, "Sentiment analysis for customer review: Case study of GO-JEK expansion," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 6, no. 1, p. 1, 2020, doi: 10.20473/jisebi.6.1.1-8.

[14] A. Adak, B. Pradhan, and N. Shukla, "Sentiment analysis of customer reviews of food delivery services using deep learning and explainable artificial intelligence: Systematic review," *Foods*, vol. 11, no. 10, p. 1500, 2022, doi: 10.3390/foods11101500.

[15] S. Moghaddam, "Beyond sentiment analysis: Mining defects and improvements from customer feedback," in *Adv. Inf. Retr., 37th Eur. Conf. IR Res. (ECIR)*, Vienna, Austria, Mar.–Apr. 2015, pp. 400–410, doi: 10.1007/978-3-319-16354-3_44.

[16] A. Mardiah, D. Defni, A. H. Lestari, J. Junaldi, and T. Ritmi, "Classification of population data of Nagari based on economic level using the K-nearest neighbor method," *Int. J. Adv. Sci. Comput. Eng.*, vol. 6, no. 1, pp. 33–37, 2024, doi: 10.62527/ijasce.6.1.191.

[17] M. A. Rahman et al., "Advancements in retail price optimization: Leveraging machine learning models for profitability and competitiveness," *J. Bus. Manage. Stud.*, vol. 6, no. 3, pp. 103–110, 2024, doi: 10.32996/jbms.2024.6.3.11.

[18] L. Fitri, D. Lestari, and F. Nurul, "Measuring public satisfaction with data services using sentiment analysis on Twitter," in *Proc. Int. Conf. Data Softw. Eng. (ICoDSE)*, 2018.

[19] N. F. Ibrahim, X. Wang, and H. Bourne, "Exploring the effect of user engagement in online brand communities: Evidence from Twitter," *Comput. Human Behav.*, vol. 72, pp. 321–338, 2017, doi:10.1016/j.chb.2017.03.005.

[20] H. S. Saragih and A. S. Girsang, "Customer engagement analysis on social media using sentiment mining: A case study on online transportation companies in Indonesia," *J. Theor. Appl. Inf. Technol.*, vol. 95, no. 7, 2017.

[21] S. Al-Otaibi, M. Ykhlef, and N. Aljohani, "A tweet advisor system for measuring customer satisfaction using sentiment analysis," *J. King Saud Univ. – Comput. Inf. Sci.*, 2018.

[22] H. Shirdastian, M. Laroche, and M.-O. Richard, "Using big data analytics to study brand authenticity sentiments: The case of Starbucks on Twitter," *Int. J. Inf. Manage.*, vol. 48, pp. 291–307, Oct. 2019, doi:10.1016/j.ijinfomgt.2017.09.007.

[23] M. S. Rana, M. M. Hossain, R. M. Jewel, and M. R. Islam, "Evaluating customers satisfaction of electronic banking: An empirical study in Bangladesh," *SIJ Trans. Ind. Financial Bus. Manage.*, vol. 5, no. 3, pp. 7–12, 2017, doi: 10.9756/SIJIFBM/V5I3/05040160101.

[24] B. P. Ghosh et al., "Deep learning in stock market forecasting: Comparative analysis of neural network architectures across NSE and NYSE," *J. Comput. Sci. Technol. Stud.*, vol. 6, no. 1, pp. 68–75, 2024, doi: 10.32996/jcsts.2024.6.1.8.

[25] M. S. Chowdhury et al., “Deep learning models for stock market forecasting: A comprehensive comparative analysis,” *J. Bus. Manage. Stud.*, vol. 6, no. 2, pp. 95–99, 2024, doi: 10.32996/jbms.2024.6.2.9.
[26] M. Bhuiyan, “Carbon footprint measurement and mitigation using AI,” *SSRN Electron. J.*, 2024, doi: 10.2139/ssrn.4746446.
[27] M. A. S. Mozumder et al., “Revolutionizing organizational decision-making for banking sector: A machine learning approach with CNNs in business intelligence and management,” *J. Bus. Manage. Stud.*, vol. 6, no. 3, pp. 111–118, 2024, doi: 10.32996/jbms.2024.6.3.12.
[28] O. Hammad, M. R. Rahman, N. Clements, S. Mishra, S. Miller, and E. Sullivan, “PureNav: A personalized navigation service for environmental justice communities impacted by planned disruptions,” in *Proc. Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Nov. 2023, pp. 56–63, doi: 10.1145/3625007.3627512.