[1]Atluri Vani Vathsala,

[2]Lakshmi H. N.

[3]LNC. Prakash K.

[4]Thaduri Venkata Ramana

[5]Kachapuram Basava Raju

Enhancing Product Perception: A Novel Meta-Model Approach for Sentiment Analysis in Product-Based Reviews

Abstract: - The proliferation of online purchasing and product evaluations, along with the rapidly expanding e-commerce industry, poses a significant problem for businesses looking to derive actionable data. High-star ratings are common, yet essential negative evaluations are sometimes overlooked. The urgent need for efficient sentiment analysis methods to enable businesses to extract sentiment from reviews and ratings is discussed in this study. By utilizing these approaches, companies may increase their comprehension of consumer input and make well-informed decisions that will lead to better product development and customer happiness. Sentiment analysis aids efforts to enhance products by revealing user perceptions of the items. However, traditional machine learning-based sentiment analysis techniques are unreliable and come with a hefty computational cost. While deep learning has demonstrated encouraging progress in sentiment analysis techniques, efficiently optimizing hyperparameters is still a hurdle. To determine which supervised machine learning classification technique produces the most reliable sentiment analysis results, this paper compares several supervised machine learning techniques used on online product reviews with the meta-model, which combines neural networks with support vector machines that have been proposed. With an amazing accuracy rate, the model reported in this study outperformed conventional methods. These findings highlight the usefulness of sentiment research in clarifying client attitudes and enhancing goods from a commercial perspective.

Keywords: Sentiment analysis, online reviews, supervised learning, Neural networks, Support vector machines, meta-model

I. INTRODUCTION

Customer product reviews have emerged as one of the most significant sources of data due to the explosive rise of e-commerce and online purchasing. Customers can rate and review goods and services on websites like Yelp, Amazon, Flipkart, TripAdvisor, and Amazon based on their experiences and satisfaction. As a result, there is now a huge reservoir of opinion-rich data that offers raw insights into how consumers see particular brands, goods, attributes, and the whole consumer experience. Organizations in marketing, product creation, customer support, and competitive intelligence can benefit greatly from mining this data to extract detailed comments and gauge general sentiment. Sentiment analysis is the efficient identification, extraction, quantification, and study of emotional expressions and particular communication from text data using natural language processing, text examination, and computational morphology. Finding a speaker's or writer's attitude towards a subject or the document's overall contextual polarity is the aim. It seeks to ascertain whether a reviewer's assessment of a product and its features is favorable, unfavorable, or neutral. This goes beyond simply taking the review rating and pulling it out of the text, which may not always match up. Reviews' sentiment analysis offers specific information about the features that buyers find appealing or objectionable. We examine the methods created especially for sentiment

[1] Professor and Head, Dept of CSE, CVR College of Engineering, Hyderabad, Telangana

[2] Professor and Head, Dept of CSE (AI&ML, DS, CS), CVR College of Engineering, Hyderabad, Telangana

[3] Assoc. Professor, Dept of CSE-DS, CVR College of Engineering, Hyderabad, Telangana

[4] Assoc. Professor, Dept of CSE-AI&ML, CVR College of Engineering, Hyderabad, Telangana

[5] Assoc. Professor, Dept of AI, Anurag University, Hyderabad, Telangana

*Corresponding author: Thaduri Venkata Ramana

analysis of online product reviews in this research. We talk about the special difficulties presented by product reviews, which are considerably more intricate than those in other fields like social media. Reviews include a variety of elements and characteristics, sarcasm, irony, and extensive domain-specific vocabulary and concepts. Deep subject knowledge is necessary for high accuracy. We provide an overview of the entire sentiment analysis pipeline, which includes aspects such as aspect-based and comparative modelling, evaluation metrics and datasets customized for product reviews, pre-processing and feature extraction, classification algorithms, and data collecting and sampling techniques. The most popular method is supervised machine learning, in which classifiers such as CNN, RNN, BERT, SVM, and Naïve Bayes are trained on labelled review data and then used to predict sentiment in unseen reviews. We describe how input is provided by natural language features such as word embeddings, POS tags, n-grams, and sentiment lexicons. Modern deep learning techniques that focus on model text and use LSTMs have advanced accuracy beyond previous limits. For fine-grained analysis, aspect extraction and matching opinions to target features are still open research issues. As alternatives, we also look at lexicon-based and semi-supervised methods. Real-world problems, including handling massive data volumes, multilingual sentiment analysis, domain adaptability, and idea drift over time, are given special attention. We also go over explainable AI methods for sentiment analysis. Key issues and prospects are covered in the paper's conclusion, along with new developments like aspect-based sentiment summarization, multimodal analysis, and fine-grained emotion identification. Our poll offers a thorough rundown of the most recent methods, use cases, and trends in sentiment analysis specific to online product reviews. This study expands on current developments in sentiment analysis methods. For sentiment categorization, previous research has mostly used traditional machine learning models such as Naive Bayes and Support Vector Machines. Instead of depending only on conventional classifiers, this research presents some other methods that improve sentiment analysis performance. Building on the results of previous research, we create another comparison framework with the goal of pushing the envelope and achieving new state-of-the-art outcomes in this area. The document's remaining sections are organized as follows: We address prior research on online review data in Section 2, along with the convergence of sentiment analysis and machine learning. We describe the proposed neural network and SVM framework in Section 3. In Section 4, we go into further detail about the tests we ran and the results we got using our model. Section 5 concludes with some last thoughts that summarize the research and future enhancements.

II. LITERATURE REVIEW

Sentiment analysis of product reviews has been an active area of research, with significant progress made in the last decade. In [1], Pang et al. published an early influential study applying machine learning for sentiment classification of movie reviews, kickstarting interest in the field. Since then, numerous papers have focused specifically on analyzing sentiments in online customer product reviews, given their business value. The task of feature-based opinion mining, which links opinion expressions to certain product attributes, was initially introduced by Hu and Liu's early work [2]. To assist with feature-level analysis, domain-specific lexical resources such as SentiWordNet [4] and Opinion Finder [3] were built. The task of extracting product attributes from reviews using relaxation labelling was taken on by Popescu and Etzioni [5]. Classifying sentiment polarity is a crucial issue in sentiment analysis [6,–8]. Sentiment analysis is the process of assigning a sentiment category like positive, negative, or neutral to a written text excerpt. The degree of categorization depends on the text being studied. The complete document is categorized into an overall sentiment at the document level. Sentence by sentence is examined for sentiment at the sentence level. Lastly, the sentiment is classified in relation to particular elements or characteristics in the text at the entity and aspect levels [9]. The general sentiment of the whole content is classified as either good or negative at its content level. Sentiment analysis is done at the sentence level, one sentence at a time, throughout the manuscript. More detailed classification of emotion towards particular textual elements or traits is possible at the entity and aspect levels. It focuses on the specific topics, things, or qualities that the opinions are favorable or unfavorable about. Pang and Lee [10] suggested that in order to enhance sentiment analysis, feature selection could involve eliminating objective statements and extracting subjective ones. They created a minimal cut text categorization technique to find subjective material. Based on Twitter data, Gann et al. [11] selected a collection of 6,799 tokens and assigned sentiment scores to each one. With a total sentiment index score known as the TSI (Total Sentiment Index), each token was assigned a positive or negative label. The sentiment classification characteristics of the tokenized dataset were utilized. In the field of employing unguided approaches to assess attitudes in product evaluations [12], These inquiries have focused on identifying and examining certain aspects of goods or services that have been brought to light in user evaluations and extracting sentiment information associated with such aspects. This method achieves less percent precision level in the context of aspect-oriented sentiment analysis systems. Real-time applications require more accuracy because inaccurate projections could negatively impact a company's reputation. Product review sentiment analysis frequently requires the use of feature extraction and visualization tools. These techniques try to identify important textual parts and provide them with a visual representation to make it easier to spot emotional patterns [13]. It's crucial to recognize a restriction, though, in that feature extraction and visualization may not fully capture the context and meaning of the evaluations despite their value. The caliber of the data used has a significant impact on how well these strategies work. Analyses that are incomplete or erroneous can arise from reviews that are poorly worded. To guarantee a comprehensive and accurate sentiment analysis of product evaluations, it is therefore necessary to combine feature extraction and visualization techniques with additional methodologies and approaches. Review sets are the first thing that is examined in the sentiment analysis process. The dataset used in the study covered in the article [14] was sourced from the official product websites. These reviews were first pre-processed, which was a process of eliminating superfluous words, conjunctions, auxiliary verbs, and punctuation. Naive Bayes and SVM algorithms were used to categorize the trained dataset after the preprocessing stage was finished. In some situations, there are a number of ten-to-fifteen-page surveys that need to be filled out. Given the significant time and effort required, reviewing every one of these questionnaires becomes both cognitively and practically impossible. Soon, the importance of audits will become clear. The author of [15] suggests a solution to this problem using a technique known as sentiment-based rating anticipation. In general, Sentiment analysis is the process of identifying, recognizing, and classifying user feelings or points of view on different services, films, issues with products, events, or any other feature and then rating them as neutral, positive, or negative [16]. When emotion is expressed as a polarity in computational linguistics, it is often treated as a classification task. However, the problem is considered a regression challenge if sentiment scores within a given range are used to represent mood. In [17], Cortis et al. (2017) reviewed the research that uses sentiment analysis as a tool for prediction or categorization. Cortis et al. (2017) discovered that when assessing feelings by assigning scores inside the interval $[-1, 1]$, there are times when forecasting is taken to mean prediction in certain circumstances and categorization in others. In order to tackle the problem of uncertainty surrounding classification and prediction, the authors present a novel approach that combines two evaluation methods to determine the similarity matrix. Wankhade et al. [18] carried out a thorough investigation of several methods for sentiment analysis, delineating four basic levels of sentiment analysis, investigating applications particular to domains and industries, and pointing out several difficulties. The study addresses crucial processes in sentiment analysis and ranks different classification approaches in order of priority. Instead of concentrating only on essential processes, a thorough rundown of all feasible approaches is very helpful in determining which is best for particular sentiment analysis models. Furthermore, sentiment analysis applications are generally classified based on the industry or subject they cater to. Prospective application domains that are specified exclusively by the dataset are rarely explored in recent reviews. Some survey articles focus exclusively on one dimension or component of sentiment analysis. Multimodal sentiment analysis, along with adjacent subjects, applications, difficulties, and prospects, were the main topics of discussion for Kaur and Kautish [19]. Verma [20] talked about employing sentiment analysis to build a progressive society that prioritizes public facilities. The author illustrated how there are a lot of opportunities for better public services when future investigative tracks and developments in sentiment analysis are grasped for the benefit of society. The results in [22] proved that deep learning is the best method for sentiment analysis. The use of the sigmoid function for activation, the weight learning process in neural networks, the addition of certain convolutional layers, and pooling procedures are all covered in detail in this study. Source [23,24] presented a deep neural network structure designed specifically for sentiment analysis of Brazilian Portuguese responses on YouTube videos in their research. This six-layer network uses operations such as pooling and convolutional layers to achieve an accuracy range of 60% to 84%. To get good results, the data should be collected properly this can be understood from the research [25-27]. Finally, product review sentiment analysis has rapidly evolved with sophisticated NLP and ML techniques focused on tackling domain-specific issues like aspect extraction, comparative opinions, explainability, and cross-domain portability. Key opportunities remain in finer-grained feature-based analysis, generalized models, and multi- modal methods leveraging the wealth of review data available online.

III. PROPOSED METHODOLOGY

Machine learning-based sentiment analysis makes use of a repository that holds sentiment-oriented phrases that include both positive and negative descriptors. The lexicon in the repository is compared with the vocabulary found in user comments to enable evaluation. The system determines the product's quality through this comparison procedure and classifies it as positive, negative, or none. Computational linguistics methods for sentiment analysis go beyond simple sentence-level word detection. It involves matching feelings to things and correctly interpreting the tone, which includes the subtle emotional undertones of a sentence. We first use a few simple algorithms on the user-provided product reviews before using more complex algorithms like Naive Bayes or Support Vector Machines. Following these inspections, the system receives the data that has been extracted and is ready for further data mining and analysis. This is the first and fundamental step in the sentiment analysis process. Effective preprocessing is a crucial step in the process that is strongly dependent on text data analysis. Reviews sometimes include redundant and superfluous information, which makes analysis more difficult. During preprocessing, data normalization acts as a filtering mechanism. Preprocessing includes actions like normalization, tokenization, stop word removal, removing extra spaces, padding, converting text to lowercase, and removing hashtags and hyperlinks. Several tasks were carried out in this study to make sure the data was structured appropriately. Gathering data is the first step in the sentiment analysis process. For this investigation, we make use of an online corpus of product reviews pertaining to apparel products. Every review includes information like rating, the title of the review, and its content. Every evaluation follows a 5-star rating system; therefore, ratings are between 1 and 5 stars; half-star or quarter-star ratings are not available. The dataset is partitioned into two portions called the training and testing dataset. The training set accounts for 80% of the entire dataset, with the testing set making up the remaining 20%. The training dataset acts as a basis to illustrate to our model how to make accurate predictions on the test dataset. To demonstrate, the training subset includes product feedback from customers, which is divided into three categories: positive, negative, and neutral. Following that, we use this dataset to properly train our model, resulting in accurate forecasts on the testing subset. In this research, the following techniques are applied to the review data.

3.1. Gaussian Naive Bayes: Gaussian Naive Bayes for sentiment analysis are described as given below.

•Probability-Based Approach: Gaussian Naive Bayes is a probabilistic method used by machine learning for classification problems such as sentiment analysis.

•Assumption of Independence: It assumes that the text's features (words or tokens) are conditionally independent provided the sentiment category (positive, negative, neutral). Gaussian Distribution: The Gaussian Distribution asserts that continuous features (such as word frequencies) follow a Gaussian (normal) distribution.

•Feature Extraction: In sentiment analysis, every record (or piece of text) can be viewed as a collection of features, such as word frequencies or TF-IDF scores.

•During the training phase, the method computes the mean and variance of each feature (word frequency) for each class (positive, negative, and neutral sentiment) using the training data.

•Classification Phase: When categorizing a new document, the method will use Bayes' theorem and the Gaussian probability density function to compute the likelihood that it belongs to each class.

•Decision Rule: It selects the class with the biggest probability as the anticipated sentiment for the document. Gaussian Naive Bayes is highly computational and performs well on small to medium-sized datasets. However, it represents a naive assumption about feature autonomy, which may not be true in all circumstances. It is often used in sentiment analysis tasks that involve categorizing text into predefined sentiment groups (positive, negative, or neutral) built on the presence of specific words or attributes. While Gaussian Naive Bayes can perform well in many sentiment analysis operations, its performance may suffer if the characteristics do not correspond to a Gaussian distribution or if there are significant connections between them.

3.2 Multinomial Naive Bayes: Multinomial Naive Bayes is a common sentiment analysis technique due to its ability to handle text data. Following is how it is applied to sentiment analysis on a review database:

•Preprocessing: Before running the Multinomial Naive Bayes method, it is needed to preprocess the text input. Usually entails procedures like eliminating punctuation, changing text to lowercase, tokenization, deleting stop words, and potentially stemming or lemmatization.

•Feature Extraction: Transform the text data into numerical feature vectors that may be passed to the Multinomial Naive Bayes method.

•Training the model: Divide your dataset into a training and test set. Using the training data, train the Multinomial Naive Bayes model. The algorithm will learn the likelihood of each word appearing in each category.

•Model Evaluation: Use the test set to evaluate the trained model's performance. Accuracy, precision, recall, and F1-score are some of the most used sentiment analysis evaluation criteria.
•Prediction: Once trained and assessed, the model may be used to forecast the sentiment of new reviews.
Multinomial Naive Bayes is a basic and efficient technique, but it may fail to capture more complicated correlations in data when compared to more sophisticated techniques such as deep learning. However, it might still be useful as a baseline for sentiment analysis operations, particularly with smaller datasets.

3.3. Logistic regression

Logistic regression is a machine learning technique that performs binary classification tasks like sentiment analysis. Here is how logistic regression operates in sentiment analysis.
•Data Preparation: Begin by gathering a dataset of reviews and their preprocessing.
•Feature Extraction: Next, it will turn the processed text data into numerical characteristics which can be fed into the logistic regression model.
•Model Training: after getting feature vectors and tags, it trains the logistic regression algorithm on the dataset. The method of logistic regression becomes adept at mapping input data to a binary sentiment tag by minimizing.
•Prediction: Upon training, the logistic regression approach may forecast the sentiment of new, previously unseen reviews. Given the numerical characteristics of a review, the model generates a probability score, which is used to predict the label of the review.
•Evaluation: This allows you to measure how well the model generalizes to new data and whether it accurately represents the sentiment of the reviews.
Overall, logistic regression is a straightforward but effective approach for sentiment analysis tasks, particularly when the sentiment classes are binary (positive or negative).

3.4. Support Vector Machines: The following are the essential aspects demonstrating how Support Vector Machines (SVMs) can be used for sentiment analysis on a customer review database. Classification: SVMs are a sort of supervised training technique utilized in classification jobs. Sentiment analysis aims to categorize text data as positive, negative, or neutral. Feature Extraction: When using SVMs, the text input must be pre-processed and translated to numerical characteristics. Every evaluation is often exemplified as a vector of numerical values using approaches. Margin Maximization: SVMs function by identifying the hyperplane which most effectively divides the various classes in the feature space while maximizing the separation among them. In sentiment analysis, this includes determining the decision boundaries that most effectively distinguish positive and negative evaluations. Kernel method: The kernel method allows SVMs to deal with nonlinear boundaries for decisions. It enables them to implicitly translate the characteristics of the input to a higher-dimensional environment where a linear separation may be conceivable. Model Evaluation: Upon learning the SVM on a labelled set of customer reviews, its efficacy should be measured using measures like accuracy, precision, recall, and F1-score.

3.5. Neural network approach: Neural networks analyze the data and find patterns by simulating brain activities. The decisions that are made are based on these patterns. It is special because it doesn't require explicit programming instructions; instead, it can learn on its own and improve over time.

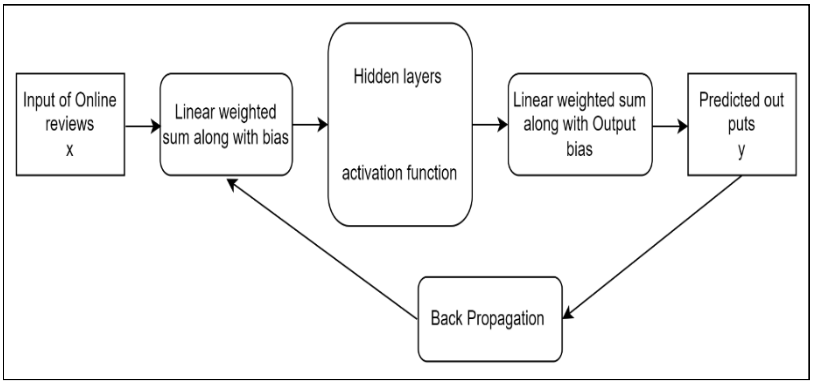

Figure 3.1: Neural Network procedure.

In this research, a neural network structure designed for processing the dataset is presented, as shown in Figure 3.1. First, the information is converted into word vectors. According to [21], Keras's embedding layer offers the potent Word2Vec approach. Two convolutional processing layers are included in our model. Identifying characteristics in the input data and creating a feature map are the tasks of the first 1D convolutional layer. The characteristics found by the first layer are then further refined by a second convolutional layer. The global max-pooling layer lowers the decision of the output features to prevent overfitting.

3.6. Proposed meta model:

The proposed metamodel combines Neural network and Support Vector Machine (SVM) techniques to perform sentiment analysis which gains the rewards of both methods. Neural networks are excellent at identifying intricate patterns in data, whereas support vector machines (SVM) are well-known for their facility to manage high-dimensional data and classification tasks. In the present method, we use a neural network as the first classifier to classify the sentiment of text data and extract the corresponding probabilities. The neural network's output is then sent into an SVM, which learns complex correlations and subtleties in the data to further refine the classification results, as shown in Figure 3.2.

Algorithm:

Let D be the input database, and D_(New) be the new database created.
Input: Dataset D of online reviews.
Output: Sentiment of unseen data after applying the meta-model.

1. Train the Neural Network using the data set D.
2. For (each $r_x \in D$)
3. Compute the probable values $r_{Positive}$, $r_{Negetive}$, $r_{Neutral}$

using Trained Neural Network.
4.D_{New} $\leftarrow$ ($r_{Positive}$, $r_{Negetive}$, $r_{Neutral}$ sentiment tag of r_x)
5. End for
6. Train the SVM model using D_{New}.
7. Predict the sentiment of unseen data using trained SVM.

Sentiment analysis tasks by combining SVM and neural network components, which are described in the algorithm, successfully capture both the local and global relationships within the data. By utilizing the complementing advantages of both methods, this hybrid methodology produces sentiment analysis results that are more accurate and dependable.

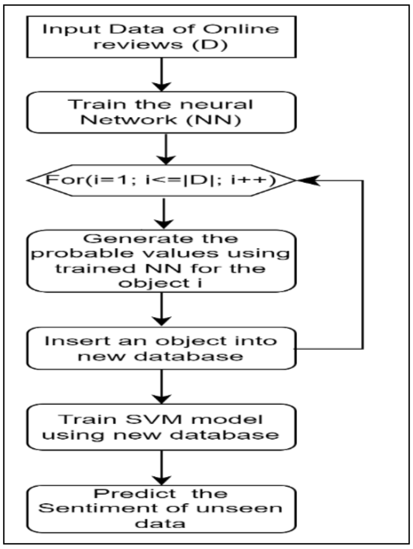

Figure 3.2: The flow of the proposed meta-model

In essence, we are employing the likelihood estimates or confidence scores of a neural network (NN) as representations of the data points when we utilize the output probabilities of an NN as input features for a Support Vector Machine (SVM). The SVM uses the NN's output probabilities as features. Here it may think of each output probability as a feature dimension. Consequently, we would have N features indicating the odds of belonging to each class if the NN has N output classes. Here, to do the Information Aggregation the probabilities show how confident the NN is in its ability to classify each input into distinct groups. This implies that each forecast has a level of uncertainty or confidence attached to it rather than merely being a hard choice.

IV. RESULT ANALYSIS

Sentiment analysis, often called opinion mining, is a natural language processing (NLP) technique that analyses and categorizes text material based on the sentiment conveyed within it. In this case, the purpose is to comprehend the attitude or viewpoint expressed in customer reviews. It should contain the steps required to preprocess text data, extract features, select models, and evaluation of results.

Data set:

Collecting data is the first stage of sentiment analysis. This research makes use of an assortment of product reviews that concentrate on apparel products that are obtained from internet resources. There are roughly 23,000 entries in the dataset, which is divided into two parts: the training dataset and the test dataset. Eighty per cent of the total dataset is the training subset, while the remaining twenty per cent is the testing subset. Each assessment includes the following information: 1) Title review, 2) review, and 3) content evaluation. All ratings are given on a 5-star scale, which means that only entire stars (from 1 to 5) are included; half-star and quarter-star ratings are not included. Our model learns to correctly predict values in the test dataset with the aid of the training dataset. One example of this is the training set, which consists of consumer feedback labelled as neutral, negative, and positive. After the model has been trained with this data, it can accurately predict the test data. The algorithms that were employed for sentiment analysis in the present study are some classic machine learning algorithms such as Support Vector Machines (SVM), logistic regression, gaussian and nominal naïve Bayes. The results obtained using these are compared with the proposed neural network-based sentiment analysis for customer reviews.

Naive Bayesian classifiers, which are often used in text classification applications, perform better than other techniques. As a result, they are widely used in applications such as spam filtering and sentiment analysis, which help identify positive and negative consumer sentiments. These classifiers determine the conditional probability of event A based on the individual probabilities of A and B, as well as the conditional probability of event B. This strategy presupposes that the features are independent of one another. The results obtained from these methods are described in Table 2. In logistic regression, the result variable is categorical, which means it has discrete values. Rather than directly predicting class labels, logistic regression calculates the likelihood of an event occurring. These probabilities are always between 0 and 1, indicating the chance of occurrence for the given event. The Logistic Regression confusion matrix graph is described in Figure 4.1.

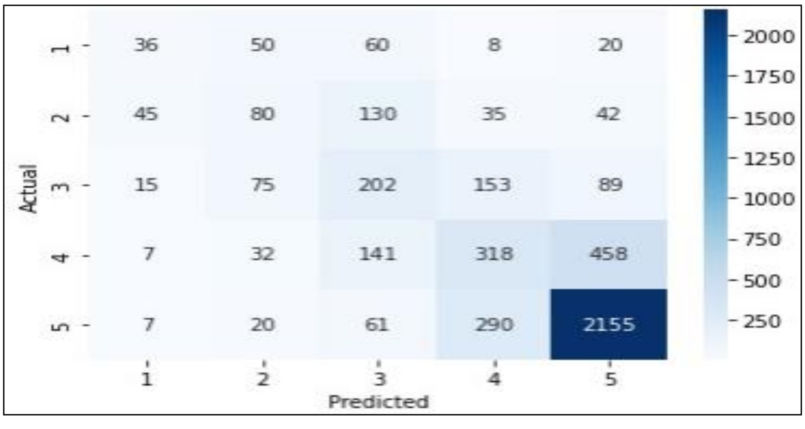

Figure 4.1: Logistic Regression confusion matrix graph

The Gaussian methodology of Naive Bayes had a poor accuracy of 23%. To enhance accuracy, the Multinomial Naive Bayes method was implemented on the same data set, resulting in three times better performance than

Gaussian. Subsequently, to enhance precision, the dataset underwent implementation of the logistic regression technique, yielding satisfactory accuracy levels comparable to those of Naive Bayes. However, this parity prompted the exploration of more sophisticated classification methodologies.

Evaluation method	Accuracy
Gaussian Naive Bayes	23.10
Support Vector Machines	58.20
Multinomial Naive Bayes	63.45
Logistic Regression	61.64
Neural Network approach	83.3
Proposed meta model	86.89

Table 2: Accuracy of different methods

In the proposed meta-model, the probabilities show how confident the NN is in its ability to classify each input into distinct groups. This implies that each forecast has a level of uncertainty or confidence attached to it rather than merely being a hard choice. If probabilities are used as features instead of just hard forecasts, the SVM may be able to learn more. Improved accuracy results are obtained from this when compared to other individual methods, particularly when the NN's ambiguity estimations closely match the actual uncertainty in the data. For a better understanding of the results, they are represented in Figure 4.2.

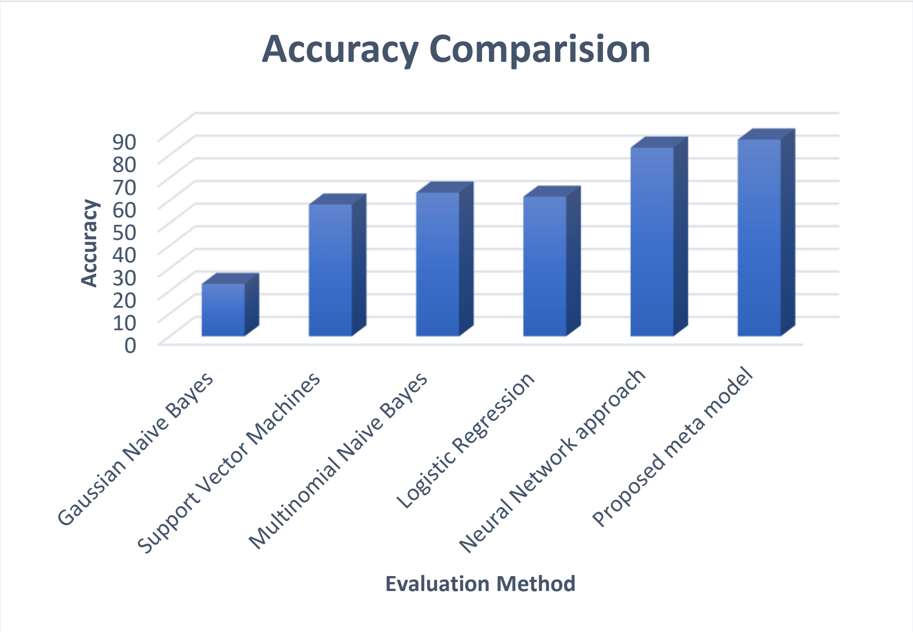

Figure 4.2: Comparison of different evaluation measures

In summary, the Multinomial Naive Bayes classifier proved to be the most accurate, closely followed by Logistic Regression; for this specific dataset, Gaussian Naive Bayes did the least well. In order to better optimize the classifiers, additional analysis can entail examining the causes of the performance variations and possibly experimenting with various feature representations or parameter values. Finally, it might be helpful to use an NN's output likelihood as input features for an SVM, particularly if the goal is to make use of both NNs' feature extraction skills and SVMs' classification abilities. However, it's important to be aware of possible hazards, including computing complexity and calibration problems.

V. CONCLUSION

Customers now rely far more on online evaluations because of the shift from conventional brick-and-mortar markets to digital platforms. These evaluations now play a crucial role in fostering trust and have a significant impact on the purchase decisions of customers. As a result of this growing dependence, there is an increasing need to handle the large volume of online reviews effectively to provide customers with reliable information. By analyzing the mood of online product reviews and classifying them as good or negative, our study aims to achieve this goal. Three categorization models were used once a balanced dataset with about equal shares of neutral, positive and negative evaluations was obtained. The proposed meta-model showed the best prediction accuracy when compared with traditional machine learning and deep learning strategies. Because of the great complexity of the algorithm, traditional neural networks are prone to converge to local optima when adjusted backwards, which reduces the accuracy of prediction. It is necessary to do further study to investigate deep learning optimization approaches that use a variety of datasets to classify reviews according to different factors, including satisfaction, preferences, and suggestions.

VI Acknowledgment & Authors contribution

We express our sincere gratitude for the combined effort and unwavering dedication of all those who contributed to this research work. Following is the contribution of all the authors.

Atluri Vani Vathsala: Conceptualization, methodology, investigation, writing - original draft preparation.

Lakshmi H N: Supervision, project administration, writing - review and editing.

LNC.Prakash K: Software, formal analysis, validation.

Thaduri Venkata Ramana: Visualization, resources, funding acquisition.

Funding: No funding.

Availability of data and materials: Data will be made available on reasonable request.

Code availability: Available on reasonable request.

Declarations Conflict of interest: There is no conflict of interest.

References

[1] Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. "Thumbs up? Sentiment classification using machine learning techniques." arXiv preprint cs/0205070 (2002).

[2] Hu, Minqing, and Bing Liu. "Mining and summarizing customer reviews." Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. 2004.

[3] Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, pages 347–354, Vancouver, British Columbia, Canada. Association for Computational Linguistics.

[4] Andrea Esuli and Fabrizio Sebastiani. 2006. SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining. In Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06), Genoa, Italy. European Language Resources Association (ELRA).

[5] Ana-Maria Popescu and Oren Etzioni. 2005. Extracting Product Features and Opinions from Reviews. In Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, pages 339–346, Vancouver, British Columbia, Canada. Association for Computational Linguistics.

[6] Choi Y, Cardie C (2009) Adapting a polarity lexicon using integer linear programming for domain-specific sentiment classification In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2 - Volume 2, EMNLP '09, 590–598. Association for Computational Linguistics, Stroudsburg, PA, USA..

[7] Jiang L, Yu M, Zhou M, Liu X, Zhao T (2011) Target-dependent Twitter sentiment classification In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1, 151–160. Association for Computational Linguistics, Stroudsburg, PA, USA..

[8] Tan LK-W, Na J-C, Theng Y-L, Chang K (2011) Sentence-level sentiment polarity classification using a linguistic approach In: Digital Libraries: For Cultural Heritage, Knowledge Dissemination, and Future Creation, 77–87. Springer, Heidelberg, Germany.

[9] Liu B (2012) Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers.

[10] Pang B, Lee L (2004) A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts In Proceedings of the 42Nd Annual Meeting on Association for Computational Linguistics, ACL '04. Association for Computational Linguistics, Stroudsburg, PA, USA.

[11] Gann W-JK, Day J, Zhou S (2014) Twitter analytics for insider trading fraud detection system In Proceedings of the second ASE international conference on Big Data. ASE.

[12] R. Sharma, S. Nigam, and R. Jain, "Mining of Product Reviews at Aspect Level," International Journal in Foundations

of Computer Science & Technology, vol. 4, no. 3, pp. 87–95, May 2014, doi: 10.5121/ijfcst.2014.4308.

[13] Bhatt, A. Patel, H. Chheda, and K. Gawande, “Amazon Review Classification and Sentiment Analysis,” International Journal of Computer Science and Information Technologies, vol. 6, no. 6, pp. 5107– 5110, 2015.

[14] J. Sadhasivam and R. B. Kalivaradhan, “Sentiment Analysis of Amazon Products Using Ensemble Machine Learning Algorithm,” International Journal of Mathematical, Engineering and Management Sciences, vol. 4, no. 2, pp. 508–520, Apr. 2019, doi: 10.33889/IJMEMS.2019.4.2-041.

[15] MLMCE, India and K. K. Thomas, “Sentiment Analysis in Product Reviews using Natural Language Processing and Machine Learning,” International Journal of Information Systems and Computer Sciences, vol. 8, no. 2, pp. 147–151, Apr. 2019, doi: 10.30534/ijiscs/2019/35822019.

[16] Mehta, Pooja, and Sharnil Pandya. "A review on sentiment analysis methodologies, practices and applications." International Journal of Scientific and Technology Research 9.2 (2020): 601-609.

[17] Cortis K, Freitas A, Daudert T, Huerlimann M, Zarrouk M, Handschuh S, Davis B (2017) Semeval-2017 task 5: Fine-grained sentiment analysis on financial microblogs and news. In: 11th International workshop on semantic evaluations (SemEval-2017): Proceedings of the workshop, Stroudsburg, PA, USA. Association for Computational Linguistics (ACL), pp 519–535.

[18] Wankhade M, Rao ACS, Kulkarni C (2022) A survey on sentiment analysis methods, applications, and challenges. Artif Intell Rev 1–50.

[19] Kaur R, Kautish S (2022) Multimodal sentiment analysis: a survey and comparison. In: Research anthology on implementing sentiment analysis across multiple disciplines. IGI Global, pp 1846–1870.

[20] Verma, Sanjeev. "Sentiment analysis of public services for smart society: Literature review and future research directions." Government Information Quarterly 39.3 (2022): 101708.

[21] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” Advances in Neural Information Processing Systems, vol. 26, 2013.

[22] K. Chakraborty, S. Bhattacharyya, R. Bag, and A. A. Hassanien, “Sentiment analysis on a set of movie reviews using deep learning techniques,” Social Network Analytics: Computational Research Methods and Techniques, vol. 127, 2018.

[23] A.A. L. Cunha, M. C. Costa, and M. A. C. Pacheco, “Sentiment analysis of YouTube video comments using deep neural networks,” International Conference on Artificial Intelligence and Soft Computing, China, 2019.

[24] D. S. Sisodia, S. Bhandari, N. K. Reddy, and A. Pujahari, “A comparative performance study of machine learning algorithms for sentiment analysis of movie viewers using open reviews,” Performance Management of Integrated Systems and its Applications in Software Engineering, Springer, Salmon Tower Building, NY, USA, pp. 107–117, 2020.

[25] K. Chakraborty, S. Bhatia, S. Bhattacharyya, J. Platos, R. Bag, and A. E. Hassanien, “Sentiment Analysis of COVID-19 Tweets by Deep Learning Classifiers—A Study to Show how popularity Affecting accuracy in social media,” Applied Soft Computing, vol. 97.

[26] Dodda, Ratnam & Maddhi, Sunitha & Thuraab, Mohammed & Reddy, Adapala & Chandra, Avadanula. (2023). NLP-Driven Strategies for Effective Email Spam Detection: A Performance Evaluation. 275-279. 10.1109/ICSCNA58489.2023.10370223.

[27] P. C. Shaker Reddy, G. Suryanarayana, L. P. K and S. Yadala, "Data Analytics in Farming: Rice price prediction in Andhra Pradesh," 2022 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT), Aligarh, India, 2022, pp. 1-5, doi: 10.1109/IMPACT55510.2022.10029009.