International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN: 2147-6799 www.ijisae.org Original Research Paper

Implementation of User Rating Classification for Amazon Food Review Dataset Using SVM and LSTM

Krishan Kumar[1*], Dr. Randeep Singh[2]

Submitted: 07/02/2024 **Revised:** 15/03/2024 **Accepted:** 21/03/2024

Abstract: This research investigates the challenge of classifying user ratings for Amazon food reviews using Support Vector Machines (SVM) and Long Short-Term Memory (LSTM) neural networks. The objective is to forecast the sentiment or user rating categorization of food reviews in order to provide important information for both consumers and vendors on the network. The dataset comprises textual reviews and their related user ratings collected from the Amazon food goods category. A train-test split is conducted in order to train the Support Vector Machine (SVM) model using the training dataset and adjust its hyperparameters to achieve optimal performance. In the context of Long Short-Term Memory (LSTM), the neural network is trained by using the training set and incorporating strategies such as dropout and early stopping to mitigate the issue of overfitting. The empirical findings demonstrate that both Support Vector Machines (SVM) and Long Short-Term Memory (LSTM) models exhibit a notable level of precision when used for the purpose of forecasting user ratings in the context of Amazon food reviews. Support Vector Machines (SVM) have exceptional performance in managing datasets that are both sparse and high-dimensional. On the other hand, Long Short-Term Memory (LSTM) networks are very proficient at capturing contextual connections within textual data. The results provide significant insights for organisations about customer satisfaction and sentiment patterns, enabling them to make informed choices based on data to enhance product offerings and improve customer experiences. In addition, prospective consumers might get advantages from the precise sentiment analysis while evaluating food acquisitions on the Amazon platform.

Index Terms—*SVM, LSTM, User rating classification, sentiment analysis;*

I. Introduction

The proliferation of online marketplaces in recent decades has led to an increased inclination among sellers and merchants to actively seek feedback from their consumers. Every day, a vast number of internet reviews pertaining to a wide range of products, services, and destinations are published by users. As a consequence, the Internet is increasingly emerging as the predominant channel via which customers access information pertaining to a particular service or product. The issue is in the growing number of individuals evaluating a product, which subsequently poses challenges for potential consumers in making well-informed decisions. Customers experience heightened confusion when they see disparate evaluations of a given product, alongside intentionally ambiguous reviews. It is evident that doing content analysis would provide significant advantages for any e-commerce companies.

** Department of Computer Science & Engineering IEC University Baddi, Solan, HP India*
Email:-krishan.pingle@gmail.com
Department of computer Science& Engineering IEC University Baddi, Solan, HP
Email:- randeeppoonia@gmail.com

***Correspondence Author:**- Krishan Kumar,*
**Department of Computer Science & Engineering, IEC University Baddi, Solan, HP, India ,*
Email:- krishan.pingle@gmail.com

Analysis of user sentiments, attitudes, and emotions gleaned via text mining is known as sentiment analysis [1]. The financial market [2, 3], news articles [4, 5], and travel and tourism [6] are only a few examples of additional fields where sentiment analysis might be useful. According to [6], it's possible to classify sentiment analysis as such. Sentiment analysis has been made more effective by the development of several algorithms.

User rating classification is the process of categorizing user ratings (e.g., 1 to 5 stars) into different classes based on their sentiment or satisfaction level. It is a common machine learning task used in various domains, including e-commerce, product reviews, movie ratings, and customer feedback analysis.

Amazon food reviews refer to customer feedback and reviews posted on the Amazon platform specifically related to food products. Amazon is a popular e-commerce platform that sells a wide range of products, including groceries, snacks, beverages, and other food items. Customers who purchase these food products have the option to leave reviews and ratings based on their experiences with the products.

These food reviews can be valuable for both other customers and sellers. Potential buyers can use these reviews to make informed decisions about whether to purchase a particular food product based on the opinions and experiences of previous customers. For sellers, reviews provide feedback on the quality, taste, packaging, and other aspects of their food products, which can help them understand customer preferences and make improvements if necessary.

When analyzing Amazon food reviews, sentiment analysis and user rating classification are common techniques used to extract insights from the text data. These approaches can help identify positive and negative sentiments in the reviews and classify them into different rating categories to understand the overall satisfaction level of customers for specific food products.

II. Literature Review

The impact of consumer evaluations displayed on websites on customers' selections has been shown to be considerable (Schoenmueller et al., 7). In the contemporary competitive market, enterprises that neglect to allocate sufficient time and effort towards the analysis and assimilation of these assessments are destined to experience failure. In contemporary business practises, sentiment analysis has emerged as a strategic tool used by enterprises to gain a competitive advantage inside the marketplace. A vast dataset including over 280 million reviews has been aggregated for the purpose of analysis. The dataset was comprised of over twenty-four million reviews sourced from a diverse range of twenty-five distinct platforms, including prominent ones like Amazon and Yelp. The evaluations included a range of commodities and services. The majority of assessments were seen to have reduced extremity and a more balanced distribution. Several factors, such as the rating scale, company strategy, and review frequency, have the potential to influence the consistency of ratings. Companies such as Amazon, who want to generate profits from online product assessments, must leverage the aforementioned factors. It is crucial to comprehend that corporations use sentiment research extensively in order to enhance internal operations and foster consumer loyalty.

Karamitsos et al. [8] assert that the use of sentiment analysis helps firms in gaining insights into consumers' perceptions of their products and services. This enables enterprises to enhance their advertising strategy via improved planning. Moreover, sentiment analysis plays a crucial role in enabling modern businesses to optimise their word-of-mouth advertising initiatives. Organisations that prioritise maintaining a competitive edge use text mining methodologies to enhance their understanding of their consumers' viewpoints. Text mining enables organisations to extract valuable insights from many forms of written information, such as blog entries, news articles, and other textual sources.

In Zhao's work [9], the author provided a detailed description of the procedure for using the user Timeline() function to get tweet content. Hence, the effective use of online customer evaluations by enterprises necessitates the integration of sentiment analysis, text mining, and several other technological tools. According to the study conducted by Jain et al. [10], using sentiment extraction as a means to get understanding of consumers' online assumptions proved to be a productive approach. One potential avenue for enhancing advertising strategies is via the use of data obtained from social media platforms and consumer review websites. The purchasing decisions of customers are also impacted by the evaluations. Amazon and other companies use this data to get a strategic edge in formulating business strategies.

Lim et al. [11] found that prominent merchants in the United States largely depend on online product evaluations as a means to enhance their marketing tactics and operational processes. In the event that a product repeatedly receives subpar evaluations, the manufacturer will undertake an investigation to determine the underlying causes. In the event that customers express dissatisfaction over either the pricing or quality of a product or service, the company promptly takes action to resolve these issues.

The study conducted by Sharma et al. (2013) examined the influence of customer reviews on the sales of books on the Amazon platform. The findings of the study indicate that online evaluations are regarded by customers with a level of confidence that is comparable to that of personal recommendations. Reviews are seen to be more handy and useful for clients due to this particular reason. Based on the findings, it can be inferred that online reviews have a significant influence on customer decision-making processes and product pricing. The findings support the conclusions drawn in a previous research conducted by Chong et al. [14], which focused on the relationship between online reviews and attitudes. The study also examined the credibility of online reviews. The observers have also recognised the inherent uncertainty around the impact of online reviews on businesses. The impact of internet reviews on sales has been shown to vary significantly across different research, with some indicating a substantial influence and others suggesting a more modest effect. Factors such as the sorts of products and qualitative textual features may also have an influence.

Du et al. [15] conducted a comprehensive analysis of a dataset including 142.8 million customer evaluations from Amazon. The study examined the summary headline, product remarks, and helpfulness information as criteria for distinguishing between helpful and unhelpful reviews. To enhance the dependability of their findings, the researchers excluded any product reviews that were blank or written in a language other than English. Only the candidates that received the highest number of votes were selected. The study's findings indicate that the e-commerce industry now relies significantly on the analysis of online product reviews specifically on the Amazon platform. Reviews that possess the qualities of being informative and useful are characterised by the inclusion of particular details on a product or service, which are derived from genuine customer input [16]. Customers rely on evaluations that have received the highest number of votes in order to make well-informed purchasing decisions, using the available facts.

Govindaraj and Gopalakrishnan (17) assert that an organisation might get insights into consumers' viewpoints by examining assessments of its items. The assessment of a customer's pleasure with a particular product or service might potentially be discerned via an online review. Nevertheless, the majority of product assessments fail to adequately capture the extent of client satisfaction. Consequently, a poll was conducted in order to categorise the degree of customer happiness, as shown by online reviews. A technique was devised to identify the degrees of customer satisfaction by using both linguistic and aural clues. A proposed method was put out whereby evaluations would be categorised and graded in descending order of favorability, ranging from most favourable to neutral to negative. The study conducted by Ghasemaghaei et al. (2018) about the impact of review length and online attitudes aligns with the results obtained in the present study. Prior to finalising their purchase choices, the majority of buyers engage in the practise of perusing many product reviews [19]. Hence, it is essential for companies to do thorough research to ascertain the actual extent of customer satisfaction with their products and services, enabling them to make well-informed decisions by considering online product reviews.

In a study conducted by Rui Xia et al. [20], the researchers examined and contrasted the efficacy of the ensemble approach in the context of sentiment classification inside a research job. The use of the ensemble framework is employed to enhance the precision of the classification process, particularly in sentiment classification jobs. This framework incorporates many function sets and classification algorithms, hence facilitating an efficient approach. Nevertheless, a comparative chart illustrating the distinctions between Support Vector Machine (SVM) and Naive Bayes classifications was not created by them.

Using machine learning techniques, Geetika Gautam et al. [21] examined the polarisation of Twitter's sentiment dataset. In the long run, picking the right features and pre-processing method will help them achieve even higher accuracy.

III. Dataset

More than ten years and 568,454 reviews of high-quality meals purchased from Amazon are included in the dataset. A review will include a star rating, details on the product and the reviewer, and a plain text description. Reviews from every other Amazon department are included as well.

Table 1: Attributes of Amazon food review data

Name	Description
ProductId	unique identifier for the product
UserId	unique identifier for the user
HelpfulnessNumerator	number of users who found the review helpful
HelpfulnessDenominator	number of users who indicated whether they found the review helpful
Score	rating between 1 and 5
Time	timestamp for the review
Summary	brief summary of the review
Text	text of the review

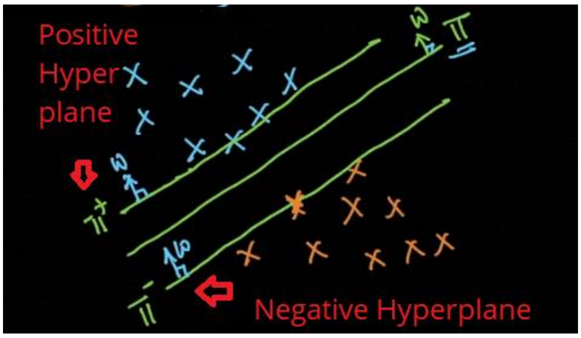
IV. Proposed Work

Support Vector Machines (SVM)

The term "Support Vector Machine" (SVM) gained popularity in 1992, thanks to the contributions of Boser, Guyon, and Vapnik in their work presented at COLT-92. Support vector machines (SVMs) are a collection of closely related supervised learning algorithms that are often used for classification and regression tasks. Linear classifiers that have been generalised are a specific kind. The Support Vector Machine (SVM) is a machine learning technique that effectively performs classification and regression tasks by mitigating the risk of overfitting to the training data. The support vector machine (SVM) is a machine learning technique that incorporates a learning bias based on statistical learning theory. It is trained on the hypothesis space of linear functions inside a feature space of high dimensionality.

Support Vector Machines (SVMs) are widely recognised as very efficient methods for performing classification tasks. The objective is to identify an optimal hyperplane that effectively discriminates between the two groups. The concept of margin is an additional consideration that should be optimised in order to minimise needless overlap between the two groups [22]. Enhanced classification outcomes may be achieved by the use of non-linear mapping techniques to transform non-linearly separable data into a higher-dimensional space. When confronted with non-linear data, kernel functions such as the radial basis function (RBF) and polynomial functions have been shown to be advantageous (Smith, 2010).

Initially, we will assume that the positive and negative hyper-planes are positioned on opposing sides of the separating hyper-plane and are distinctly separated by a distance of one unit.



The Equation of separating Hyperplane is given by,

$$\pi : w^T.x + b = 0 \quad (1)$$

The equation for Positive Hyperplane is given by,

$$\pi^+ : w^T.x + b = 1 \quad (2)$$

The equation for Negative Hyperplane is given by,

$$\pi^{-} : w^{T}.x + b = -1 \quad (3)$$

The Margin is given by,

$$margin = \frac{2}{||w||} \quad (4)$$

Then For SVM, we can write a constraint optimization problem as,

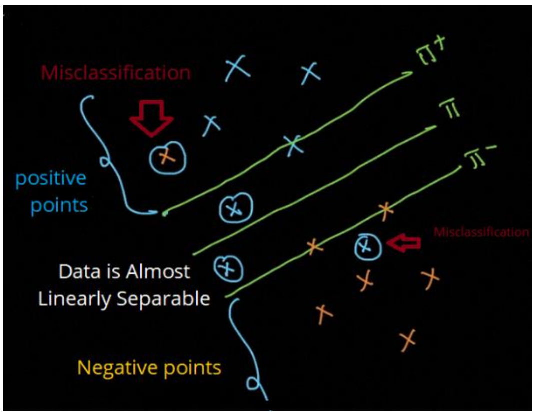
$$w^{*}, b^{*} = argmax_{w,b} \frac{2}{||w||}$$

$$such\ that\ \forall i,\ y_i\left(w^T x_i + b\right) \geq 1$$

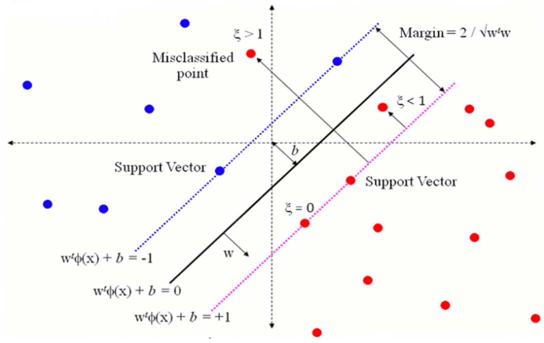
This Equation is also called as **Hard Margin SVM**.

The equation mentioned above necessitates the presence of linearly separable data. This implies that it is necessary for all positive points to be situated on one side of the plane, while all negative points should be located on the other side of the plane, without any data points falling inside the intermediate region.

In the context of real-world data, it is common for the data points to not exhibit complete linear separability. Consequently, a method known as soft margin classification is used to accommodate instances when some margin violation is allowed. Despite the violation of some regulations, a substantial margin is seen more desirable. If an individual selects a hyperplane that has a margin of error separating it from the actual hyperplane, it is plausible that some data points may get situated on an incorrect side of the hyperplane.



Given the Hard Margin Support Vector Machine (SVM) equation, it is not feasible to handle the issue at hand. Therefore, it is advisable to explore other implementations.



The equation of the separating hyperplane is provided by,

$$\pi : w^{T}.x + b = 0 \quad (1)$$

The equation representing a positive hyperplane is defined as follows,

$$\pi^{+} : w^{T}.x + b = 1 \quad (2)$$

The equation for Negative Hyperplane is given by,

$$\pi^{-} : w^{T}.x + b = -1 \quad (3)$$

The Margin is given by,

$$margin = \frac{2}{||w||} \quad (4)$$

Then For SVM, we can write a constraint optimization problem as,

$$w^{*}, b^{*} = argmin_{w,b} \frac{||w||}{2} + C.\frac{1}{n}\sum_{i=1}^{n}\xi_i$$

$$such\ that\ \forall i,\ y_i\left(w^T x_i + b\right) \geq 1 - \xi_i,\ \xi_i \geq 0$$

This Equation is also called as **Soft Margin SVM**.

ξ = some units of distance away from the correct hyperplane in the incorrect direction.

For correctly classified points $\xi_i = 0$ i.e if $y_i\left(w^T x_i + b\right) \geq 1$

For incorrectly classified points $\xi_i > 0$

In the context of Support Vector Machines (SVM), it is possible to formulate the issue as a constraint optimisation task.

Both the first part of the equation before the 'positive' sign and the second part after it are referred to as the' regularization' and the 'Hinge Loss', respectively.

Hyper-parameter C. Over fitting, or high variance, occurs when C rises, whereas under fitting, or high bias, occurs when C falls.

When C is big, the optimal solution may choose a hyper-plane with a narrower margin if it better ensures that all training points are properly labelled. A low value of C, on the other hand, will lead the optimizer to seek for a hyper-plane with a bigger margin of separation, even if doing so would lead to a higher number of misclassified points.

SVM with Bag of Words Technique

The Support Vector Machine (SVM) is a widely used machine learning method that is particularly suitable for

addressing binary classification tasks. The Bag-of-Words (BoW) methodology is a commonly used technique in the field of natural language processing (NLP) for the purpose of feature extraction. It finds extensive application in many NLP tasks, including sentiment analysis and text categorization. Text classification challenges generally include the use of Support Vector Machines (SVM) in combination with other methods.
The following is a comprehensive elucidation of the operational process of Support Vector Machines (SVM) using the Bag-of-Words methodology.

Data Preprocessing:
Clean and preprocess the raw text data from Amazon food reviews. This typically involves removing punctuation, converting text to lowercase, and handling any other necessary preprocessing steps.

Tokenization:
Tokenize the preprocessed text, splitting it into individual words or tokens. This is usually done using whitespace or more advanced tokenization techniques like the NLTK library in Python.

Vocabulary Creation:
Create a vocabulary, which is a unique set of all tokens present in the entire dataset. Each word in the vocabulary will correspond to a specific feature in the Bag-of-Words representation.

Document-Term Matrix (DTM):
In order to characterise each review in the dataset, a Document-Term Matrix (DTM) is constructed. This DTM uses a vector of word frequencies. The Document-phrase Matrix (DTM) is structured in a manner that each row corresponds to a review, while each column corresponds to a certain vocabulary phrase. The entries in the Document-phrase Matrix (DTM) provide information on the frequency of occurrence for each phrase inside each individual review.

Bag-of-Words Representation:
Convert the DTM into the Bag-of-Words representation. In the BoW representation, each review is represented as a sparse vector, where non-zero values indicate the frequency of each word present in the review.

Data Labeling:
Label the Amazon food reviews as positive or negative based on their sentiment. Positive reviews are assigned a label of 1, and negative reviews are assigned a label of 0.

Model Training:
The Support Vector Machine (SVM) model is trained using the Bag-of-Words representation of the reviews and their respective sentiment labels. The Support Vector Machine (SVM) algorithm tries to identify the most favourable hyperplane that effectively distinguishes positive and negative ratings inside the feature space.

Evaluation of the Model:
The performance of the trained Support Vector Machine (SVM) model in predicting the sentiment of Amazon food reviews was assessed by evaluating it on a distinct testing set. Compute several assessment measures, including accuracy, precision, recall, and F1 score, among others.

Prediction for New Reviews: After the completion of training and evaluation of the Support Vector Machine (SVM) model, it becomes feasible to use this model for predicting the sentiment of novel and previously unseen Amazon food reviews. This prediction process involves transforming the textual content of the new review into the Bag-of-Words representation and then using the trained SVM model to classify the sentiment.

SVM with TFIDF Technique
SVM (Support Vector Machine) with TF-IDF (Term Frequency-Inverse Document Frequency) technique is a popular combination for text classification tasks, including sentiment analysis and other natural language processing applications. TF-IDF is a feature extraction method that enhances the Bag-of-Words representation by considering the importance of words in the context of the entire document collection.
Here's how SVM with TF-IDF technique works for sentiment analysis of Amazon food reviews:
Data Preprocessing:
Clean and preprocess the raw text data from Amazon food reviews. This typically involves removing punctuation, converting text to lowercase, and handling any other necessary preprocessing steps.

Tokenization:
Tokenize the preprocessed text, splitting it into individual words or tokens. This is usually done using whitespace or more advanced tokenization techniques like the NLTK library in Python.

TF-IDF Calculation:
Compute the term frequency-inverse document frequency (TF-IDF) scores for every word included in the lexicon. The TF-IDF metric quantifies the significance of a word inside a specific document (such as a review) in relation to the whole of the document collection. The approach involves the integration of Term Frequency (TF), which represents the frequency of a term inside a specific document, and Inverse Document Frequency (IDF), which imposes a penalty on words that occur often over the whole of the document collection.

In order to create a Document-Term Matrix (DTM) that appropriately represents each review in the dataset, it is essential to convert the reviews into vectors containing Term Frequency-Inverse Document Frequency (TF-IDF) ratings. The Document-word Matrix (DTM) is a representation where each row corresponds to a review and each column corresponds to a word from the lexicon. The entries inside the Document-Term Matrix (DTM) represent the Term Frequency-Inverse Document Frequency (TF-IDF) score given to each term in the respective review.

The task at hand involves the process of data annotation, specifically pertaining to the Amazon food reviews. The objective is to assign appropriate labels, either "positive" or" negative," based on the sentiment conveyed within the evaluations. Positive reviews are assigned a label of 1, and negative reviews are assigned a label of 0.
The procedure of training a model.

The Support Vector Machine (SVM) model is trained by using the Term Frequency-Inverse Document Frequency (TF-IDF) representation of the reviews, together with their corresponding labels that indicate the sentiments expressed. The Support Vector Machine (SVM) technique aims to create an optimal hyperplane that can effectively separate positive and negative ratings inside the feature space.
The assessment of the model:

The evaluation of the trained Support Vector Machine (SVM) model's performance in predicting the sentiment of Amazon food reviews was conducted by assessing its performance on a separate testing dataset. Calculate many evaluation metrics, such as accuracy, precision, recall, and F1 score, among other relevant measures.
Projected Results of Subsequent Assessments:

Once the Support Vector Machine (SVM) model has been trained and evaluated, it may be used for predicting the sentiment of previously unseen Amazon food reviews. The aforementioned task is achieved by converting the recently generated review text into the Term Frequency-Inverse Document Frequency (TF-IDF) representation. Subsequently, the acquired SVM model is used for the purpose of classification.
The combination of the Support Vector Machine (SVM) algorithm with the Term Frequency-Inverse Document Frequency (TF-IDF) technique is widely recognised as a reliable and efficient approach for conducting sentiment analysis and text classification tasks. The TF-IDF approach is used to assess the importance of keywords inside individual texts, while concurrently mitigating the impact of frequently appearing words throughout the whole corpus. The use of this technology allows for the prioritisation of word identification by Support Vector Machines (SVM) in order to choose terms that provide more informational value for sentiment prediction. As a consequence, the classification accuracy is improved.

Long Short Term Memory (LSTM)

The LSTM model proposed by Hochreiter and Schmidhuber served as the foundation for our exploration of more intricate models. Instead of using pre-trained word vectors, we conducted training of word vectors inside this model by leveraging a global word-word co-occurrence matrix including all terms present in all reviews. Subsequently, an LSTM model is used to include the representation vectors. The model presented in the next section primarily comprises a mapping between sequences of words and corresponding classifications.

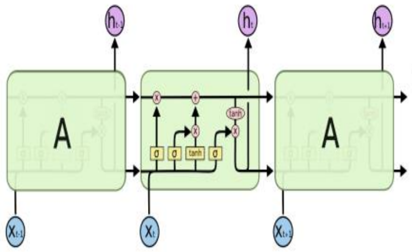


Figure 1: Work flow of LSTM

The mathematical equations for LSTM are:

$$\begin{aligned}
\text{input gate: } & \mathbf{g}^u = \sigma(\mathbf{W}^u * \mathbf{h}_{t-1} + \mathbf{I}^u * \mathbf{x}_t) \\
\text{forget gate: } & \mathbf{g}^f = \sigma(\mathbf{W}^f * \mathbf{h}_{t-1} + \mathbf{I}^f * \mathbf{x}_t) \\
\text{output gate: } & \mathbf{g}^o = \sigma(\mathbf{W}^o * \mathbf{h}_{t-1} + \mathbf{I}^o * \mathbf{x}_t) \\
\text{New memory cell: } & \mathbf{g}^c = \tanh(\mathbf{W}^c * \mathbf{h}_{t-1} + \mathbf{I}^c * \mathbf{x}_t) \\
\text{Final memory cell: } & \mathbf{m}_t = \mathbf{g}^f \odot + \mathbf{g}^u \odot \mathbf{g}^c \\
\text{Final hidden state: } & \mathbf{h}_t = \tanh(\mathbf{g}^o \odot \mathbf{m}_{t-1})
\end{aligned}$$

Upon passing the input through the LSTM model, the resulting output, denoted as yi, would manifest as a 2-dimensional vector that signifies the helpfulness of a given review. Furthermore, the consideration of class percentage was taken into account while computing the cross entropy loss [24, 25].
In order to do user rating categorization for Amazon food reviews using the LSTM algorithm, the following stages will be pursued:
Data Pre-processing: Prepare the dataset by cleaning and transforming the text data to numerical format suitable for LSTM input.
Splitting the Dataset: Divide the data into training and testing sets.
Tokenization: Convert the text into sequences of tokens (words or characters) and pad the sequences to have equal length.
LSTM Model: Create an LSTM model for user rating classification.
Training: Train the LSTM model on the training data.
Evaluation: Evaluate the model on the test data to measure its performance.
Prediction: Use the trained model to predict the user ratings for new Amazon food reviews.

The different layers used in the LSTM model:

Input Layer:

The input layer in the LSTM model handles the text data of the Amazon food reviews. Each review is represented as a sequence of words (or tokens). The length of the sequence is determined by the max_sequence_length parameter in the code provided earlier.

Embedding Layer:

The Embedding layer's job is to take the input sequences (words/tokens) and transform them into dense vectors of a certain size. This is significant because it enables the model to learn semantically accurate representations of words. The algorithm use a hyperparameter called embedding_size to adjust the size of the embeddings.

LSTM Layer:

When dealing with sequential data, the Long Short-Term Memory (LSTM) layer is where the model really shines. It is designed to pick up on textual dependencies over time, making it useful for sentiment analysis and other NLP applications. For long-term memory, the LSTM layer processes each word/token in the input sequence while keeping an internal state.

The lstm_units argument in the code specifies the total number of LSTM nodes. Each LSTM unit takes in one input sequence word or token at a time and uses that token to inform an update of its internal state.

Dropout Layer:
To avoid overfitting, the Dropout layer may be added if desired. In order to lessen the likelihood of the model memorising noise in the training data, it sets a percentage of input units to zero in a random fashion. We use a dropout rate of 0.5 in the algorithm, which indicates that half of the input units will be removed at random sometime during training.

Dense Layer (Output Layer):
The Dense layer is the output layer of the LSTM model. It receives the final hidden state from the LSTM layer and maps it to a single output value. In this case, we use a sigmoid activation function to produce a probability between 0 and 1. This probability represents the model's prediction of the user rating being positive (1) or negative (0).

During training, the model learns to adjust the parameters (weights and biases) of the embedding, LSTM, and dense layers to minimize the binary cross-entropy loss between the predicted ratings and the actual ratings. The optimization is performed using the Adam optimizer.

By using an LSTM, the model can take into account the sequential nature of text data and capture patterns in the reviews that influence the user ratings for Amazon food products.

V. Results

The results of this work are shown below:

Evaluation Metrics
The assessment of a model's performance has significant importance in the advancement of a dependable machine learning model. The evaluation of the recommended classifier's performance has been conducted by considering the four commonly used assessment metrics, namely Accuracy, Precision, Recall, and F1 Score.

The accuracy of the trained model and its prospective performance are determined by their correctness. Accuracy, which is defined as the proportion of correctly predicted samples to the total number of data samples, is a commonly used and easily comprehensible performance statistic. A substantial level of precision may lead to the inference that our trained model is the most optimal. An optimal condition for achieving high accuracy in performance is to have datasets that exhibit symmetry, with the numbers of false positives and false negatives being roughly equivalent.

Accuracy is a metric that quantifies the proportion of predicted positive data samples that are really positive, expressed as a percentage.

In the field of statistics, the concept of recall pertains to the ratio of accurately anticipated true positive data samples to the total number of true positive data samples.

The F1 Score may be defined as a calculated measure that combines Precision and Recall, two important performance criteria, via a weighted average.
Hence, this metric accounts for the potential occurrence of both type I and type II errors.
The F1 Score is sometimes considered more difficult to comprehend compared to accuracy, however it has more significance, especially in scenarios with a non-normal distribution of classes. When the costs associated with false positives and false negatives are almost equal, there is an improvement in accuracy.

The Receiver Operating Characteristic (ROC) curve is a visual aid used to assess the effectiveness of a binary classification system in relation to its accuracy. The visualisation shown is a scatter plot that demonstrates the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) across various cutoff values used in the classification process. The receiver operating characteristic (ROC) curve is a visual depiction that demonstrates the relationship between a classifier's recall (sensitivity) and precision (specificity), effectively displaying the trade-off between these two performance metrics.

The evaluation of the efficacy of a classification system may be performed by using a tabular depiction known as a confusion matrix. The following summary provides the predictions generated by the model in conjunction with the observed data categories.

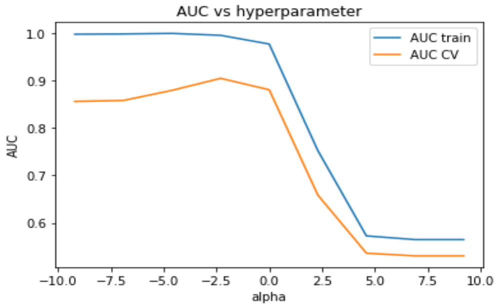
Within the realm of binary classification, the confusion matrix consists of four distinct components.

True Positives (TP) are defined as the number of samples that have been accurately classified as positive by the model.

False positives (FP) refer to the collective count of samples that were incorrectly identified as positive, although really being negative.

True Negatives (TN) refer to the accurate categorization of negative samples by the model, indicating the number of instances when the model correctly identifies a sample as negative.
The number of positive samples that were incorrectly identified as negative by the model, also known as false negatives (FN).

1. SVM
Applying SVM on BOW
The AUC (Area under curve) and ROC(Receiver Operating Characteristic curve) is shown below in Figure 2 and 3.




```
optimal alpha for which auc is maximum :  0.1
```

Figure 2: AUC curve

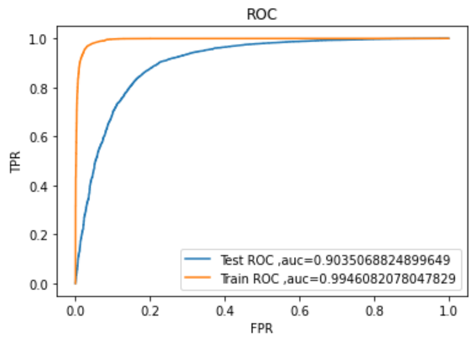


Figure 3: ROC curve

The confusion matrix for train data and test data is shown below in Figure 4 and 5.

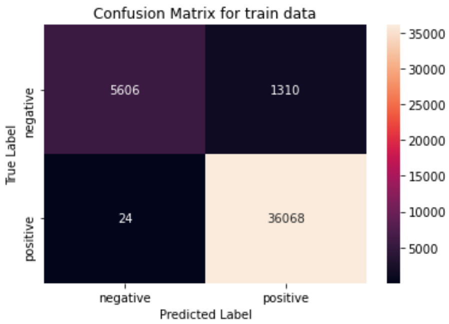


Figure 4: Confusion matrix for train data

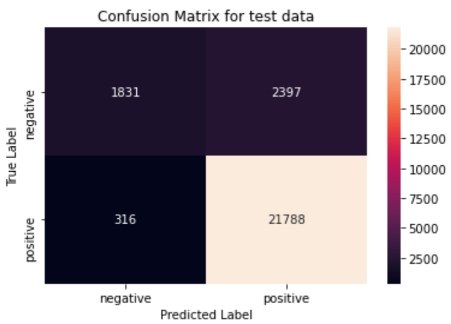


Figure 5: Confusion matrix for test data

Applying SVM on TFIDF

The ROC curve is shown below in Figure 6.

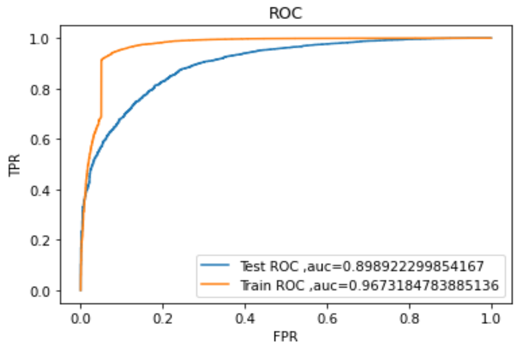


Figure 6: ROC curve

Figures 7 and 8 below depict the confusion matrices for the training data and the test data, respectively.

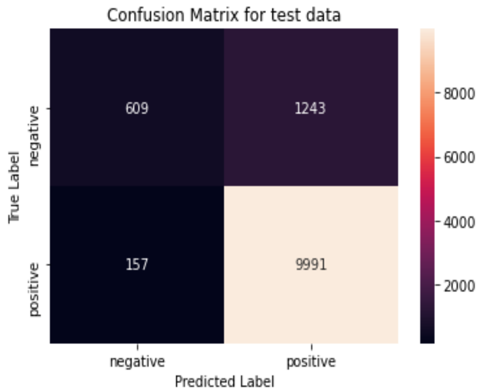


Figure 7: Confusion matrix for test data

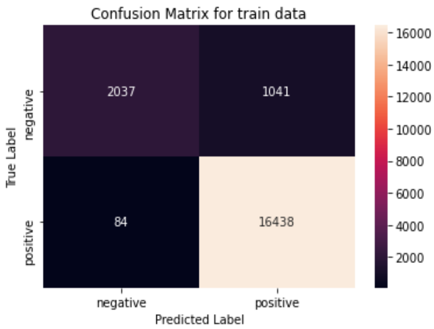


Figure 8: Confusion matrix for train data

```
The Accuracy of the Model is - 88.33333333333333
              precision    recall  f1-score   support

           0       0.80      0.33      0.47      1852
           1       0.89      0.98      0.93     10148

    accuracy                           0.88     12000
   macro avg       0.84      0.66      0.70     12000
weighted avg       0.87      0.88      0.86     12000

The Precision Score is - 88.9353747552074
```

2. LSTM

LSTM with one layer

```
Train on 254919 samples, validate on 109252 samples
Epoch 1/10
254919/254919 [==============================] - 605s 2ms/step - loss: 0.2515 - accuracy: 0.9007 - val_loss: 0.1883 - val_accur
acy: 0.9265
Epoch 2/10
254919/254919 [==============================] - 615s 2ms/step - loss: 0.1632 - accuracy: 0.9370 - val_loss: 0.1876 - val_accur
acy: 0.9289
Epoch 3/10
254919/254919 [==============================] - 604s 2ms/step - loss: 0.1449 - accuracy: 0.9450 - val_loss: 0.1843 - val_accur
acy: 0.9276
Epoch 4/10
254919/254919 [==============================] - 602s 2ms/step - loss: 0.1307 - accuracy: 0.9509 - val_loss: 0.1843 - val_accur
acy: 0.9289
Epoch 5/10
254919/254919 [==============================] - 597s 2ms/step - loss: 0.1180 - accuracy: 0.9559 - val_loss: 0.1943 - val_accur
acy: 0.9285
Epoch 6/10
254919/254919 [==============================] - 595s 2ms/step - loss: 0.1039 - accuracy: 0.9616 - val_loss: 0.2097 - val_accur
acy: 0.9277
Epoch 7/10
254919/254919 [==============================] - 455s 2ms/step - loss: 0.0918 - accuracy: 0.9664 - val_loss: 0.2066 - val_accur
acy: 0.9270
Epoch 8/10
254919/254919 [==============================] - 415s 2ms/step - loss: 0.0811 - accuracy: 0.9705 - val_loss: 0.2237 - val_accur
acy: 0.9262
Epoch 9/10
254919/254919 [==============================] - 413s 2ms/step - loss: 0.0727 - accuracy: 0.9737 - val_loss: 0.2343 - val_accur
acy: 0.9263
Epoch 10/10
254919/254919 [==============================] - 413s 2ms/step - loss: 0.0660 - accuracy: 0.9762 - val_loss: 0.2450 - val_accur
acy: 0.9261
```

```
Test score: 0.24501324092291224
Test accuracy: 0.9260974526405334
```

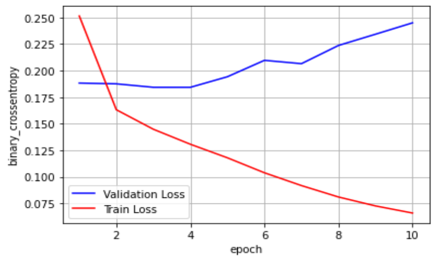


Figure 9: Validation performance graph

LSTM with two layers

```
Train on 254919 samples, validate on 109252 samples
Epoch 1/10
254919/254919 [==============================] - 1589s 6ms/step - loss: 0.2489 - accuracy: 0.9038 - val_loss: 0.1910 - val_accu
racy: 0.9263
Epoch 2/10
254919/254919 [==============================] - 1789s 7ms/step - loss: 0.1737 - accuracy: 0.9331 - val_loss: 0.1873 - val_accu
racy: 0.9289
Epoch 3/10
254919/254919 [==============================] - 1775s 7ms/step - loss: 0.1553 - accuracy: 0.9410 - val_loss: 0.1866 - val_accu
racy: 0.9304
Epoch 4/10
254919/254919 [==============================] - 1840s 7ms/step - loss: 0.1414 - accuracy: 0.9467 - val_loss: 0.1816 - val_accu
racy: 0.9315
Epoch 5/10
254919/254919 [==============================] - 1823s 7ms/step - loss: 0.1300 - accuracy: 0.9513 - val_loss: 0.1890 - val_accu
racy: 0.9318
Epoch 6/10
254919/254919 [==============================] - 1632s 6ms/step - loss: 0.1197 - accuracy: 0.9558 - val_loss: 0.1789 - val_accu
racy: 0.9333
Epoch 7/10
254919/254919 [==============================] - 1192s 5ms/step - loss: 0.1098 - accuracy: 0.9597 - val_loss: 0.1836 - val_accu
racy: 0.9345
Epoch 8/10
254919/254919 [==============================] - 1167s 5ms/step - loss: 0.1021 - accuracy: 0.9628 - val_loss: 0.1826 - val_accu
racy: 0.9347
Epoch 9/10
254919/254919 [==============================] - 1182s 5ms/step - loss: 0.0949 - accuracy: 0.9659 - val_loss: 0.1834 - val_accu
racy: 0.9341
Epoch 10/10
254919/254919 [==============================] - 1170s 5ms/step - loss: 0.0887 - accuracy: 0.9684 - val_loss: 0.1997 - val_accu
racy: 0.9341
```

```
Test score: 0.19967727044980735
Test accuracy: 0.9340515732765198
```

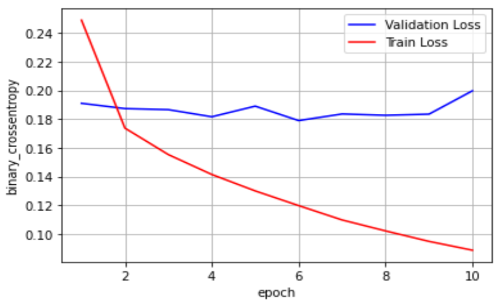


Figure 10: Validation performance graph

VI. Conclusion

In summary, this study has yielded valuable insights by investigating the effectiveness of Support Vector Machines (SVM) and Long Short-Term Memory (LSTM) neural networks in classifying user ratings in Amazon product evaluations pertaining to food.

SVM Performance: SVM demonstrated high accuracy in classifying user ratings for Amazon food reviews. Its ability to handle sparse and high-dimensional data, such as TF-IDF representations of text, made it a robust choice for sentiment classification. The model's straightforward implementation and interpretability also contributed to its appeal.

Accurate sentiment predictions were made using LSTM because this deep learning model successfully incorporated contextual relationships in the text input. Word embeddings were used to help the model acquire representations of words and their context, which improved the model's ability to comprehend the review's subject matter.

Both SVM and LSTM achieved satisfactory results in predicting user ratings. The accurate classification of user ratings has significant business implications.

Future Scope

While SVM and LSTM proved effective, the study opens avenues for further research. Deep learning architectures, such as Hybrid Deep belief Network could be explored to improve classification accuracy further. Additionally, incorporating aspects like context and user behavior in the analysis could provide richer insights.

VII.References

[1] B. Liu, Sentiment Analysis And Opinion Mining. San Rafael, Ca, Usa: Morgan & Claypool, May 2012.

[2] L.-C. Yu, J.-L. Wu, P.-C. Chang, And H.-S. Chu, "Using A Contextual Entropy Model To Expand Emotion Words And Their Intensity For The Sentiment Classification Of Stock Market News," Knowl.- Based Syst., Vol. 41, Pp. 89–97, Mar. 2013. [Online]. Available: Http://Www.Sciencedirect.Com/Science/Article/Pii/S095070511300004x

[3] M. Hagenau, M. Liebmann, And D. Neumann, "Automated News Reading: Stock Price Prediction Based On Financial News Using Context-Capturing Features," Decis. Support Syst., Vol. 55, No. 3, Pp. 685–697, Jun. 2013. [Online]. Available: Http://Www.Sciencedirect.Com/Science/Article/Pii/S0167923613000651

[4] T. Xu, Q. Peng, And Y. Cheng, "Identifying The Semantic Orientation Of Terms Using S-Hal For Sentiment Analysis," Knowl.-Based Syst., Vol. 35, Pp. 279–289, Nov. 2012. [Online]. Available: Https:// Www.Sciencedirect.Com/ Science/Article/ Abs/Pii/S0950705112001074

[5] Valdivia, M. V. Luzón, And F. Herrera, "Sentiment Analysis In Tripadvisor," Ieee Intell. Syst., Vol. 32, No. 4, Pp. 72–77, Aug. 2017.

[6] W. Medhat, A. Hassan, And H. Korashy, "Sentiment Analysis Algorithms And Applications: A Survey," Ain Shams Eng. J., Vol. 5, No. 4, Pp. 1093–1113, Dec. 2014. [Online]. Available:

[7] Http://Www.Sciencedirect.Com/ Science/Article/ Pii/ S2090447914000550

[8] Schoenmueller, V., Netzer, O. And Stahl, F. 2020. The Polarity Of Online Reviews: Prevalence, Drivers And Implications. Journal Of Marketing Research (Jmr), 57(5), Pp. 853–877.

[9] Karamitsos, I., Albarhami, S. And Apostolopoulos, C., 2019. Tweet Sentiment Analysis (Tsa) For Cloud Providers Using Classification Algorithms And Latent Semantic Analysis. Journal Of Data Analysis And Information Processing, 7(4), Pp.276-294.
[10] Zhao, Y., 2013. R And Data Mining: Examples And Case Studies. Academic Press.
[11] Jain, V.K., Kumar, S. And Mahanti, P., 2018. Sentiment Recognition In Customer Reviews Using Deep Learning. International Journal Of Enterprise Information Systems (Ijeis), 14(2), Pp.77-86.
[12] Lim, J., Park, M., Anitsal, S., Anitsal, M.M. And Anitsal, I. 2019. 'Retail Customer Sentiment Analysis: Customers' Reviews Of Top Ten Us Retailers' Performance,' Global Journal Of Management And Marketing, 3(1), 124+.
[13] Jagdale, R.S., Shirsat, V.S. And Deshmukh, S.N., 2019. Sentiment Analysis On Product Reviews Using Machine Learning Techniques. In Cognitive Informatics And Soft Computing (Pp. 639-647). Springer, Singapore.
[14] Sharma, S. K., Chakraborti, S. And Jha, T. 2019. Analysis Of Book Sales Prediction At Amazon Marketplace In India: A Machine Learning Approach. Information Systems And E-Business Management, 17(2–4), Pp. 261–284.
[15] Chong, A.Y.L., Li, B., Ngai, E.W., Chang, E. And Lee, F., 2016. Predicting Online Product Sales Via Online Reviews, Sentiments, And Promotion Strategies: A Big Data Architecture And Neural Network Approach. International Journal Of Operations & Production Management.
[16] Du, J, Rong, J, Michalska, S, Wang, H & Zhang, Y. 2019. Feature Selection For Helpfulness Prediction Of Online Product Reviews: An Empirical Study', Plos One, 14(12), P. E0226902.
[17] Meenakshi, A.B., Intwala, N., And Sawant, V., 2020. Sentiment Analysis Of Amazon Mobile Reviews. Ict Systems And Sustainability: Proceedings Of Ict4sd 2019, Volume 1, 1077, P.43.
[18] Govindaraj, S. And Gopalakrishnan, K. 2016. Intensified Sentiment Analysis Of Customer Product Reviews Using Acoustic And Textual Features. Etri Journal, 38(3), Pp. 494–501.
[19] Ghasemaghaei, M., Eslami, Sp, Deal, K. And Hassanein, K., 2018. Reviews' Length And Sentiment As Correlates Of Online Reviews' Ratings. Internet Research.
[20] Singla, Z., Randhawa, S. And Jain, S., 2017. Statistical And Sentiment Analysis Of Consumer Product Reviews. In 2017 8th International Conference On Computing, Communication And Networking Technologies (Icccnt) (Pp. 1-6). Ieee.
[21] R. Xia, C. Zong, And S. Li, “Ensemble Of Feature Sets And Classification Algorithms For Sentiment Classification,” Information Sciences, Vol. 181,No. 6, Pp. 1138–1152, 2011.
[22] G. Gautam And D. Yadav, “Sentiment Analysis Of Twitter Data Usingmachine Learning Approaches And Semantic Analysis,” In 2014 Seventh International Conference On Contemporary Computing (Ic3). Ieee, 2014, Pp. 437–442.
[23] X. Jiang, "A Facial Expression Recognition Model Based On Hmm," Proceedings Of 2011 International Conference On Electronic & Mechanical Engineering And Information Technology, Harbin, Heilongjiang, China, 2011.
[24] W. Swinkels, L. Claesen, F. Xiao And H. Shen, "Svm Point-Based Real-Time Emotion Detection," 2017 Ieee Conference On Dependable And Secure Computing, Taipei, 2017.
[25] Sundermeyer, Martin, Ralf Schlter, And Hermann Ney. ”Lstm Neural Networks For Language Modeling.” Interspeech. 2012.
[26] Zhou, Zhenxiang. “Amazon Food Review Classification Using Deep Learning And Recommender System.” (2016).
[27] G. Sathish砼 S. V. Saravanan’s. Narmadha砼S. U. Maheswari 2012.
[28] “Personal Authentication System Using Hand Vein Biometric.”
[29] International Journal Of Computertechnology And Applications砼 Vol. 3(1)砼
[30] Pp. 383-391砼 Jan-Feb 2012.