International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 4 Volume No.6 July – 2022
DOI:10.46647/ijetms.2022.v06i04.0071 ISSN: 2581-4621

Restaurant Review Analysis System Using Collective Blended Mechanism

Mr. Pavan Movva[1], Dr. Akbar Shaik[2], Prabhu Patnala[3], Priyanka S[4], Shehanaj Shaik[5], Jagadeesh Marri[6]

[1]Associate Professor, CSE, PSCMRCET, Vijayawada, Andhra Pradesh
[2]Professor, CSE, PSCMRCET, Vijayawada, Andhra Pradesh
[3,4,5,6]Student, CSE, PSCMRCET, Vijayawada, Andhra Pradesh

Abstract
Actually, most of businesses remain as failures due to lack of sufficient profits, lack of proper improvement measures. Specifically, the restaurant owners are facing a lot of problems and difficulties to improve their businesses due to lack of productivity in their business. Based on a report, the number of restaurants that have been shutting down in the first year are upto 60% and the restaurant businesses which are stopping their services within the first 5 years are about 80% in the world. So, this is becoming a universal problem now. There are various reasons involved in the failure of restaurant businesses, but serving low-quality food and lack of proper taste in the food items they are serving is a major one. To improve their business, they should know drawbacks of their food items and improve them by taking quality measures. To resolve the problem of business loss which is due to various drawbacks of food items, a solution is to implement a "RESTAURANT REVIEW ANALYSIS SYSTEM USING COLLECTIVE BLENDED MECHANISM". Each food item in the restaurant is assessed through the text review given by the customer where the text review is verified as a positive review or a negative review by the implementation of collection of classification models (Collective ML Model) in machine learning after handling textual data with natural language processing (to convert textual data to numerical data). And this data is stored into the database where each food item consists of number of customers, number of positive reviews, number of negative reviews, positive rate and negative rate. Now, the owner checks the least positively rated food item from the database, and takes necessary quality measurements which involve replacement of chefs or updating the ingredients used etc.

Keywords - Lemmatization, Stemming, Count vectorization, Universal Sentence encoder layer, Back propagation, Blending, Meta classifier, Collective Blended Model.

I. INTRODUCTION

The great opportunities and challenges for textual research has been opened up by the substantial increase in the volumes of unordered or unstructured textual data accompanied by the proliferation of tools to analyse them. Significantly, In the last few years, the scientific researches and study of the sentiment analysis has became very famous. Now a days, Business managers not only want to know about their product marketing strategies and the profit measures based on the number of transactions completed with customers but also want to know about the reviews/feedbacks and thoughts of the people based on their mindsets on using these products. The reviews they receive via social media and other internet services has became a key factor and a very important aspect to measure the quality of a product they are providing. Textual Sentiment analysis is a specific area or domain where the analysis is primarily aims on the extraction of feedback/review and thoughts of the users towards a specific aspect from a structured, or unstructured textual data. In this paper, we try to focus our

DOI:10.46647/ijetms.2022.v06i04.0071 ISSN: 2581-4621

effort on resolving the problems mentioned above and on implementing textual sentiment analysis on restaurant review databases. We used the textual sentiment expression to classify the reviews from the customers of the restaurant business whether it is negative or positive and perform identifying least positively rated food items and use these details for updating, maintenance and successful running of the restaurant business.

The proposed application involves the working process as follows, each food item in a restaurant is assessed through the text review given by the customer where the text review is verified as a positive review or a negative review by the implementation of collection of classification models (Collective ML Model) in machine learning after handling textual data with natural language processing (to convert textual data to numerical data). And this data is stored into the database where each food item consists of number of customers, number of positive reviews, number of negative reviews, positive rate and negative rate. Now, the owner checks the least positively rated food item from the database, and takes necessary quality measurements which involve changing of chefs or updating the ingredients used etc. The main sections present in our proposed method in analysing various reviews from the customers on food items involves data pre-processing, Building and implementing a collective model.

The data is collected from the reviews given by the customers on different online portal and sentiment depicting English words. The data pre-processing involves the implementation of removal of stop words, applying stemming and lemmatization of words using natural language processing techniques. The Collective ML Model is built using 5 classification mechanisms i.e. Logistic Regression (LR), Decision Trees(DT), Support Vector Machine (SVM), Artificial Neural Network(ANN), XGBoost(XGB).

II. LITERATURE REVIEW

[1]. Nina Lao et.al, This paper focused on building a useful app that compliments the most popular dishes with the help of the Yelp database. Whenever a user chooses something they want in a particular restaurant, that review is collected on their website. Now, these updates are sorted by Machine Learning Algorithms to classify reviews as positive or negative. A high-quality meal is the most recommended .Also, these highly recommended food items are returned to the user for further exploring. But there is no need for food recommendations to identify drawbacks of food items in the restaurant.

[2].AlhassanMabrouk et.al, compared the estimation results of greater than hundred Deep Learning based textual Sentiment analysis Classification approaches by using some of the public datasets of feedbacks given by consumers or customers within three particular areas or application domains (products, movies and restaurants). These datasets have various features (balanced/imbalanced, size, etc.) which lead to give a universal vision for our research study. Also, explained the various kinds of literature based comparisons to identify the most significant key factors in the three phases such as: feature representation, data preparation and classification techniques. The comparison demonstrates how the proposed data factors and other measures quantitatively affect the performance or estimation ability of the studied Deep Learning based textual Sentiment Classification approaches.

[3].RubaObiedat et.al, explained the textual sentence based sentiment analysis of customers feedbacks or reviews using a hybrid evolutionary SVM based approach in an unstructured an non-uniform Data Distribution of the information. The data is collected from the reviews for some

restaurants from the Jeeran website. After successfully completing the data preparation process, four different kinds which are in the form of versions of the dataset are presented using the different textual tokenization and embedding methods on the data. Earlier, individual random k value and a random weights for the sampling parameter are developed and initialised. By applying a technique called PSO optimization to determine the best k value and the best possible weights for the dataset attributes for every oversampling technique, after that, applied the SVM classification mechanism to the weighted and oversampled dataset to determine and extract the best possible sentiment results of the customer reviews of a specific restaurant.

[4].ShilpaShendre et.al, suggested a way to classify business reviews by using Sentiment analysis. Sentiment analysis focuses on the output of user's feedback and opinions on a particular topic from formal, or informal text data. In this case, major focus is on sentiment analysis at the restaurant review site. By testing the sentiment phrase, restaurant business reviews as distinguished pretty much as good or bad and make an output feature and use these features to review and maintain the business with Naïve Bayes and Linear Support Vector Classification (SVC).

[5].MarouaneBirjali et.al, described the categorization of the foremost used Sentiment analysis approaches briefly to possess a overview of accessible techniques and approaches like Machine Learning, Lexicon based and hybrid approach. Semantic orientation of words which are present in a text are used to extract sentiments from them in the lexicon based approach. Three key mechanisms for developing, annotating and creating sentiment lexicons are manual approach, corpus based approach, and dictionary based approach. Collection of common lexicons utilized by many researchers are WordNet, SenticNet, Sentiment WordNet, MPQA based subjectivity lexicon. Classification mechanism with Naive Bayes and SVM algorithms are taken as key factor and base features to compare newly proposed techniques

III. PROPOSED METHODOLOGY

Our Proposed Methodology involves the mechanism of the Collective Blended Model, which takes the use of the blending ensemble method. Every individual model provides results to some extent. But, the results provided by an individual model might not be optimal. To acquire optimal or better accurate results, there is a need to use a model which combines the properties of all individual models i.e., an ensemble model. So, we introduce a "Collective blended model" which involves the mechanism of the Blending ensemble method. The prediction of customer's textual review as whether +ve/-ve can be done by our collective blended model which involves the following phases.

- Data Pre-processing
- Prediction using individual models
- Developing blended model
- Performing final prediction

  - *Data Pre-processing*

An ML model can't make predictions on textual data as it operates only on numerical data. So, to transform a textual data to the numerical data, different Natural Language Processing techniques were used. Firstly, the removal of stop words (like 'the', 'and', 'an',….) from the customer's textual review gets happened followed by stemming the each word in the review. Finally, all the selected words from the review can undergo count vectorization technique to transform into the numerical data. The numerical data is formed on the basis of relative positions and corresponding frequencies of all words with respect to all other words in English language.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 4 Volume No.6 July – 2022
DOI:10.46647/ijetms.2022.v06i04.0071 ISSN: 2581-4621

Table1: Comparative Study On Existed Systems

S.No.	Author name	Algorithm	Advantages	Disadvantages
1	Nina Lao et.al	Logistic Regression	1. Very fast prediction of unknown records.	1.Difficult to filter out food items from the given reviews. 2. Difficult to handle non-linear data.
2.	Alhassan Mabrouk et.al	Neural Networks (Deep Learning)	1.It produces more accurate results compared to others.	1.It needs more time and memory compared to other techniques.
3.	RubaObiedat et.al	1.SVM(support vector machine)	1.Handles non-linear data efficiently.	1.Long training time 2.Requires feature scaling 3.Memory requirement is very high 4.This is not suitable for large data sets
4.	Shilpa Shendre et.al	1.Naïve Bayes	1.Easy to work and understand the mechanism. 2.The number of data points an estimators and predictor estimators have made it lot more scalable. 3. Very fast in making predictions as it uses a probability based model	1.Naive Bayes considers that all features are independent, which rarely occur in real world life scenario. 2.Its predictions can be Wrong or not accurate in some scenarios, so we should not accept its results in serious manner.
5.	Marouane Birjali et.al	SVM, Naive Bayes with lexicon based approach	1.The lexicons can handle different informal and slang words.	1. Performance is low during large datasets. 2. Difficult(Costly) to implement.

Stemming is widely used for text processing. Stemming is the process of finding root of words by removing either prefix or suffix. It links words with similar meanings to one word.
Text before stemming: amazing
Text after Stemming: amaze

DOI:10.46647/ijetms.2022.v06i04.0071 ISSN: 2581-4621

A machine learning model can't understand the textual input data on which it has to perform computations to make prediction results. So, we can't transfer textual data directly to any ML model. Thus, we have to convert the textual data into the numerical form. So, to perform this process we need to use some of the proper NLP (natural language processing) techniques. The key mechanism involved in doing this is to use a bag-of-words (BOW) model.

The bag-of-words model removes or discards the all information involved in the textual data and position of the textual characters present in each word of the sentence and only considers the frequent occurrences of the word. It converts specific kind of textual documents to the specific length list of numbers.

A global vector is implemented by considering all possible words that can occur in English Language, and each word is assigned with a unique number (commonly index of an array) besides the frequency count which demonstrating the number of occurrences of that specific word in the given sentence. This kind of encoding mechanism on the textual data, in which the key focus is to demonstrate the representation of the data but not on the position of the data.

Count Vectorizer is responsible for the tokenization of the textual sentences. This breaks the each textual sentence into collection of key words after performing the important data pre-processing techniques which involve removal of special characters, transforming the all words into lowercase, removing stopwords, applying stemming and lemmatization which find root words of the given word which are further referred to as keywords, etc.

A. *Prediction using individual models*

Each individual model which is already trained on training data from the dataset make predictions on review given by the customer after performing NLP techniques on review.

The following are the algorithms involved in our Collective Blended Model.

- Decision Tree
- Logistic Regression

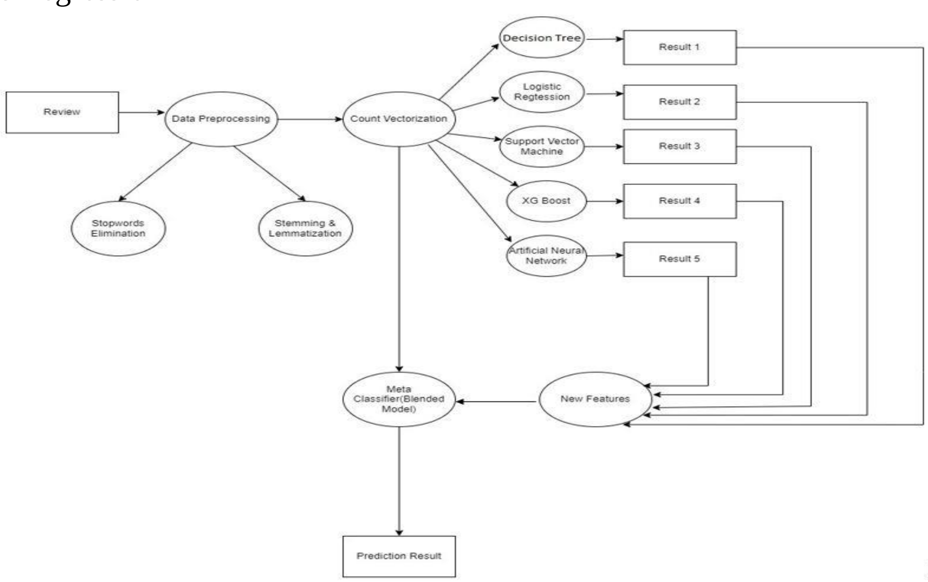

Fig1 : Mechanism for predicting customer reviews

- Support Vector Machine
- XG Boost
- Artificial Neural Networks

Let us dive into the working process involved in each individual model. Detailed Overview of each individual model is as follows,

Decision Trees:

A decision tree is a hierarchical representation of all feasible solutions to a condition or decision based on particular conditions which is done by the different features involved in the data. Review is the root node involved in the decision tree to classify the restaurant reviews. Apart from that, the prediction results made by each algorithm act as additional nodes (internal nodes) which play a key role in finding whether the review is positive or negative.

Logistic Regression:

The review in the numerical form is mapped to random weights initially. Sigmoid function involves in converting numerical results into suitable class label. The cost function updates the weights using gradient descent algorithm until the optimised results generated.

One of the best merits of Logistic Regression is that when we have a complex linear problem and not a whole lot of data it's still even able to produce and achieve much more useful predictions. On the other side, however, the traditional modelling predictions and decisions may lead to underfitting for complex and rich datasets.

Support Vector Machine (SVM):

Support Vector Machines (SVMs) are best useful for finding the best line in the 2D geometry and best plane in greater than 2D (like 3D,...) to differentiate the data space into different kind of classes. Linear SVM used in classifying the reviews as the data collected appears to be in linear form. These are best useful for implementing classification problems, however it can also solve regression problems. The hyperplane is identified or determined through the depiction of maximum margin, i.e., the maximum possible distance between data objects of the all kind of classes.

Hard margin refers to selecting the two hyperplanes with max possible distance among data objects of all kind of classes to separate the data objects according to the corresponding classes. Soft Margin Allowance of margin violation under the usage of non-linearly separable data.

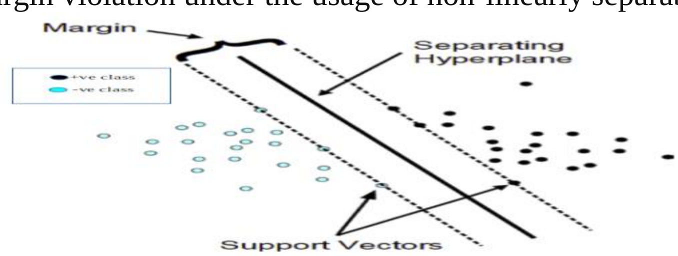

Fig2 :Support Vector Machine for linearly Separable data

Extreme Gradient Boosting (XG Boosting):

XGBoost is an efficient and reliable machine learning ensemble technique which is implemented based on the working of gradient boosting. It uses gradient descent algorithm to perform minimization of loss occurred during training and updates the new parameters. It is a decision tree ensemble based on gradient boosting. Variation in loss function results in identifying and controlling the complexity of performance when constructing the decision trees.

After performing proper data pre-processing on the restaurant review data, it gets subjected to the classification by the boosting technique. This loss function plays a crucial role in the integration

DOI:10.46647/ijetms.2022.v06i04.0071 ISSN: 2581-4621

of the split mechanism involved in the decision trees which lead to the strategical process of pre-pruning. The simpler trees are formed on the basis of higher values of γ result. The minimum loss reduction gain can be found to be controlled by using the values of γ which is highly needed to split a node. An important benefit of using these kind of algorithms is that the models require less storage space and trained relatively faster as compared to other kind of models.

Additionally, randomization techniques are highly implemented in Extreme Gradient Boosting to minimize the probability of occurrence of overfitting and to increase or improve the training process speed. Random subsamples to train the individual decision trees and Column subsampling at decision tree and tree node levels are some of the randomization techniques or mechanisms that were included in the Extreme Gradient Boosting (XGBoost) in classifying restaurant reviews.

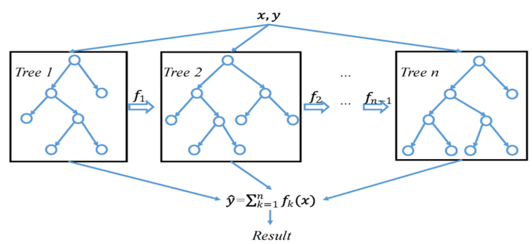

Fig3 : Mechanism of XGBoost

Artificial Neural Networks (ANN):

An Artificial Neural Network (ANNs) contains collection of layers in linear manner in which every layer consists of some no. of units which performs a specific task under the usage of an activation function. The input layer takes the numerical review data. The data transfers from the input layer units into the one or more hidden unit layers. The hidden units are responsible performing mathematical computations and transform the data from the input unit to the output unit. The artificial neural networks are depicted as fully connected networks from one layer to another layer as they contain some millions of neurons which transfer the data. Some weights are added to each connection. As each information passes through each layer unit in the network is learning more about the information or input data. At output unit's side, the network responds to the given data and process it.

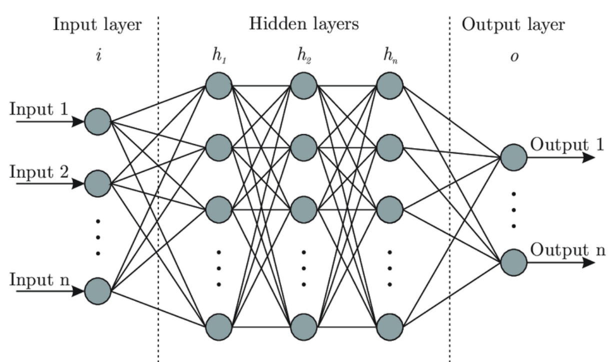

Fig4 : Artificial Neural Networks for Binary Classification

In the above ANN Architecture, the primary layer in the hidden layer set is used as Universal Sentence Encoder Layer. As we know hidden layer is less responsible for learning of the model for sentiment analysis on numerical data. So, we use Universal Sentence Encoder Layer to encode the review sentences from the training data and convert them into numerical data which is operable by

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 4 Volume No.6 July – 2022
DOI:10.46647/ijetms.2022.v06i04.0071 ISSN: 2581-4621

ANN model.

In universal sentence encoder layer, The model architecture consists of two encoders, which have two different goals. First one focusses on higher accuracy, but requires huge resource consumption. The other focusses on efficient inference, but with little lesser accuracy. The model transformer, a textual sentence model used for encoding frames the every textual sentence embedding process using the encoding sub-graph in the architecture of the transformer. The encoder receives the tokenized string in the lowercase order as the input data and produces a 512 dimensional vector as the sentence embedding in the form of output. The transformed based sentence encoding model achieves overall best task performance. But it requires more computational time and also consumes more memory when the sentence length is more.

The second encoding model which is known as the Deep Averaging Network (DAN) where input textual data embeddings for the words and bi-grams are formed as an average and then transferred through a feed-forward deep neural network to generate the finest textual sentence embeddings. DAN achieves strong baseline performance for classification of text reviews. For training encoder data, we have collected a huge responses from customers in different restaurants through online web sources and thus created a dataset.

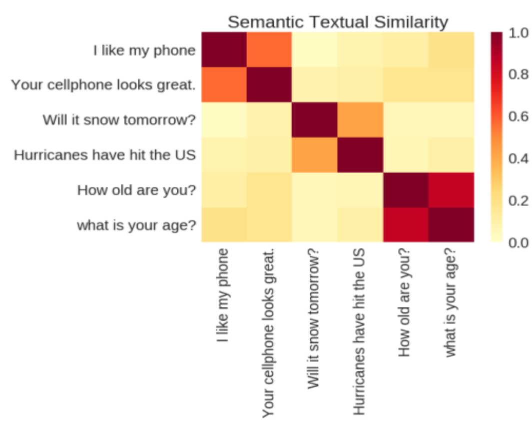

Fig5 : Similarity scores of the different kind of sentences using corresponding embeddings from the universal sentence encoder layer.

A. Developing Blended model:

Every trained individual model make predictions on cross validation and test set data and the predicted results by each individual model for each record of data are extracted as new features to the same record. Now the new attributes in the cross validation and test sets are (review, LR, DT, SVM, XGB, ANN). Now, a new kind of training data is created which is used to train the meta classifier.

Working of Collective Blended Model:

Collective Blended model uses the Blending ensemble method in the process of sentiment analysis of customer reviews and predicting the accurate status of the customer reviews. Following is the brief overview about blending technique.

“Blending” is a kind of ensemble technique which is also known as stacking generalization. In blending, Initially each individual model perform under training phase on the training data and performs predictions on the validation set and test set data. The results made by the each individual

DOI:10.46647/ijetms.2022.v06i04.0071 ISSN: 2581-4621

model are formed as new features and added to the corresponding subsequent input data present on the validation set and test set. Now, the meta classifier (collective blended model) model gets trained on the predictions results made by each individual model (base model) that were made on a separate cross validation set of the dataset instead of full and folded training set. And finally, our collective blended model which is also known as the meta classifier model perform predictions on newly formed test set with added features as base model predictions. It makes the predictions which are more accurate than the results achieved by each and every single individual model used.

Steps involved in Working Mechanism of Blending Ensemble Model:

Step-1: Training dataset is split into base train data, validation data and test data.

Step-2: Train the individual models i.e., Decision Tree, Logistic Regression, Support Vector Machine (SVM), XG Boost, Artificial Neural Networks (ANN) on the training data of our dataset and perform predictions on validation set and test set data. The creation of new predictions made by this mechanism.

Step-3: A Collective Blended Model, a new meta-classifier model is then fitted on validation/holdout set using individual model prediction attributes which are the results made by each individual model. For this, both actual features/attributes and new meta features from the holdout or validation set will be used.

Step-4: Finally, this trained collective blended model is used to perform the crucial final predictions on the test set data which involves the usage of original and new blended model (meta) features.

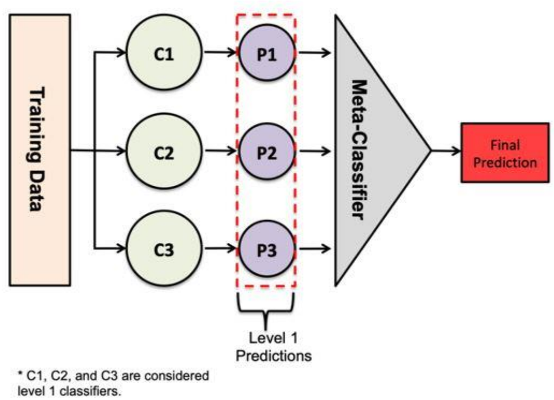

Fig6 : Blending Ensemble Technique

Table 2 : Data With Predictions As New Features

Review	X_{LR}	X_{DT}	X_{SVM}	X_{XGB}	X_{ANN}	Y
Numerical_review	Pred_result (LR)	Pred_result (DT)	Pred_result (SVM)	Pred_result (SVM)	Pred_result (ANN)	Actual_ result

D. Performing final prediction:

Now the trained blended model take the use of prediction results of each individual models and make final prediction based on given review and results predicted by each individual models. In this way the final prediction result is fetched through the blended model.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 4 Volume No.6 July – 2022
DOI:10.46647/ijetms.2022.v06i04.0071 ISSN: 2581-4621

IV. RECOMMENDING TOP RELEVANT REVIEWS

The process of recommending top relevant positive and negative reviews based on the reviews given by the customers to our system. The important and key words present in each review are extracted and corresponding frequencies are maintained in a database. Now, the frequency cost is calculated for each review as the sum of the frequencies of all keywords present in the customer review. And finally the positive reviews with highest frequency costs are recommended as top positive relevant reviews and the negative reviews with highest frequency costs are recommended as top negative relevant reviews to the owner of the restaurant through our system which gets displayed on the dashboard.

$$\textbf{freqCost [review]} = \sum \textbf{freq [keyword}_i\textbf{]}$$

where, $keyword_i$ is the ith key word present in the given customer review.

V. RESULTS AND PERFORMANCE EVOLUTION

Our Collective Blended Model achieves better accurate results (80.2893%) as compared to the other individual machine learning classification models such as Decision Tree (71.7902%), Logistic Regression (71.9711%), Support Vector Machine (72.3327%), XGBoost (71.9711%), Artificial Neural Networks (79.9277%). Our Collective Blended Model uses the prediction results achieved by the each individual classification model an builds a meta classifier to predict the accurate results collectively.

```
#########################    Decision Tree     #########################

Accuracy       Precision       Recall  F1-Score
71.7902        0.7268       0.7179        0.6744

#########################    Logistic Regression     #########################

Accuracy       Precision       Recall  F1-Score
71.9711        0.7607       0.7197        0.662

#########################    Support Vector Machine     #########################

Accuracy       Precision       Recall  F1-Score
72.3327        0.7416       0.7233        0.6775

#########################    XGBoost     #########################

Accuracy       Precision       Recall  F1-Score
71.9711        0.7758       0.7197        0.6573

#########################    Artificial Neural Networks     #########################

Accuracy       Precision       Recall  F1-Score
79.9277        0.8002       0.7993        0.7888

#########################    Collective Blended Model     #########################

Accuracy       Precision       Recall  F1-Score
80.2893        0.8025       0.8029        0.794
```

Fig 7: Results for Blended Model Compared to each individual model

Accuracy: The Accuracy is demonstrated as the absolute accuracy of pattern and got estimated as the total of specific prediction factors. In the proposed model, we have applied the model on two datasets.

The Customer Reviews dataset contains 5530 records. The confusion matrix for the Customer Reviews dataset for intelligent ensemble (collective blended) algorithm can be described as follows and the computation of accuracy is shown in equation.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 4 Volume No.6 July – 2022
DOI:10.46647/ijetms.2022.v06i04.0071 ISSN: 2581-4621

Table 3: Confusion Matrix

	Actual=true	**Actual=false**
Predicted=true	332	28
Predicted=false	81	112

Accuracy = (True Positives + True Negatives) / Total Number of records
= (TP+TN)/(TP+FP+TN+FN)
= 80.28%

Recall: Recall is referred to as "How many are predicted positive out of all actual positive results". The formula and computation for recall is as follows:

Recall = TP/TP+FN
= 80.29%

Precision: Precision is referred to as "How many are actually positive out of all positive predictions". The formula and computation for precision is as follows:

Precision= TP/TP+FP
= 80.25%

F-Measure: It referred to as the harmonic mean of the recall and precision.

F-measure = 2*((Recall*Precision) / (Recall +Precision))

= 2*((0.8029*0.8025)/(0.8029+0.8025))

= 79.4%.

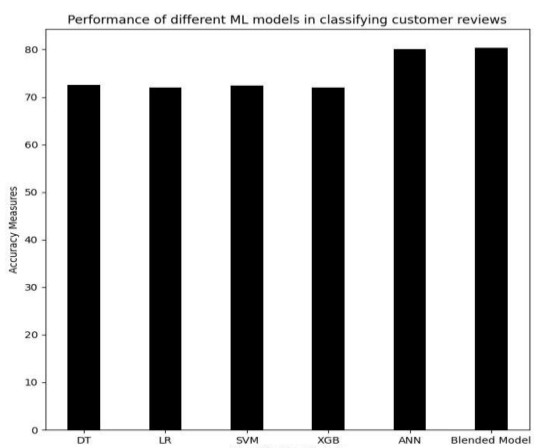

Fig8 : Bar plot depicting the performances of each model

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 4 Volume No.6 July – 2022
DOI:10.46647/ijetms.2022.v06i04.0071 ISSN: 2581-4621

VI. CONCLUSION

This application perfectly helps to the business managers who wants or needs to improve their productivity, which in turn indirectly improves their business standards and significantly lead to good profits. By taking reviews from customers, the business problems got diminished through the development of customer-friendly service standards and productivity. This system can be kept and used in any food hotel or restaurant. It perfectly plays an effective role in taking food improvement measurements which directly lead to improvement one's business. It is Consumer-friendly application. The feasibility of the project lies in business performance of restaurants because this system provide better results and make a greater impact on one's business if and only if number of visiting customers are more. Also, as long as owner want to know the drawbacks of his own restaurant food items, the system is used. Finally, the overall impact made by the application to a restaurant is relatively greater as compared to manually maintaining the data.

REFERENCES

[1] Nina Lao et.al, Analyzing and filtering food items in Restaurant reviews, https://aircconline.com/csit/papers/vol10/csit101208.pdf
[2]Alhassan Mabrouk et.al, Deep Learning Based Sentiment Classification,
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9085334
[3] Ruba Obiedat et.al, Sentiment Analysis of Customer Reviews Using a Hybrid Evolutionary SVM-Based Approah in an Imbalanced data distribution, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&
arnumber=9706209
[4] Shilpa Shendre et.al, Classification of Business reviews using Sentiment Analysis, https://www.irjet.net/archives/V6/i6/IRJET-V6I6571.pdf
[5] Maroune Birjali et.al, A Comprehensive Survey on Sentiment Analysis, https://www.sciencedirect.com/science/article/pii/S095070512100397X
[6] Cui, H., V. Mittal and M. Datar. Comparative experiments on sentiment classification for online product reviews. In Proceedings of AAAI-2006,2006.28.
[7] Das, S. and M. Chen. Yahoo! For Amazon: Extracting market sentiment from stock message boards. In Procedings of APFA-2001, 2001.
[8]Alm, C.O. Subjective natural language problems: motivations, applications, characterizations, and implications. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: short papers (ACL-2011), 2011.
[9] Andreevskaia, A. and S. Bergler. Mining Wordnet for fuzzy sentiment: Sentiment tag extraction from WordNet glosses. In Proceedings of Conference of the European Chapter of the Association for Computational Linguistics (EACL-06), 2006.
[10] Daniel Ellis, Using Sentiment Analysis to explore emotions within text methods, towardsdatascience.com.