Food Ordering System based on Human Computer Interaction and Machine Learning Techniques

Anmol Sajnani[1], Nirmal Patel[2]
[1]anmolsajnani04@gmail.com, [2]nirmalpatel783@gmail.com

***Abstract*—There is a saying,"the best interface is no interface". Many people consider voice interface as an excellent approach to communicate with computer system. In the following paper, we are trying to build up a system which is based on Human Computer Interaction to put in an order for food in restaurants. People tend to go to a restaurant where service and convenience of the customer are given top priority. A major problem faced by the customers in a restaurant is waiting for the staff to take the order, this is because during rush hours the availability of staff decreases. This mishap can be removed by implementing a Food ordering system, a device which is built using Raspberry Pi 3; along with voice controlled interface, as presented in this paper. Techniques like Natural Language Processing, Rule-based system, Sentiment Analysis are used in the proposed system.**

***Keywords*—Automatic Speech Recognition(ASR), Deep Learn- ing(DL), Machine Learning(ML), Natural Language Process- ing(NLP), Sentiment Analysis(SA), Text to Speech(TTS)**

I. INTRODUCTION

Nowadays, we see many restaurants competing on the basis of taste rather than their service. That being said, we see many restaurants having excellency in food taste, but average to poor performance in customer service. Many restaurants find it difficult to cope with all the orders during rush hours. So we came up with a solution to this problem with the help of a Food Ordering System deployed using NLP, ML, DL and Rule-Based System.

The focus of this research is to build a device which can help restaurants in optimizing their profits by providing the following functionalities:

- A computing device for interacting with the customers, taking orders from them, asking their preferences. This reduces human labour and increases the enjoyment of Bon Vivante in the restaurant.
- Applying Sentiment Analysis to the customers' reviews and feedback which can provide restaurants knowledge about areas of improvement and different price ranges to consider.

NLP based systems takes order as voice commands and converts them to text, this data is stored in the DB and sent to management server where the data is processed. The order is sent to the kitchen and when it is ready, the customer is served with it directly. Once the customer is satisfied with the food and asks for the bill, along with the option of providing feedback, the bill is presented.

If the customer opts for providing the feedback, Sentiment Analysis algorithm is orchestrated in the background to scru- tinize the feedback provided by the user. Now when sufficient feedbacks are gathered from various customers, restaurant authorities can come to know what kind of food people like, what is currently in demand. Raspberry Pi version 3 is used as device hardware to support our system.

The implementation of such a system will lead to major advantages:

- During rush hours multiple customer requests will be handled efficiently without the customer having to wait to place the order.
- This system eliminates the overhead of restaurant staff taking orders from the customer and the miscommunica- tions happening in the process.

The structure of the paper is as follows, Section II give a brief of the Literature Survey done and Section III deals with the Proposed System for Food Ordering, which is followed by Results and Discussion in Section IV, Conclusion and Future Work are provided in Section V and Section VI respectively.

II. LITERATURE SURVEY

A. Speech Recognition

In Speech Recognition technology speech is given as input and is converted to text, Automatic Speech Recognition is one of the technology used for the same. Automatic Speech Recog- nition (ASR) can be defined as the independent, computer- driven transcription of spoken language into readable text in real time. ASR technology allows a computer to identify the words that a person speaks into a microphone or telephone and converts it into written text. Although ASR technology is not yet at the point where machines understand all speech, in any acoustic environment, or by any person. It is used on a day-to-day basis in a number of Natural Language Processing applications and services.

The goal of ASR research is to allow a computer to recognize in real-time, with 100% accuracy, all words that are intelligibly spoken by any person, independent of vocabulary size, noise, speaker characteristics or accent. Today, if the system is trained to learn an individual speakers voice, then larger vocabularies are possible and accuracy can be increased above 90%. Most commercial companies claim that recogni- tion software can achieve between 98% to 99% accuracy if operated under optimal conditions. Optimal conditions usually assume that users have speech characteristics which match the training data, can achieve proper speaker adaptation, and work in a clean noise environment [1].

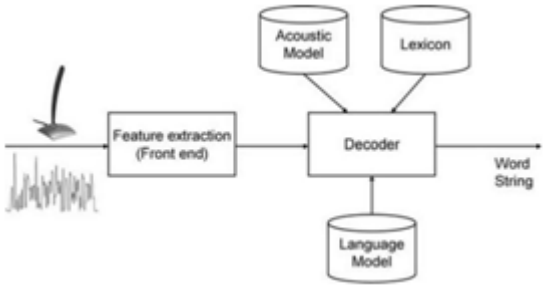

Fig. 1. Block diagram of a typical ASR system

Figure 1 shows the block diagram of a typical ASR system. It composes of two major components: the front end and the decoder. The front end block extracts the spectrum representation of the speech waveform. The most widely used features are Mel Frequency Cepstral Coefficients (MFCC) [2]. The decoder block searches the best match of word sequences for the input acoustic features based on acoustic model, lexicon, and language model.

B. Database (MongoDB)

MongoDB is a document-store database designed for best scalability, high availability and good performance. It allows data persistence in a nested state and has the ability to query the nested data in an undefined fashion with embedded queries. In addition, it does not inflict schema, allowing it to adapt quickly as applications. Moreover, a MongoDB document can contain field types that other documents of the same collec- tion do not have. Anyhow of this flexibility, MongoDB still ensures expected functionalities such as full query language and consistency.

MongoDB is in the forefront of NoSQL databases, provid- ing agility and scalability to businesses. More than thousand companies and new start-up companies have acquired and are using MongoDB to develop new applications as it refines client experience, speeds up marketing time and decreases costs. Most large-scale companies like Facebook, Amazon, Google etc. have already implemented MongoDB.

MongoDB stores data as a document in a binary-encoded form called BSON or simply Binary JSON. Like in relational databases, MongoDB organizes documents that tend to have a similar structure as collections. A Collection in MongoDB corresponds to a table in relational databases, a document is a row, and a field is a column.

Dynamic Schema:

The structure of a MongoDB document can vary from document to document in the form of JSON, unlike in a relational database where the structure for a row must be defined. For instance, all documents describing Twitter Users might contain user ID, tweets and followers. However, some documents do not necessarily have to require user ID for one or more third- party applications. Hence, fields can be added to a document if needed, without disrupting other documents or updating the central system catalog or having system downtime.

Because of these major advantages of MongoDB over other Database systems, it is chosen to support our Food Ordering System

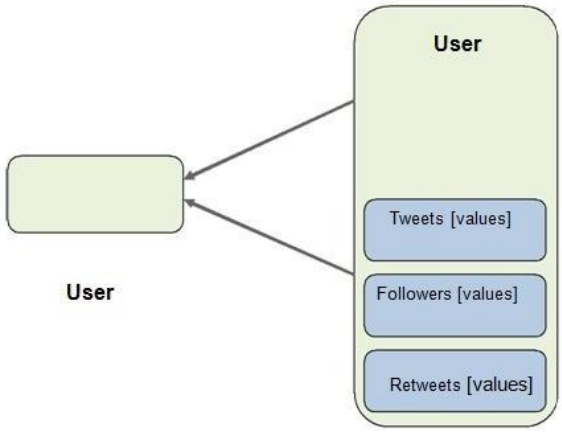

Fig. 2. Demonstration of MongoDB [4]

SQL QUERIES	MongoDB Queries
SELECT QUERY: Select * From Students where id = "200A"	**SELECT QUERY:** db.Students.find({},{t_id:1})
INSERTION QUERY: Insert into students (s_id,name,course,branch,email) values("200", "Lokesh", "M.Tech", "CSE", "lkbansal1993@gmail.com")	**INSERTION QUERY:** db.students.insert({s_id : "200", name: "Lokesh", course: "M.Tech", branch: "CSE", email: "lkbansal1993@gmail.com"})
DROP QUERY: Drop table students;	**DROP QUERY:** db.students.drop();
DELETE QUERY: Delete From students where s_id = "200"	**DELETE QUERY:** db.students.delete({s_id: "200"})

Fig. 3. Difference between SQL queries and NoSQL(MongoDB) queries [4]

C. Seq2Seq

Sequence To Sequence model introduced in Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation has become the Go-To model for Ma- chine Translations [5]. There are two RNNs (Recurrent Neural Network) : An Encoder and a Decoder. The encoder an input as a sequence of sentences and processes one word at each timestep. Its objective is to convert a sequence of symbols into a fixed size feature vector that encodes only the important information in the sequence while losing the unnecessary information. You can visualize data flow in the encoder along the time axis, as the flow of local information from one end of the sequence to another.

Each hidden state influences the next hidden state and the final hidden state can be seen as the summary of the sequence. This state is called the context or thought vector, as it represents the intention of the sequence. From the context, the decoder generates another sequence, one symbol(word) at a time. Here, at each time step, the decoder is influenced by the context and the previously generated symbols.

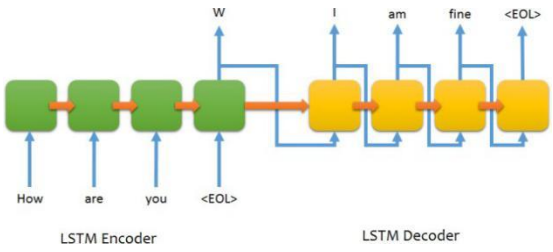

Fig. 4. Seq2Seq input and output [6]

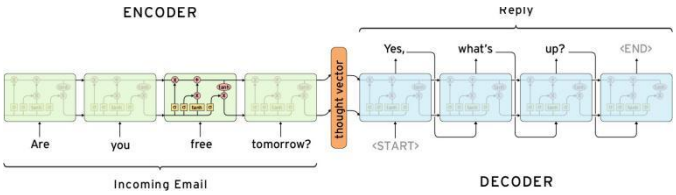

Fig. 5. Seq2Seq as chatbot [7]

1) Data Padding:

Before training, we work on the dataset to convert the variable length sequences into fixed length sequences, by padding. We use a few special symbols to fill in the sequence.

- EOS : End of sentence
- PAD : Filler
- GO : Start decoding
- UNK : Unknown; word not in vocabulary

D. Feedback Analysis

Restaurants always want their customers to feel good about their food. This can be achieved by analyzing the feedbacks given by the customer. From this knowledge, restaurants can improve themselves and serve the customers well.

Articulating the emotions and feelings with the succor of words makes human beings inimitable. Technobabbly these feelings are known as sentiments and the process of analyzing these statements is called Sentiment Analysis/Opinion Mining [8].

1) Survey of Product Reviews using Sentiment Analysis [9]:

In this paper, the author has proposed a technique called sentiment orientation, which automatically finds the frequently used terms for an aspect of a product from online customer reviews. The methodology which is put forward provides an efficient way of predicting the users' opinion and thereby suggesting them. Classifying the product review/opinion on the basis of positive and negative is the major task in one of the supervised machine learning approach called lexicon-based approach. This work is said to be sentiment orientation.

The problems which arose are confection (i.e) multiple ways to arrive at the solution and the second one is to deal with the context based words. To overcome these problems, the sentiment orientation algorithm is proposed by the author. It includes two major approaches:

- Corpus-based approach: It is classified as a linguistic- based approach and mainly identifies the emotional sim- ilarity of the word. It follows 3A perspective which is: Annotation, Abstraction and Analysis.
- Dictionary-based approach: This approach makes use of datasets like wordnet. By using the resources of words provided by wordnet it is possible to identify the large set of text from the comments that are retrieved. Some of the major advantages of this approach include ease of use, increases time efficiency, removes unwanted, duplicate content and the dataset.

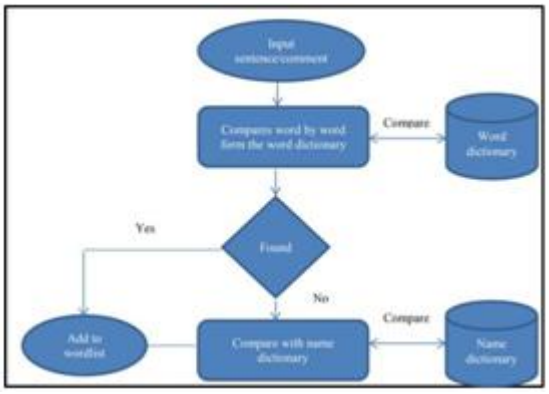

Fig. 6. Flowchart for Semantic Orientation [9]

Figure 6 shows the flowchart explanation for sentiment orientation. The Wordlist is the output of this system which is created by comparing words in word dictionary and name dictionary respectively. Figure 7 shows the difference of percentage accuracy between existing product aspect ranking model and the proposed model.

Parameters	Existing product aspect ranking model	Semantic orientation model
Time Efficiency	70%	90%
Accuracy	70%	100%

Fig. 7. Difference of percentage accuracy between existing product aspect ranking model and the proposed model [9]

2) Application of Machine Learning Techniques to Sentiment Analysis [10]

Sentiment Analysis/Opinion Mining is the process of identifying whether the user-generated text expresses positive, negative or neutral opinion about an en- tity which can be product, people, topic, event etc. In this paper, the author has proposed several Sentiment Analysis approaches which include following:

Lexicon Based approach: This approach deals with counting of positive and negative words in the output text. If the text consists of more positive words, then the text is assigned a positive score and if there are more number of negative words, then it is assigned as negative score. There must be some cases where there are equal number of positive and negative words then it is assigned as neutral score. This process is called building an opinion lexicon. Two main approaches for this purpose are:

  - Dictionary-based approach
  - Corpus-based approach
- Machine-Learning Based approach: In this approach, the dataset is divided into training and testing dataset. Various Machine Learning Algorithms which are usually used for classification of text are: Maximum Entropy, Nave Bayes, Support vector machines
- Hybrid approach: This approach combines both lexicon and machine learning based sentiment analysis tech- niques. The advantage of this is best of both worlds can be attained.

In [10], the author has briefly described each step which needs to be followed for Sentiment analysis procedure, below described steps are applied on twitter dataset, hence describing them in detail-

- Data Collection: Twitter data is collected using Twitter API
- Data Preprocessing: Steps that are carried out in prepro- cessing are as follows-
  - Case Conversion
  - Stop-words Removal
  - Punctuation Removal
  - Stemming
  - Lemmatization
  - Spelling Correction
- Feature Extraction: Once tweets are preprocessed, fea- tures are extracted which are relevant for sentiment analysis. Some example features are- team presence and frequency, parts of speech tagging, opinion words and phrases, negation, presence of emoticons in tweets, hash- tags which are all twitter specific features.
- Training and Testing Machine Learning classifier: After features are selected, a Machine Learning classifier is chosen for sentiment Analysis. List of Machine Learning Classifiers:
  - Naive Bayes Classifier
  - Support vector machines
  - Decision trees

In the paper [10], authors have used Apache Spark to obtain quicker results. In the proposed work, analysis is performed on datasets of different sizes and domains to demonstrate that the proposed framework works on data of all sizes and domains. Three different datasets of varying sizes and domains are considered for analysis.

The following points can be elaborated based on outcome perusal:-

- Multinomial Naive Bayes does not perform as expected when supplied with a small training dataset.
- Decision tree takes longer training time than Naive Bayes.
- Decision tree takes very less time for predicting unseen data compared to multinomial Naive Bayes.
- Decision tree performs better than Multinomial Naive Bayes for datasets of varied sizes and domains.
- This work uses Apache Spark Cluster for processing. Hence, the proposed framework is able to produce the results of analysis quickly.

Performance Analysis of Supervised Machine Learning Techniques for Sentiment Analysis [11]: Human beings always tries to make the machine smarter to solve all its problem smartly and efficiently within some stipulated period of time. Therefore, machine learning techniques come to action by using which machines get trained and expected to work accordingly. Machine Learning (ML) is a field of Computer Science in which machines of computers are able to learn without being programmed explicitly.

In [11], the author has proposed a methodology where Sentiment Analysis is done on movie reviews. Step-by-Step procedure for the same is described below:

- Collecting Movie Reviews Data Sets
- Cleaning the Data Sets: Characters, numbers, special characters, and unrecognized characters are all removed from the dataset, this process is called cleaning or pre- processing of data.
- Data Categorization: Supervised Machine Learning tech- nique take labelled data, which needs categorizing the data into positive or negative. In this, author has used python for labelling the data.
- Preparing training and testing datasets: 70% of data is considered as training dataset and 30% as testing data
- Training the Model with training Data Sets
- Testing the Model with Training DataSets

Classifiers	10600	25000	35600	50000	85600
Naive Bayes	87.34295643	93.89467593	94.1190275	94.82060185	95.58089034
Multinomial naive Bayes	86.9152633	92.25280417	93.06768047	94.39380787	95.21534564
Bernoulli naive Bayes	87.12910986	92.41898148	93.29296912	94.23466435	95.15019906
Logistic Regression	88.62603582	96.81712963	97.8550643	98.94748264	99.46073109
SGDClassifier	84.92381716	94.48784722	95.06242373	97.62008102	98.55591748
LinearSVC	89.00026731	96.81834491	96.61513189	98.91319444	100
NuSVC	86.56776263	92.41898148	92.61710316	93.81510417	94.37929786

Fig. 8. Experimental results of 6 classifiers [11]

Result: Figure 8 shows the experimental results of 6 classifiers. From which LinearSVC performs the best.

E. Speech Synthesis

Speech synthesis is the artificial production of human speech. A computer system used for this purpose is called a speech computer or speech synthesizer and can be imple- mented in software or hardware products. A text-to-speech (TTS) system converts normal language text into speech. A text to speech output is based on generating a corresponding sound output when the text is inputted [12].Wide range of applications use text to speech technique in medicals, telecom- munications fields, etc. Each spoken word is created from the phonetic combination of a set of vowel and consonant speech sound units. Producing an artificial human speech is known as speech synthesis. The Various speech synthesis methods that have been used for text to speech output for obtaining intelligible and natural output are Concatenative, Formant, Articulatory, Hidden Markov model (HMM) [14].

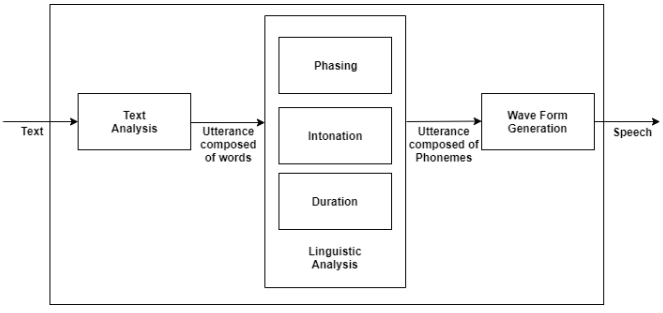

Fig. 9. Speech Synthesis Module [13]

III. PROPOSED SYSTEM

All the above literature survey was used to build a Food Ordering System which works as follows: The system first analyzes the speech given by the user. The Google text to speech algorithm in python (https://pypi.org/project/gTTS/) will convert the spoken sentence into a text format. Once the speech format is converted into text, the text will be sent to the rule-based system for processing, this request will be sent using a post request (rest api) which will take an input of the text as the headers and return multiple attributes such as the response, intent as well as entities. Figure 10 shows the overview of the proposed system.

Our AI response module has the following working modules.

- rule-based system
- seq2seq outlier request handling model

Rule-based system

Our rule-based system will first get the input text from the ”speech to text module” and this text will be run on an entity detection module in order to generate the given meaning of the sentence. The intent will signify the meaning of the sentence in order to maintain the conversation flow. Once the intent is determined the rule-based system will generate an output response in relation to the intent. While the text is being processed, the entities will be recorded such as ”1 burger” in order to calculate the bill.

IV. SEQ2SEQ OUTLIER REQUEST HANDLING MODEL [15]

Seq2seq is a neural network model [15] which will help in handling the outlier queries. The model will receive the queries once the rule-based system detects that the intent for the given input text is an outlier intent. The seq2seq model will generate a suitable response instead of a default response.

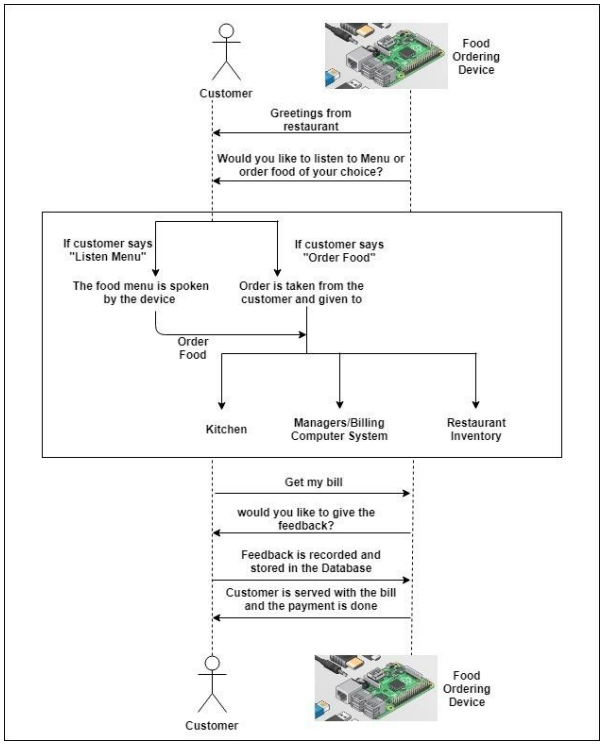

Fig. 10. Overview of the Proposed System

1) Training data: The model was trained on the twitter data corpus in order to have a generalized dataset. This dataset captures normal conversations that might not be specific to the food ordering system.

2) Model generation: The code is written in python along with the Tensorflow library.

The data is provided in an encoding and a decoding format. Where the data will be tagged according to whether it is for an encoding format or a decoding format.

```
input_seqs : ['how', 'are', 'you', '<PAD_ID'>]
decode_seqs : ['<START_ID>', 'I', 'am', 'fine', '<PAD_ID'>]
target_seqs : ['I', 'am', 'fine', '<END_ID>', '<PAD_ID'>]
target_mask : [1, 1, 1, 1, 0]
```

Fig. 11. Data Format

Our system will then process the text and detect intents out of the given input, generating an apt response for the user query. Meanwhile based on the given intent the order and the bill will be recorded for futher storage into the database. Once the user is done with his food he will be given an option to give a feedback about his experience, during this time the user will also be prompted about his total bill. The bill will be calculated by getting the default prices from the MongoDB server and multiplying them by their quantities. If the user feels like giving a feedback, the feedback will be processed.

Feedback Analysis

Feedback Analysis/ Sentiment Analysis is done using *Nltk* with python 3.5.2, using spyder ide, preprocessing steps per- formed are [8]:

1) Tokenizing
2) Stop Words
3) Stemming
4) Part of Speech Tagging
5) Chunking
6) Lemmatizing
7) Named Entity Recognition

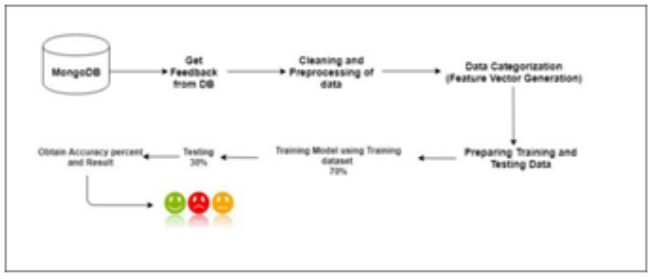

Fig. 12. Steps for Feedback Analysis

The Dataset which is used for training and testing is taken from nltk corpora. Words are converted to featuresets, to compare with the test data. Classifiers which are used for classifying the text are:

3) Naive Bayes(original):

It is a classification technique based on Bayes Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

4) Multinominal Naive Bayes:

It is used for discrete counts. For example, let us say, we have a text classification problem. Here we can consider Bernoulli trials which is one step further and instead of 'word occurring in the document', we have 'count how often word occurs in the document', you can think of it as 'number of times outcome number x i is observed over the n trials'.

5) Bernoulli Naive Bayes:

The binomial model is useful if your feature vectors are binary (i.e. zeros and ones). One application would be text classification with 'bag of words' model where the 1s & 0s are 'word occurs in the document' and 'word does not occur in the document' respectively.

6) Linear Regression:

In this technique, the dependent variable is continuous, inde- pendent variable(s) can be continuous or discrete, and nature of regression line is linear. Linear Regression establishes a relationship between dependent variable (Y) and one or more independent variables (X) using a best fit straight line (also known as regression line).

It is represented by an equation Y=a+b*X + e, where a is intercept, b is the slope of the line and e is the error term. This equation can be used to predict the value of target variable based on given predictor variable(s).

7) SGD() Classifier:

Stochastic Gradient Descent (SGD) is a simple yet very effi- cient approach to discriminative learning of linear classifiers under convex loss functions such as (linear) Support Vector Machines and Logistic Regression.

8) SVC:

Uses a one-against-one approach

9) Linear SVC:

Uses one-against-rest approach

10) NuSVC:

The nu-SVM has the advantage of using a parameter nu for controlling the number of support vectors. The parameter nu represents the lower and upper bound on the number of examples that are support vectors and that lie on the wrong side of the hyperplane, respectively.

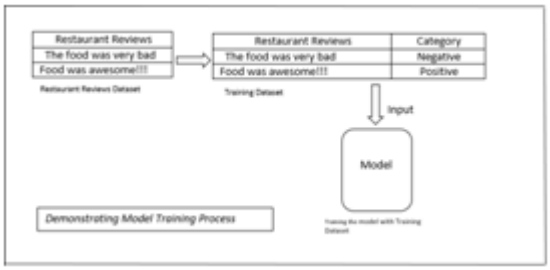

Fig. 13. Training the Model with training Data Sets

Figure 13 shows a flowchart which depicts the Training process which is used for training the SA Model.

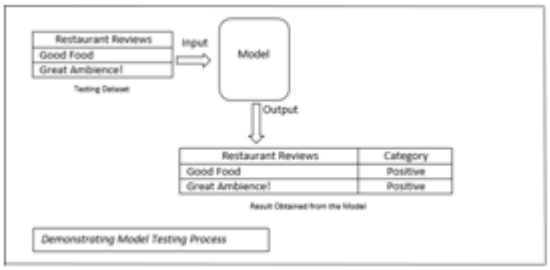

Fig. 14. Testing the Model with Training DataSets

Figure 14 shows a flowchart which depicts the Testing process which is used for testing the SA Model.

Python library called 'Pickle' is used for storing the classifiers so that we don't need to run the classifiers again and again. 'Voted Classifier' is used as a highlight of the system, where the highest vote is taken from the above-defined classifiers and given as the output.

The generated query will then be converted into speech again using the Espeak module. The response time of the Food

Ordering System is based on the internet connectivity of the user.

Once the user is done with everything his details are stored in the database with attributes such as:

- total bill
- name
- feedback
- sentiment processed

V. DEPLOYMENT HARDWARE

The AI response module is deployed on a raspberry pi model 3. The Raspberry Pi has a speaker and a microphone module attached for user input as voice and a voice output from the AI response module. The raspberry pi runs Linux and the AI response module is run on the raspberry pi in order to get input from the user.

VI. RESULTS AND DISCUSSION

Fig. 15. Chat between Customer and Device

Fig. 16. Sentiment Analysis of Feedback

Figure 15 shows the chat between customer and device which starts from asking name, their order and finally after asking the bill, the customer is asked to give feedback which is then stored in the database(MongoDB).

Figure 16 shows the query, in the form which is stored in the database.

Figure 17 shows the comparison between all the classifiers which are used for training the Sentiment Analysis Model. The ultimate result i.e mean of all classifiers to the Voted Classifier wisely supported by BernoulliNB, Logistic Regression, and Liner SVC is seen at 67%.

This training model will help us diagnose the sentiment/feedback of the customer as positive, negative, or neutral.

VII. CONCLUSION

A familiar human-like language makes people more comfortable to interact to a machine. Furthermore with the proliferation of Smart Speaker like Amazon Echo, Google Home, Apple Homepod, the use of voice is starting to become a new trend. So following the footsteps of the same, we have tried to build up a system which can be used in restaurants for removing the bottleneck caused while there are less number of waiters and more customers. The input and the output of the food ordering system is handled through NLP and Rule Based System. The database which is used in the backend to store the user input is Mongo DB which is a no-sql database. Through proper scrutinization we have concluded that using a no-sql database will be better for our future releases than a relational database since a Machine Learning module for feedback analysis is also included in the food ordering device. This module collects the feedback of the customers and analyzes it which can be then further utilized by restaurants for improving themselves. From the various results obtained through testing the device in different environments, we conclude that the system provides the output in an average span of 5 seconds.

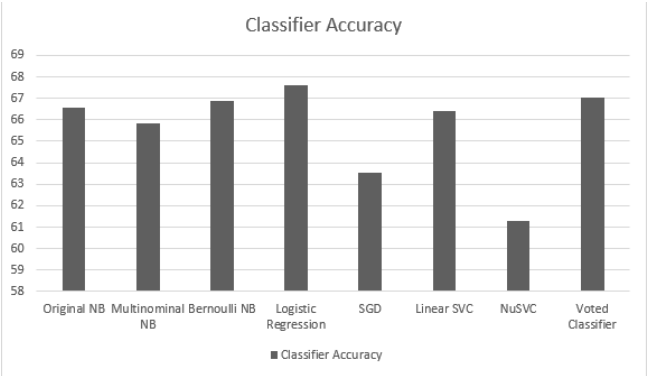

Fig. 17. Graphical representation of all classifiers performance comparison

VIII. FUTURE WORK

In future, combining Machine Leaning algorithms along with data analysis we are planning to propose a system which can be used for ordering food in the restaurants along with handling the inventory management for them. Furthermore, through the feedback obtained by the customers we can get different types of valueable information from the data by ap- plying different data analyzer algorithms; for example, details about popular food items among the customers, their opinion about the price ranges or food or ambience or customer sevices, etc can be gathered from the data which is stored in the database. This information can be further taken into consideration for various purposes like marketing which will be favorable for restaurants.

REFERENCES

[1] ASR [Online] Available:http://support.docsoft.com/help/whitepaper-asr.pdf (2018, April 21)

[2] MFCC [Online] Available:https://en.wikipedia.org/wiki/Mel-frequency- cepstrum (2018, April 21)

[3] Speech Recognition [Online] (2018, April 21) Available:http://what-when-how.com/video-search-engines/speech-recognition-audio-processing-video-search-engines/

[4] Lokesh Kumar, Dr. Shalini Rajawat, and Krati Joshi, ”Comparative analysis of NoSQL (MongoDB) with MySQL Database”, International Journal of Modern Trends in Engineering and Research

[5] Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, Yoshua Bengio, ”Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation”, CoRR 2014

[6] github [Online] Available: https://github.com/farizrahman4u/seq2seq

[7] wildml [Online] Available: http://www.wildml.com/2016/04/deep-learning-for-chatbots-part-1-introduction/

[8] Eril Cambria, Soujanya Poria, Alexander Gelbukh, and Mike Thelwall, Sentiment Analysis Is a Big Suitcase, IEEE 2017

[9] R. Jenitha Sri and P. Ajitha, ”Survey of Product Reviews using Sentiment Analysis”, International Journal of Science and Technology

[10] Anuja P.Jain and Asst. Prof Padma Dandannavar, ”Application of Machine Learning Techniques to Sentiment Analysis”, IEEE 2016

[11] BiswaRanjanSamal, Anil Kumar Behera, and Mrutyunjaya Panda, Per- formance Analysis of Supervised Machine Learning Techniques for Sentiment Analysis, IEEE 3rd International Conference on Sensing, Signal Processing and Security(ICSSS)

[12] Sneha Lukose and Savitha S. Upadhya,”Text to Speech Synthesizer-Formant Synthesis”, International Conference on Nascent Technologies in the Engineering Field, IEEE 2017 (2018, April 21)

[13] Text to Speech Flow [Online] Available: http://www.wezs.com/ dan-guy/monguy/TTS.html (2018, April 21)

[14] Speech Synthesis [Online] Available: https://en.wikipedia.org/wiki/Speech.synthesis (2018, April 21)

[15] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. 2014. Sequence to sequence learning with neural networks. In NIPS.