Intent Classification in Conversational System using Machine Learning Techniques

Monica Menda
Student
Department of CSE
GVPCE

G. Satya Keerthi, PhD
Faculty
Department of CSE
GVPCE

ABSTRACT
The improvement of chatbots has ended up greater famous in current days. Many conversational chatbots were created for that reason so far. Chatbots are computer programs that speak with humans in herbal language. These chatbot structures were programmed with a little expertise to understand sentences and make a judgment as a reaction to a query. On the chatbot system, a class method was changed into offered on this studies to discover the reason, person input, or reason class. To decide the reason and examine it for every other, the Naive Bayes method, Logistic Regression method, and SVM have been used in this study. The quantity of accuracy, precision, recall, and f1-score are obtained, and don't forget of each method is the parameter utilized to assess the outcomes. Based on categorization outcome, we can decide which algorithm is great for a powerful bot primarily based totally on the metrics.

Keywords
Chatbot, Intent Classification, Natural language Processing, Multinomial Naïve Bayes, SVM, Logistic Regression.

1. INTRODUCTION
Human communication relies heavily on speech and textual information. According to a recent article in "The New York Times," adults now spend more than 8 hours per day on computer or mobile device screens. As a result, major human communication takes place through web applications such as WhatsApp, Facebook, and Twitter, especially voice and text conversations. The need for chatbot systems has grown due to the widespread use of personal systems that want to communicate, as well as the desire of their creators to provide natural language for them to communicate with humans. In the implementation of artificially intelligent systems, the development of information technology and communication has been complex. Decision support systems, robotics, natural language processing, expert systems, and other systems are attempting to mimic human activities.

Furthermore, there is now a hybrid of natural language and intelligent systems that can understand human natural language. These systems can learn and refresh their knowledge by reading all electronic articles available on the internet. As a user, you can ask questions to the system in the same way that you would to another human. These systems are commonly referred to as internet answering machines.

Chatbots should be able to understand and respond to human natural language. Within the domain of Artificial Intelligence, advanced automated chatbots commonly use Machine Learning (ML) in conjunction with Natural Language Processing (NLP). Generally, the anatomy of information-seeking chatbots consists of two components: Natural Language Understanding (NLU) is used to categorize the

user's intent, and Natural Language Generation (NLG) is used to generate a natural language response.

A chatbot is a popular AI-created application. It is used to facilitate human work in humanoid robots, personal assistants, car assistance, and other applications. The system can understand natural language that is input by the user; in this study, the classification method can understand the user's text and proposes using classification to determine intent classification so that the system can provide answers based on intent classification.

2. BACKGROUND AND RELATED WORK
When designing a chatbot, it is important to understand the key characteristics of the chatbot. These traits were discovered by investigating people's expectations of chatbots.. M.Anupam designed a chatbot that employed an ensemble learning method as a random forest. The validation system offers an average F-measure on various K-values under random forest for this chatbot[1].B.Setiaji designed a knowledge-based chatbot where the researcher make sentence similarity calculation using the bigram method which divides input sentence into two strings of input sentence[2].V.Selvi proposed a bully-free chatbot such a way, the concept of cyberbullying in two-way chat was constructed using Linear Regression[3].P.Jitendra accelerates the interview process chatbot that would conduct interviews by analyzing the curriculumvitae(CV)[4].Y.W.Chandra built a question answering system using a sequence-to-sequence model[5].Polignano presented HealthAssistantBot which acts as a conversational agent for monitoring treatments, suggests doctors, and performs self-diagnosis. Symptom Checker module represented to identify the patients' disease with a degree of accuracy by taking a set of symptoms[6].Prissadang Suta gives an overview of Machine Learning in Chatbots This paper outlines many machine learning techniques which could improve the performance of chatbots because they allow chatbots to learn and adapt through experience[7].

3. INTENT CLASSIFICATION
The automated categorization of textual data based on customer goals is known as intent classification. In general, an intent classifier automatically analyses texts and categorizes them into customer-specified intents such as Purchase, Downgrade, Unsubscribe, and Demo Request. This is useful for deciphering the motivations behind customer inquiries, automating the process, and gaining valuable insights.
It employs machine learning and natural language processing to generate words or expressions with specific intent. A machine learning model can learn that words like buying or acquiring, which are frequently associated with the intent to purchase, are associated with the intent to purchase. And the intent here refers to the user's intention. For example, if a user

types "show me yesterday's sports news," this indicates the user intends is to retrieve a list of sports news from the whole corpus.

4. ALGORITHMS USED TO BUILD THE MODEL

4.1. NAÏVE BAYES

Naive Bayes classifiers are a type of probabilistic classifier that is based on the Bayes Theorem and makes a strong assumption of feature independence. The Naive Bayes classifier is a collection of many algorithms that all share one common principle: each feature being classified is unrelated to any other feature. The presence or absence of one feature does not affect the presence or absence of another.Naive Multinomial Bayes is one of the most widely used supervised learning classifications for categorical text data analysis. These are not only fast and reliable but also simple and easy-to-use classifiers that are proving their worth in the world of machine learning. Despite its simplicity, it predicts correctly in text classification problems. They are probabilistic classifiers that use the Bayes theorem to calculate the conditional probability of each label given a given text, with the label with the highest conditional probability being output. It has been widely used in text classification in recent years. These are extremely simple, fast, readable, and dependable algorithms.

4.2. SVM

A support vector machine is an algorithm that determines the optimal decision boundary between vectors that belong to a particular group (or category) and vectors that do not. This can be applied to any type of vector which encodes any type of data. In other words, to take advantage of the SVM text classification feature, you need to convert the text to a vector. Therefore, when the SVM determines the above decision limits, it determines where to draw the optimal "line" (or optimal hyperplane) that divides the space into two subspaces. One is for vectors that belong to a particular category and the other is for vectors that do not.

4.3. LOGISTIC REGRESSION

Logistic Regression is a supervised machine learning algorithm that predicts the probability of a target variable. This type of regression allows us to test the effect of numerical factors on binary responses.It includes the sigmoid function, which is used to calculate the logistic model where each value is provided as limited input and output in the range of 0 and 1. Logistic Regression is used to study data where there are one or more independent variables that know the results.
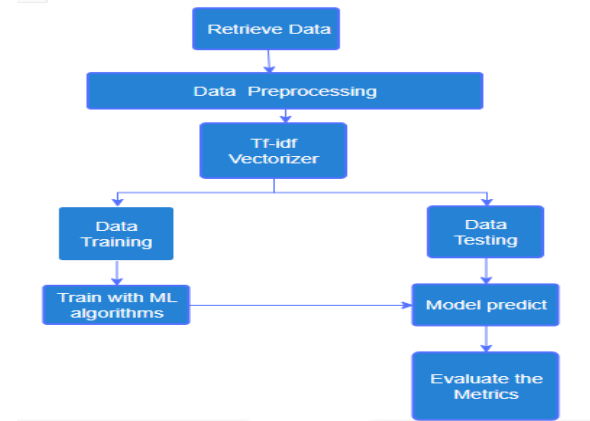
5. ARCHITECTURE OF SYSTEM

Fig1:Overview of the model

In this project, intent classification has been done by using three algorithms i.e. Multinomial Naïve Bayes, Logistic Regression, and SVM. The first process is to retrieve the data, where the intents are labeled. After that some of the preprocessing methods are applied like text lowering, removal of punctuations, stop words, frequent words, and rare words. The procedures of NLP, stemming and lemmatization are also applied for normalizing data. The preprocessed data are used for feature extraction. The next step is to split the data for training and testing. In training, the data had been trained with ML algorithms for the detection of classification. Later evaluation of the model was done by considering the metrics.

6. MATERIALS AND METHOD

6.1. DATASET

ATIS dataset provides a large number of utterances and their associated intents that can be used in training a classifier.The number of attributes in the dataset is 21 and the number of records is 4978.

Table1: Dataset

Intents	Sentences
atis_flight	I want to fly from Boston at 838 and arrive in Denver at 1110 in the morning
atis_flight	What flights are available from Pittsburgh to Baltimore on Thursday morning
atis_flight_time	What is the arrival time in SanFrancisco for the 735 am flight leaving Washington
atis_airfare	The Cheapest airfare from Tacoma to Orlando

6.2. PREPROCESSING

Preprocessing is a method of cleaning up text data so that it can be fed into the model. Text data contains noise in various forms such as emotions, punctuation, and text in each case. Text lowering, stop word removal, frequent wordremoval,stemming and lemmatization were completed in this step.

Fig2: Preprocessing Steps

```
0                  want fli 838 arriv 1110 morn
1                  avail baltimor thursday morn
2                  arriv time 755 leav washington
3                  cheapest airfar tacoma orlando
4        round trip fare philadelphia 1000 dollar
Name: text_stemmed, dtype: object
```

Fig3: Stemming

```
0         want fly boston 838 arrive denver 1110 morning
1         flight available pittsburgh baltimore thursday...
2         arrival time san francisco 755 flight leaving ...
3                          cheapest airfare tacoma orlando
4         round trip fare pittsburgh philadelphia 1000 d...
Name: text_lemmatized, dtype: object
```

Fig4:Lemmatization

6.3. FEATURE EXTRACTION

Machine Learning algorithms generate output for test data by learning from a pre-defined set of features in the training data. The main issue with working with language processing is that machine learning algorithms cannot work directly on raw text. As a result, feature extraction techniques are required to convert text into a matrix (or vector) of features. TF-IDF is one of its techniques.

Term Frequency-Inverse Document Frequency: It consists of two parts.

- Term frequency defines how frequently a term occurs in a document.

- Inverse Document Frequency describes the weight of rare words. The words which occur rarely in the corpus have a high IDF score.

doc_1="i want to fly from Boston at 838 am and arrive in Denver at 1110 in the morning"

doc_2 = "what flights are available from Pittsburgh to Balti more on Thursday morning"

Fig5: Tf-Idf Vectorizer

6.4 CLASSIFICATION TECHNIQUES

To predict the intent of the text, given by the user, the Multinomial Naïve Bayes, SVM, and Logistic Regression are used to determine the intents of different labels. 75% is given as training data set and 25% is given as testing data set input to the algorithms.

7. RESULTS AND DISCUSSION

In this section, the results of the experiment and their significance are discussed based on the following figures, where they refer to the complete evaluation metrics results (accuracy, precision, recall, F1-Score) on each algorithm.

Fig6:Result of Multinomial Naïve Bayes

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Fig7:Result of SVM

Fig8:Result of Logistic Regression

Table2:Result Analysis

ALGORITHMS	ACCURACY
Logistic Regression	0.92
Support Vector Machine	0.95
Naïve Bayes	0.86

8. EVALUATION METRICS

The effectiveness of a proposed model can be determined by applying a few evaluation metrics to calculate how accurately the model can determine the intents. In this research, three machine learning algorithms have been constructed. So, to review these models the standard evaluation metrics used by the research community are applied to them.

Accuracy
Accuracy is a popular metric that represents the percentage of correctly predicted observations, whether true or false.
Precision

The proportion of correctly predicted positive observations to all predicted positive observations is defined as precision.

$$Precision = \frac{TP}{TP + FP}$$

Recall
The Recall is the ratio of correctly predicted positive observations to all observations in an actual class.

$$Recall = \frac{TP}{TP + FN}$$

F1-Score

F1 Score is the weighted average of precision and recall.

$$F1\ Score = \frac{2 * (Precision * recall)}{Precision + recall}$$

Algorithms	Precision	Recall	F1 Score
Logistic Regression	0.922	0.922	0.923
Support Vector Machine	0.945	0.945	0.955
Naïve Bayes	0.857	0.856	0.867

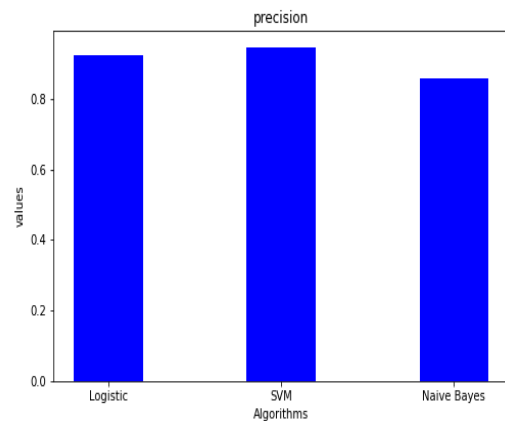**Table3: Precision, Recall, F1 score of each algorithm**
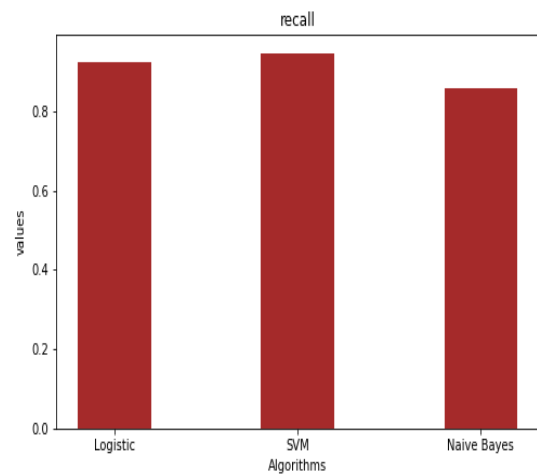

Fig 10:Graph Representation of Precision Test

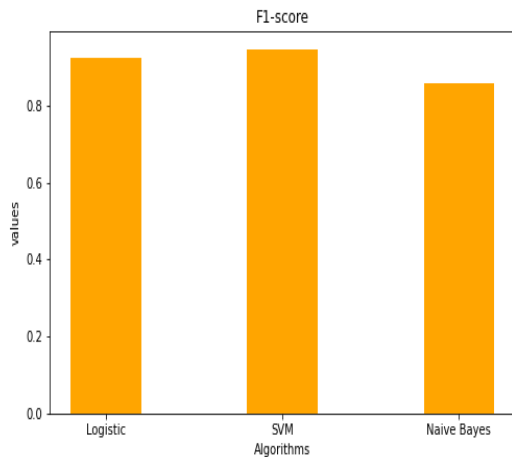**Fig 11:Graph Representation of Recall Test**

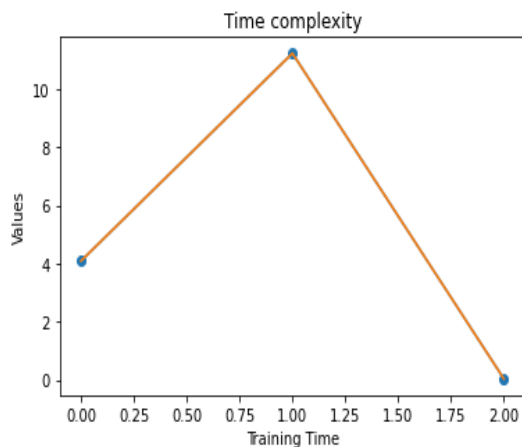**Fig 12:Graph Representation of F1 score Test**

Fig 13:Graph Representation of Time Complexity

9. CONCLUSION

 The result analysis of the experimental classification model is based on categorization outcome of three algorithms that are Naïve Bayes, Logistic Regression and Support Vector Machine are compared to decide the intent classification on chatbot system. The accuracy of Naïve Bayes model is 0.86, the accuracy of logistic regression is 0.92 and SVM is 0.95.In the evaluation results, of this experiment Precision, Recall and F1-score are also obtained. SVM is giving a better performance for the classification model. The classification of chatbot system was done on textual data. The present trending are voice notes and pictures that are using in chatbot. So that the voice notes and picture data can be added along with textual data.
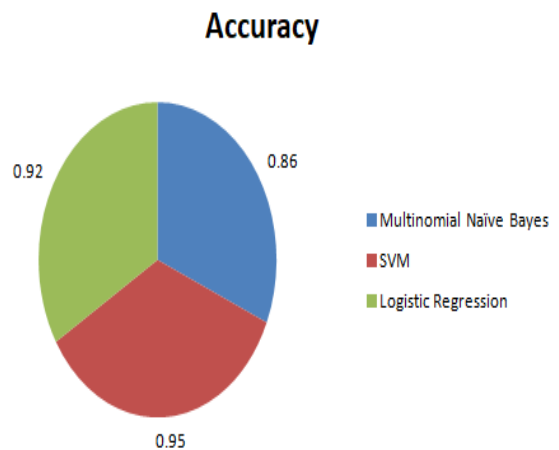

Fig 14:Pie Chart Representation of Accuracy

10. REFERENCES

[1] A. Mondal, M. Dey, D. Das, S. Nagpal and K. Garda, "Chatbot: An automated conversation system for the educational domain," *2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP),* Pattaya, Thailand, 2018, pp. 1-5, doi: 10.1109/iSAI-NLP.2018.8692927.

[2] B. Setiaji and F. W. Wibowo, "Chatbot Using a Knowledge in Database: Human-to-Machine Conversation Modeling*," 2016 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, Bangkok, Thailand, 2016, pp. 72-77, doi: 10.1109/ISMS.2016.53.

[3] V. Selvi, S. Saranya, K. Chidida and R. Abarna, "Chatbot and bully-free Chat," *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN),* Pondicherry, India, 2019, pp. 1-5, doi: 10.1109/ICSCAN.2019.8878779

[4] J. Purohit, A. Bagwe, R. Mehta, O. Mangaonkar and E. George, "Natural Language Processing based Jaro-The Interviewing Chatbot," *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC),* Erode, India, 2019, pp. 134-136, doi: 10.1109/ICCMC.2019.8819708.

[5] Yogi Wisesa Chandra, Suyanto Suyanto "Indonesian Chatbot of University Admission Using a QuestionAnswering System Based on Sequence-to-Sequence Model," 4th International Conference on Computer Science and Computational Intelligence 2019(ICCSCI).

[6] M. Polignano, F. Narducci, A. Iovine, C. Musto, M. De Gemmis and G. Semeraro, "HealthAssistantBot: A Personal Health Assistant for the Italian Language," in IEEE Access, vol. 8, pp. 107479-107497, 2020, doi: 10.1109/ACCESS.2020.3000815.

[7] Prissadang Suta , Xi Lan , Biting Wu, Pornchai Mongkolnam and Jonathan H. Chan, "An Overview of Machine Learning in Chatbot," International Journal of Mechanical Engineering and Robotics Research Vol. 9, No. 4, April 2020.

[8] E. Handoyo, M. Arfan, Y. A. A. Soetrisno, M. Somantri, A. Sofwan and E. W. Sinuraya, "Ticketing Chatbot Service using Serverless NLP Technology," *2018 5th*

International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), Semarang, Indonesia, 2018, pp. 325-330, doi: 10.1109/ICITACEE.2018.8576921.

[9] A. Argal, S. Gupta, A. Modi, P. Pandey, S. Shim and C. Choo, "Intelligent travel chatbot for predictive recommendation in echo platform," 2018 IEEE 8th Annual Computing and Communication Workshop and

Conference (CCWC), Las Vegas, NV, USA, 2018, pp. 176-183, doi: 10.1109/CCWC.2018.8301732

[10] E. Handoyo, M. Arfan, Y. A. A. Soetrisno, M. Somantri, A. Sofwan and E. W. Sinuraya, "Ticketing Chatbot Service using Serverless NLP Technology," *2018 5th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, Semarang, Indonesia, 2018, pp. 325-330, doi: 10.1109/ICITACEE.2018.8576921.