Project Report On

“Intent Recognition and Classification from Conversations”

Project submitted

in partial fulfilment of the necessities of the degree of

Master of Computer Application

By

Sumit Kumar Mandal

Roll No.**: 001910503013**

Registration No.**: 149876 of 2019-20**

Under the supervision of

Dr. Dipankar Das

Department of Computer Science and Engineering

Faculty of Engineering and Technology

Jadavpur University

Kolkata – 700032, India

JADAVPUR UNIVERSITY
FACULTY OF ENGINEERING AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Certificate of Recommendation

This is to certify that Sumit Kumar Mandal (Registration No.: 149876 of 2019-20, Roll No.: 001910503013, Exam Roll No.: MCA226013) is a student in Master of Computer Application course and the project "***Intent Recognition and Classification from Conversations***" is a record of work carried out by him, is accepted in partial fulfilment of the requirement for the degree of Master of Computer Application from the Department of Computer Science and Engineering, Jadavpur University during the academic year 2021-2022. He has been able to follow all the instructions in a calm and responsible way and successfully carried out his research work. All the best wishes to him for his future endeavours.

Dr. Dipankar Das (Project Supervisor)

Assistant Professor, Dept. of Comp. Science & Engineering

Jadavpur University, Kolkata-700032

Prof. Anupam Sinha

Head of the Department, Dept. of Comp. Science & Engineering

Jadavpur University, Kolkata-700032

Prof. Chandan Majumdar

Dean, Faculty Council of Engineering & Technology

Jadavpur University, Kolkata-700032

JADAVPUR UNIVERSITY
FACULTY OF ENGINEERING AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Certificate of Approval

The foregoing project entitled as "***Intent Recognition and Classification from Conversations***" is hereby approved as a creditable study of Master of Computer Applications and presented in a manner satisfactory to warrant its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion therein but approve this project only for the purpose for which it is submitted.

Signature of Examiner 1	Signature of Examiner 2
Date:	Date:
Name:	Name:

JADAVPUR UNIVERSITY
FACULTY OF ENGINEERING AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Declaration of Originality and Compliance of Academic Ethics

I, hereby declare that this project contains literature survey and original research work by the undersigned candidate, as part of his Master of Computer Application studies.

All information in this document have been obtained and presented in accordance with academic rule and ethical conduct.

I also declare that, as required by this rules and conduct, I have fully cited and reference all the materials that are not original to this work.

Signature

Name: **Sumit Kumar Mandal**

Class Roll No.: **001910503013**

Exam Roll No.: **MCA226013**

Project Title: "***Intent Recognition and Classification from Conversations***"

JADAVPUR UNIVERSITY
FACULTY OF ENGINEERING AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Acknowledgement

With my most sincere and gratitude, I would like to thank Prof. (Dr.) Dipankar Das, Department of Computer Science and Engineering, my supervisor, for his overwhelming support throughout the duration of the project. His motivation always gave me the required inputs and momentum to continue with my work, without which the project work would not take this current shape. His valuable suggestion and numerous discussions have always inspired new ways of thinking. I feel deeply honoured that I got this opportunity to work under him.

I would like to express my sincere thanks to all my teachers for providing sound knowledge base and cooperation.

I would like to thank all the faculty members of the Department of Computer Science and Engineering of Jadavpur University for their continuous support.

Last, but not the least, I would like to thank my batch mates for always staying by my side.

Thank you.

Sumit Kumar Mandal

Roll No.: 001910503013

Exam Roll No.: MCA226013

Table of Contents

Abstract

In this digital world, an automated chatbot that could understand human intent and provide necessary information and advices will definitely reduce human workload. For training a chatbot, a large number of data is required. Along with it, intent words are also required that will teach the model how to predict the outputs correctly. Here, we have two conversational datasets, one about the awareness of COVID, and the other about legal domain. The COVID dataset consists of 36 conversations, with 713 utterances. The legal dataset consists of 235 conversations, with 3178 utterances. We have implemented different machine learning and deep learning methods. Random Forest and Support Vector Machine giving the best results of accuracy just over 77%.

1. INTRODUCTION:

Natural Language Processing or NLP is a part of Artificial Intelligence that is used by machines to understand, analyse, manipulate, and interpret human languages. There are five main phases of NLP – morphological analysis, syntax analysis, semantic analysis, discourse analysis, and pragmatic analysis. The main application of NLP are email filters, smart assistant, text predictions, etc.

Discourse analysis is one of the phases of NLP. Discourse analysis can be defined as the process of determining contextual information. It involves the examination of how the ways of speaking about things normalizes and privileges some frames of thinking about things while marginalizing others. Discourse analysis is widely used for television analysis, market text examining, film review examining, etc.

Intent Recognition is a component of discourse analysis. It is the task of taking a spoken or written input and classifying it based on what the user wants. Intent recognition is used in many places. The most example will be the google search bar. Google takes up the search text and identifies intent words and provide a search result. Intent recognition is also used in chatbots. The most common example of chatbot is customer support of Amazon. Where they take the input text of customer and through intent recognition recognizes the intent of the user and provides necessary support. For example, if we consider a sentence, “In the year 2020, a pandemic hit the world and the world came to stand still”, intent word will be, “pandemic”.

In the present work, we have developed two datasets; one is on COVID awareness and another on the legal domain. We employed various machine learning and deep learning techniques to identify and classify intents from the chat utterances. We have observed that the performance of the system drops while the number or classes of the fine-grained intents increases. The ambiguities occur in case of close level intents.

Finally, the rest of the draft is organized as follows. The literature survey is discussed in Section 2. Two datasets are described in Section 3. Section 4 illustrates the machine and deep learning frameworks used to classify the intents whereas the experiments and results are discussed in Section 5. Error analysis and important observations are mentioned in Section 6 and Section 7, respectively. Finally, Section 8 concludes the draft with future directions.

1. 1. CHALLENGES:

A chatbot that can understand the human intent and can converse with the user to get to the exact topic and help the user with appropriate solutions is always helpful. Be it the information and awareness about COVID or any kind of legal advice, anyone would take it if they can get that information by a single touch of their finger. Thus, it can be really helpful if it can give solutions accurately. A large amount of data is needed for training the model.

1. 2. GAP IDENTIFICATION:

In the recent time, many researchers have been successful in intent recognition. But a dataset on dialogue between two individuals has never been worked on before. Thus, I have chosen this project to work on a dataset that is build upon conversations between individuals.

1. 3. MOTIVATION:

The chatbots are becoming a kind of personal assistant for all of us. Google assistant is a brilliant example of chatbot. Intent recognition, being an essential component of chatbot, has always been a field of interests for the researchers. Several researches were carried out on IR prior to my work, but I have chosen the conversational dataset to work on IR.

1. 4. PROBLEM STATEMENT:

In a time when everything is digitized, automation is the future. If there can be chatbot that can give solution, or legal advices, or may be some information about the latest pandemic, it would be very helpful. Intent Recognition can help developing a chatbot that can solve this problem. Intent recognition is an essential component for chatbots. There are many models for intent recognition and classification, like Logistic Regression, Support Vector Machine, Random Forest, etc.

For this huge amount of data is required. Then the raw data needs to be processed and then trained before feeding it into the model. Well-labelled data will be fitted into the model. This well-labelled data will teach the model how to predict the output correctly. The result will show the accuracy of how well the model will predict the output.

1. 5. OBJECTIVE:

The objective of this project is to build a model that can accurately predict the topic of what a user wants to achieve. This model will further help in building a chatbot system that can accurately predict the user intent and can give accurate solutions to the user. The goal of this project is to use different machine learning and deep learning methods that will train the model in predicting the output correctly.

1. 6. CONTRIBUTION:

The biggest and primary challenge for this project was to collect a large amount of data. We collected the data by chatting with friends, through text conversations, etc. I have presented in the report two separate datasets. Both the datasets have utterances, their corresponding intents and sentiments. I played the role of Annotator2 in the sentiment classification part and noted down the sentiments of each utterance. I also calculated the fine grained detailed of both the datasets. Fine grained details include the total number of utterances for both the datasets, total number of conversations for both the datasets, the total number of intents in both the datasets, and many more. I have mentioned the complete details in Sub-section 3.1 and Sub-section 3.2. I have calculated the contribution matrix that helped in sentiment analysis. I have also calculated the total number of utterances for each sentiment class for both the annotators, the total number of utterances where both the annotators have noted same sentiment in a relaxed manner (either positive, or negative, or neutral), the total number of utterances where both the annotators have noted exact same sentiment in a strict manner.

2. RELATED WORK:

Intent Recognition and Classification is a subfield of Natural Language Processing which is further a part of Artificial Intelligence. Researchers have always shown their interests on this topic.

In [i], a language model had been proposed where a set of words was defined. The set of words contained the main keywords for queries. The model used difference distribution, mutual information, the usage rate as anchor texts and the POS information for Intent.

In [ii], authors have implements LSTM model on question-answering problem. First the problem was broken down into simpler parts and then each simpler part was solved one by one. By this modular approach the solution was confined to only giving answers to questions.

In [iii], Intent classification methods were implemented to understand the intent of users behind their web searches to make the search results more efficient. The approaches were divided into three categories – navigational, informational, and transactional. Naïve Bayes, Random Forest were used. Naïve Bayes gave an accuracy of 88%, while Random Forest gave an accuracy of 91%.

In [iv] authors used LSTM followed by CRF and SoftMax that explicitly models the dependencies between semantic levels for better understanding. An accuracy of 94.85% was achieved by LSTM model.

3. DATASET PREPARATION:

The idea is to develop a chatbot that will read the user text and will identify the topic or the context of the text. The chatbot will also have to be able to converse with the user and find out exact topic and provide the necessary solutions to the user.

For this the models implemented requires a large set of data, and collecting this amount of data is a challenge itself. In order to cope with this challenge, I have prepared two separate conversational datasets; one on the effects of COVID-19 and the other is on the legal domain.

Few guidelines have been followed in preparation of the dataset. For every utterance, two annotators, say Annotator1 and Annotator2, will read the utterance and manually note down the intent words. If there are more than one intent word for any utterance, then the intent words should be separated by an '@'. There is no restriction given to the annotators regarding the length of the intent words. So, the intent words can be a word, or even a phrase. Apart from intent word(s) for each utterance, special intents like 'greet', 'thank', 'bye', 'personal_info', etc. are also used according to the annotators' accordance. For those utterances where no intent word(s) has been found and any of the special intents cannot be applied, N/A has been given in the intent column. Two annotators also noted the sentiments for each utterance from a range of -5 to +5. -5 being the most negative sentiment while +5 being the most positive sentiment.

3.1. DATASET ON COVID EFFECTS:

The main goal is to create a chatbot that will spread awareness, provide necessary solutions and information to the user. The dataset is based on the conversation between two users, say User1 and User2. There are 36 conversations and 713 total utterances out of which User1 has 357 utterances and User2 has 356 utterances. Two annotators noted down the intents of each utterance, and two annotators noted down the sentiments of each utterance. The sentiment score is given in a range of -5 to +5. The detailed description of the dataset is given in the table below.

The confusion matrix for the dataset is shown below:

Fig. 1: Confusion matrix for COVID dataset.

		Annotator2										
		-5	-4	-3	-2	-1	0	1	2	3	4	5
Annotator1	-5	22	20	5	0	0	0	1	0	0	0	0
	-4	0	23	23	2	0	0	0	3	0	0	0
	-3	2	8	41	8	1	0	1	1	0	0	0
	-2	0	2	16	31	14	1	2	4	0	0	0
	-1	2	4	6	36	14	1	12	5	2	0	0
	0	0	0	1	0	0	3	0	0	0	0	0
	1	0	0	7	14	13	19	49	39	1	1	0
	2	0	0	3	9	7	4	27	65	14	0	0
	3	0	0	1	6	0	2	13	19	19	0	0
	4	0	0	0	2	0	0	10	9	9	5	0
	5	0	0	0	1	0	0	14	5	3	2	4

Table 1: Detailed Description of Dataset.

Feature	**Size**
Total number of utterances	713
Total number of utterances by User1	357
Highest Length of a single utterance of User1 (by characters)	287
Lowest Length of a single utterance of User1 (by characters)	3
Average Length of an utterance of User1 (by characters)	78.011
Total number of utterances by User2	356
Highest Length of a single utterance of User2 (by characters)	481
Lowest Length of a single utterance of User2 (by characters)	3
Average Length of an utterance by User2 (by characters)	97.005
Total number of conversations	36
Highest number of utterances in a conversation	41

Lowest number of utterances in a conversation	12
Average number of utterances in a conversation	19.805
Total number of Intents of Annotator1	1136
Highest number of Intents in a single utterance by Annotator1	8
Lowest number of Intents in a single utterance by Annotator1	1
Average number of Intents per utterance for Annotator1	1.593
Total number of Intents of Annotator2	1073
Highest number of Intents in a single utterance by Annotator2	6
Lowest number of Intents in a single utterance by Annotator2	1
Average number of Intents per utterance by Annotator2	1.505
Total number of positive utterances according to Annotator1	396
Total number of negative utterances according to Annotaotr1	313
Total number of neutral utterances according to Annotator1	4
Total number of positive utterances according to Annotator2	339
Total number of negative utterances according to Annotaotr2	344
Total number of neutral utterances according to Annotator2	30
Total number of occasions where both annotators agreed on whether an utterance is positive, or negative, or neutral	591
Percentage of agreement	82.889
Total number of occasions where both annotators gave the exact same Sentiment score for an utterance	276
Percentage of agreement	38.709

Fig. 2: Number of utterances for each Sentiment Class according to Annotator1

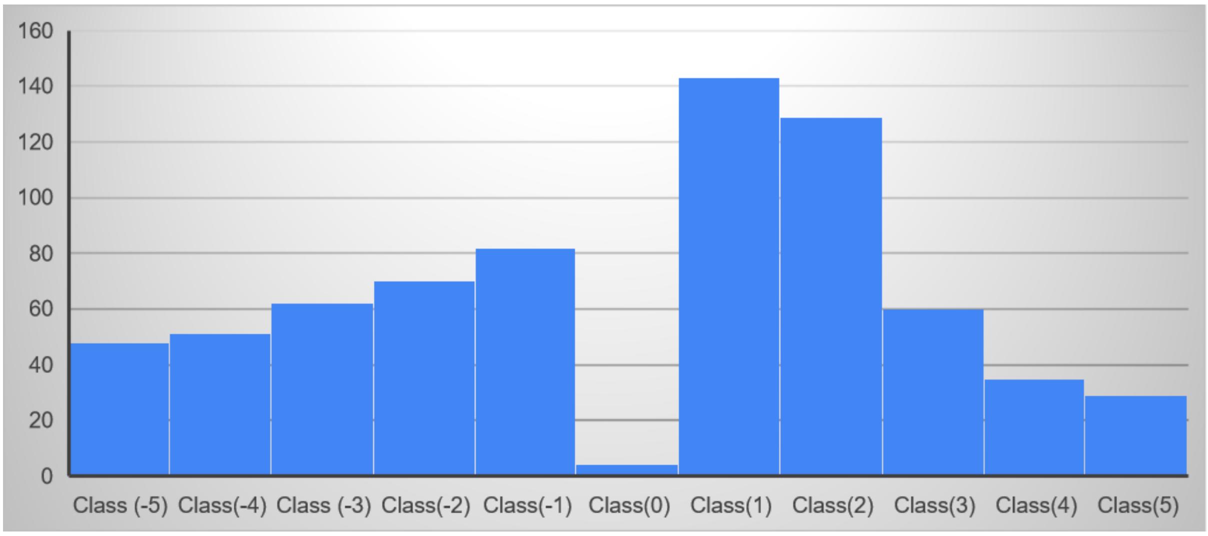
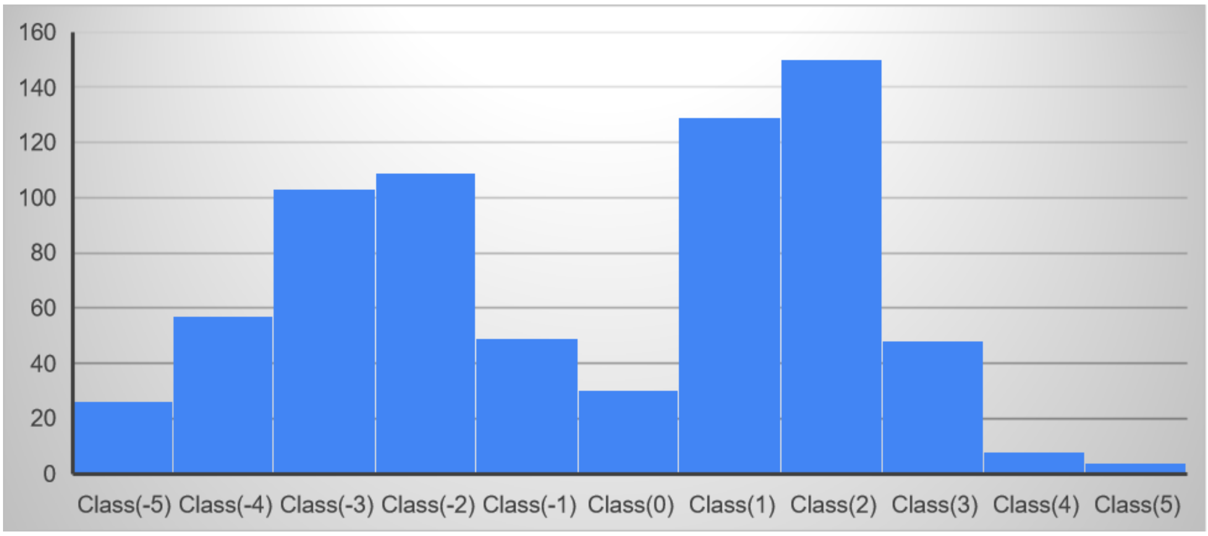


Fig. 3: Number of utterances for each Sentiment Class according to Annotator2



3.2. DATASET ON LEGAL DOMAIN:

The main goal is to provide legal advices and legal knowledge to the general public and spread awareness about the functions of the judiciary. The dataset is based on conversation between user and bot. Unlike the dataset on COVID effects, only one annotator noted down the intent words in this dataset and one annotator also noted the sentiments of each utterance. There are 235 conversations with 3178 total utterances. Out of those 3178 utterances, 1604 are user utterances and 1574 are bot utterances. Detailed description of the dataset is given in the below table.

Table 2: Detailed Description of Dataset.

Feature	Size
Total number of utterances	3178
Total number of utterances by User	1604
Highest Length of a single utterance by User (by character)	1239
Lowest Length of a single utterance by User (by character)	2
Average Length of an utterance by User (by character)	137.136
Total number of utterances by Bot	1574
Highest Length of a single utterance (by character)	1192
Lowest Length of a single utterance (by character)	4
Average Length of an utterance by Bot (by character)	115.582
Total number of conversations	235
Highest number of utterances in a single conversation	53
Lowest number of utterances in a single conversation	4

Average number of utterances in a conversation	13.523
Total number of intents	8505
Highest number of intents for a single utterance	13
Lowest number of intents for a single utterance	1
Average number of intents for an utterance	2.676
Total number of positive sentiments	243
Total number of negative sentiments	625
Total number of neutral sentiments	747

4. METHODOLOGY:

For intent recognition and classification, four supervised learning models and one deep learning model have been used. Before implementing any model pre-processing of the text data was done.

4.1. PRE-PROCESSING:

Pre-processing of the text data is the primary task of intent recognition and classification. Text data is considered as one the most unstructured data. So, the data needs to be cleaned up before fitting it into any model. Text data contains noise of various kinds, for example, mixture of uppercase and lowercase letters, punctuations, etc. When we are dealing with human language, there are many ways of saying a same thing. But a computer will understand all these things differently.

There are five steps for text pre-processing. Those are –

i. Loading the data
ii. Lowercasing the characters
iii. Tokenization
iv. Punctuation removal
v. Removal of stop-words
vi. Stemming or Lemmatization

4. 2. INTENT CLASSIFICATION MODELS:

Intent Recognition or Intent Classification is a subfield of Natural Language Processing or NLP, which is further a subfield of Artificical Intelligence. Intent recognition is the task of taking inputs from the user and recognizing the topic or context on what the user wants to convey. Intent recognition is an important component for chatbots and are widely used for customer support, sales, etc.

To work an intent classification model, we have to provide text data and alongside it we also need to provide their intents to a Machine Learning model. This is known as training the model. After training the model, the testing data is to be fit in the model and the model will predict the intents of the test data.

There are many algorithms for intent recognition, but for my project I have chosen four supervised learning models and one deep learning model.

4. 2. 1. SUPERVISED LEARNING MODELS:

Supervised learning is a type of machine learning model in which a machine is trained with well-labelled data and provides an output based on the data provided. Labelled data means those data that are already tagged with correct output. The input data works as a supervisor for the model that teaches the machines to predict the output correctly. The goal of supervised learning algorithms is to find a mapping function that will map the input variables (say x) with the output variables (say y).

The working principle of supervised learning is given in the figure below.

Let us assume that we have a dataset of shapes of different types like hexagon, square, rectangle, pentagon, triangle, etc. Firstly, we need to train the model with labelled data. For example, the model is made to learn that if the shape has four equal sides, then it is a square, if it has five equal sides then it is a pentagon, if it has three sides then it is a triangle. Then the model uses this lesson and identifies the shapes given as test data and gives the output.

The supervised learning models used in this project are Naïve Bayes, Logistic Regression, Random Forest and Support Vector Machine.

4. 2. 1. 1. NAÏVE BAYES:

Naïve Bayes algorithm is a supervised learning algorithm that is used for solving classification problems. This algorithm is based on Bayes theorem. Naïve Bayes classifier is a probabilistic classifier that predicts on the basis of probability of an object.

This algorithm is called ‘Naïve’ because it assumes that the occurrence of certain features is independent of occurrence of other features. It is called Bayes as it is based on Bayes theorem.

4. 2. 1. 2. LOGISTIC REGRESSION:

Logistic Regression is also a supervised learning algorithm and one of the most popular machine learning algorithms. This algorithm is used to predict the categorical dependent variables using a given set of independent variables. Logistic regression gives probabilistic values that lie between 0 and 1.

Two assumptions are considered in Logistic Regression. Those are:

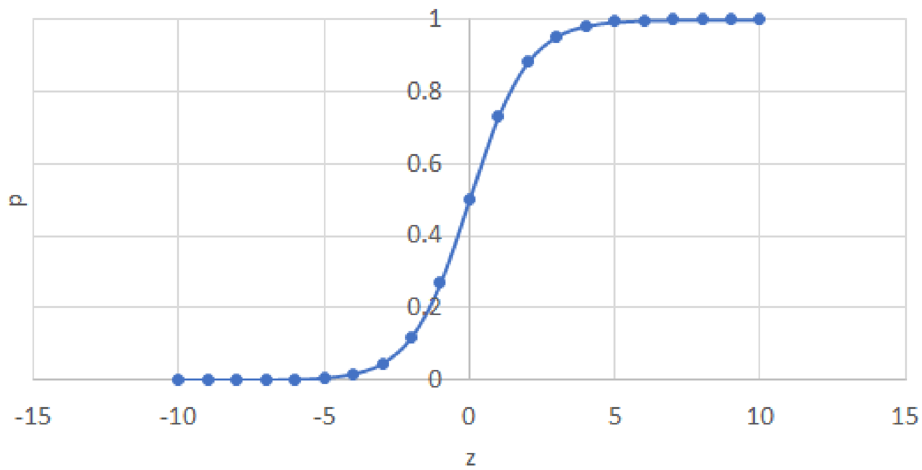
i. Dependent Variables must be categorical in nature.

ii. The independent variable should not have multi-collinearity.

$$\boldsymbol{log\left(\frac{y}{1-y}\right) = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n}$$

This is the final equation for logistic regression.

Fig. 4: Logistic Function.



4. 2. 1. 3. RANDOM FOREST:

Random Forest is also a supervised learning algorithm and one of the most popular machine learning algorithms. It can be used for both regression and classification problems. Random Forest is based on ensemble learning. Ensemble learning is a process of combining more than one classifier to improve the accuracy of the model.

As the name suggests, “Random Forest is a classifier which contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of the dataset.” Random Forest takes the prediction from each tree in the forest, instead of relying on the decision of only one tree. Thus, the greater number of trees, the better the accuracy.

4. 2. 1. 4. SUPPORT VECTOR MACHINE:

Support Vector Machine or SVM is a supervised learning algorithm. It is one of the most popular machine learning algorithms. Although it is primarily used for classification problems, SVM can be also be used for regression problems.

The primary aim of SVM is to create a hyperplane that can segregate n-dimensional space into classes so that the new data point can easily be put in the correct category. SVM chooses the extreme vectors that help in creating the hyperplane. Those extreme vectors are called support vectors.

4. 2. 2. DEEP LEARNING MODEL:

Deep learning is a machine learning process and artificial intelligence that behaves identically as a human brain gathers information or knowledge. Neural networks are the type of deep learning models that uses interconnected nodes or neurons in a layer structured identical to the human brain. Neural network is a method in artificial intelligence that make the computer learn to process the data similar to the human brain.

4. 2. 2. 1. LONG SHORT TERM MEMORY:

Long Short Term Memory or LSTM is a kind of Recurrent Neural Network or RNN that can learn order dependence. In RNN, the output of the last step is fed as an input of current step. Hochreiter and Schmidhuber designed the LSTM in order to tackle the long term dependency problem of RNN in which RNN cannot predict the word stored in memory but can give more accurate prediction from the recent information. As the gap length increase, the accuracy decreases. LSTM can by default retain the information for a long period of time. It is used for predicting, processing, and classifying on the basis of time-series data.

5. EXPERIMENTS AND RESULTS:

Four supervised learning algorithms and one deep learning algorithm has been implemented on both the datasets. Since, we have more than one intent for an utterance, thus we considered an assumption that the first intent written is the intent with the most priority. Since, the output results are varying every time, the each and every model is trained and tested five times and the average accuracy has been considered as final output. Training size of 0.8 and testing size of 0.2 has been taken into consideration.

Table 3: Experimental results for COVID dataset (Annotator1).

	Naïve Bayes	Logistic Regression	Random Forest	Support Vector Machine	Long Short Term Memory
Precision	0.0515	0.0872	0.1050	**0.1091**	0.00029
Recall	0.0824	0.1123	**0.1270**	0.1244	0.00063
F1 score	0.0567	0.0865	0.1113	**0.1134**	0.00057
Accuracy	0.1545	0.2000	**0.2818**	0.2636	0.00455

In the above table it can be seen that that accuracy and recall of Random Forest is the best, while precision and f1 score of Support Vector Machine is the best.

Table 4. Experiment Results for COVID Dataset (Annotator 2).

	Naïve Bayes	Logistic Regression	Random Forest	Support Vector Machine	Long Short Term Memory
Precision	0.0695	0.0696	**0.1562**	0.1370	0.0436
Recall	0.1165	0.1061	**0.1903**	0.1681	0.0255
F1 score	0.0822	0.0743	**0.1617**	0.1381	0.0381
Accuracy	0.1923	0.1635	**0.3019**	0.2692	0.0483

It can be seen from above table that Random Forest has the best precision, recall, f1 score and accuracy value.

Table 5: Experiment Results for Legal Dataset:

	Naïve Bayes	**Logistic Regression**	**Random Forest**	**Support Vector Machine**	**Long Short Term Memory**
Precision	0.6713	0.7187	0.7535	**0.7679**	0.5225
Recall	0.6283	0.7055	**0.7133**	0.7101	0.4391
F1 score	0.6594	0.7118	**0.7431**	0.7134	0.4823
Accuracy	0.6974	0.7405	0.7712	**0.7743**	0.5489

It can be seen from the above table that accuracy and precision score of Support Vector is the best while Random Forest has the best recall and f1 score result.

6. ERROR ANALYSIS:

Table 6: Error Analysis for Naïve Bayes.

Actual	Predict
case_file	marriage
divorce	marriage
acknowledge	thank

Table 7: Error Analysis for Logistic Regression.

Actual	Predict
harassment	inlaws_harassment
acknowledge	legal_enquiry
case_file	marriage

Table 8: Error Analysis for Support Vector Machine.

Actual	Predict
FIR_file	legal_enquiry
marriage	personal_info
legal_enquiry	marriage

Table 9: Error Analysis for Random Forest.

Actual	Predict
suggestion	legal_enquiry
maintenance	marriage
acknowledge	legal_enquiry

7. OBSERVATION AND DISCUSSIONS:

After implementing all the models on both the dataset, it is observed that the dataset for legal domain has better accuracy than the dataset for COVID. This is because legal dataset has more data than the COVID dataset. Thus, it can be clearly seen that more the data, the better the accuracy.

In the legal dataset, out of the five models, it can be seen that Support Vector Machine and Random Forest Classifier has the best accuracy, both over 77%. Long Short Term Memory, inspite of being a deep learning model has a way poorer accuracy than the supervised learning models. This is because of the lack of data. Deep learning models are very much data hungry and requires a very large amount of data to work properly. But the dataset I worked on have only over 3000 data. This is the main cause of such poor accuracy.

8. CONCLUSION:

Based on the experiments that included supervised learning models and deep learning models, it can be concluded that supervised learning models performed better than the deep learning model. Supervised learning models, Random Forest and Support Vector Machine achieved an accuracy of over 77%. Deep learning model, Long Short Term Memory achieved an accuracy score of only 55%. Thus, it can be concluded that for short datasets, supervised learning models will give better accuracy than deep learning model.

REFERENCES:

i. Kang, I.H. and Kin, G.C. *Query type classification for web document retrieval.* In proceedings of the 26th annual International ACM SIGIR conference on Research and development in information retrieval, 64-71, 2003.

ii. Gennaro, G. D., Buonanno, A., Girolamo, A. D., Ospedale, A., Palmieri, F. *Intent Classification in Question – Answering Using LSTM architechture.* Presented at the 2019 Italian Workshop on Neural Networks, 1-10, 2020.

iii. Das, A., Mandal, C., and Reade, C. *Determining the User Intent Behind Web Search Queries by Learning from Past User Interactions with Search Results*. Proceedings of the 19th International Conference on Management of Data, 1-4, 2013.

iv. Yao, K., Peng, B., Zhang, Y., Yu, D., Zweig, G., and Shi, Y. *Spoken language understanding using long short term memory neural networks.* IEEE Spoken Language Technology Workshop (SLT), 189-194, 2014.