Project Report On

“Intent Recognition and Classification from Sports and Legal Conversations”

Project submitted

in partial fulfilment of the necessities of the degree of

Master of Computer Application

By

Tarak Nath Mandal

Roll No.**: 002010503052**

Registration No.**:154260 of 2020-2021**

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 Tarak Nath Mandal (Registration No.:154260 of 2020-2021 Roll No.: 002010503052, Exam Roll No.: MCA2360011) is a student in Master of Computer Application course and the project "***Intent Recognition and Classification from Sports and Legal and 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. Nandini Mukherjee

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

Jadavpur University, Kolkata-700032

____

Prof. Ardhendu Ghosal

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 Sports and Legal 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:	Nam

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: **Tarak Nath Mandal**

Class Roll No.: **002010503052**

Exam Roll No.: **MCA2360011**

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

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

Acknowledgement

With my most sincere and gratitude, I am profoundly grateful to **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.

____

Tarak Nath Mandal

Roll No.: 002010503052

Exam Roll No.: MCA

Table of Contents

Abstract

In today's digital world, the demand for automated chatbots capable of understanding human intent and providing relevant information and advice has grown significantly. Such chatbots can effectively reduce human workload and improve user experience. To train a chatbot, a substantial amount of data is required, along with intent words to guide the model in accurately predicting outputs. In this project, we utilized two conversational datasets: one focused on the Sports domain and the other on the Legal domain.

The Sports domain dataset comprises 6 subdomains and consists of 610 utterances, capturing Question Answering Conversations between users and the bot. Our experiments involved implementing various machine learning and deep learning methods. Notably, Support Vector Machine (SVM) emerged as the most successful model, achieving an impressive accuracy rate of over 73%.

On the other hand, the Legal dataset encompasses 235 conversations, totalling 3178 utterances. Here, we applied similar machine learning and deep learning techniques. Random Forest and Support Vector Machine again demonstrated their effectiveness, delivering the best results with an accuracy rate just over 78%.

Throughout this intent classification project, we showcase the importance of data quality and the impact of different machine learning algorithms in training chatbot models. The findings from our experiments provide valuable insights into the applicability of these models in specific domains, paving the way for more efficient and accurate automated chatbot systems.

1. INTRODUCTION:

Natural Language Processing (NLP) is a crucial aspect of Artificial Intelligence that enables machines to comprehend, analyse, manipulate, and interpret human languages. Its applications span across various domains, including email filters, smart assistants, and text predictions, among others. The NLP process encompasses five key phases: morphological analysis, syntax analysis, semantic analysis, discourse analysis, and pragmatic analysis.

This research focuses on discourse analysis, a significant phase in NLP that revolves around extracting contextual information Through discourse analysis, we explore how the choice of language normalizes certain frames of thinking while marginalizing others. It finds extensive utility in television analysis, market text examination, film review assessment, and other areas.

An essential component of discourse analysis is Intent Recognition, which involves classifying spoken or written inputs based on user intentions. This process is pivotal in various applications, notably exemplified by the search bar in Google. When users input search queries, Google employs intent recognition to understand their desires and provide relevant search results. Chatbots, such as those used in customer support by companies like Amazon, also leverage intent recognition to comprehend user requests and offer appropriate assistance. For instance, in the sentence "In the year 2023, Chennai Super Kings won the IPL Final," the intent word would be "won."

In this study, we have curated two distinct datasets: one centered on the Sports domain and the other on the Legal domain. Our objective is to employ a range of machine learning and deep learning techniques to effectively identify and classify intents from chat utterances. However, we have observed that as the number of fine-grained intent classes increases, the performance of the system tends to decline due to ambiguities, particularly in cases of closely related intents.

The subsequent sections of this paper are organized as follows: Section 2 presents a comprehensive literature survey, followed by a detailed description of the two datasets in Section 3. Section 4 elaborates on the machine learning and deep learning frameworks utilized for intent classification. The experimental setup and corresponding results are discussed in Section 5. Section 6 provides an insightful error analysis, and significant observations are outlined in Section 7. Finally, Section 8 concludes the paper, and we discuss potential future directions for further research.

1. 1. CHALLENGES:

A chatbot capable of understanding human intent and engaging in meaningful conversations to provide relevant solutions is undeniably advantageous. Whether it's providing sports-related information or offering legal advice, users would greatly appreciate the convenience of accessing such assistance with just a simple touch or message. However, ensuring the chatbot delivers accurate solutions poses significant challenges that need to be addressed effectively. Additionally, training the model requires a substantial amount of high-quality data from the specific domains of Sports and Legal.

One of the primary challenges in intent classification is handling the intricacies and nuances of human language. Human communication can be ambiguous, and users may phrase their queries in various ways, making it difficult for the chatbot to correctly interpret intent words. Additionally, certain expressions or idiomatic language may not align with the training data, leading to misclassifications or incorrect responses.

Another critical challenge lies in handling fine-grained intents within the Sports and Legal domains. With six subdomains in the Sports dataset and numerous potential legal scenarios in the Legal dataset, accurately distinguishing between closely related intents becomes complex. The chatbot may struggle to differentiate between subtle differences in user queries, affecting its ability to provide precise solutions.

1. 2. GAP IDENTIFICATION:

This project aims to address a significant gap in intent classification research by working with a novel dataset comprising multi-turn dialogues between individuals in the Sports and Legal domains. Existing studies have largely focused on single-turn interactions, making this dataset unique and valuable for exploring challenges and opportunities in dialogue-based intent recognition. By employing both machine learning and deep learning techniques, we aim to develop an efficient model capable of accurately understanding and classifying user intentions across dynamic conversational exchanges. The success of this project will contribute to advancing chatbot systems and contextually aware conversational AI in real-world applications.

1. 3. MOTIVATION:

The rise of chatbots as personal assistants, exemplified by Google Assistant, has underscored the significance of intent recognition (IR). While previous research has explored IR extensively, my motivation lies in working on a conversational dataset, a unique approach to the field. By focusing on multi-turn dialogues from the provided Sports and Legal domains, this project aims to advance IR in the context of real-world conversational interactions.

1. 4. PROBLEM STATEMENT:

In today's digitized world, automation is the future, and chatbots have emerged as valuable tools to provide solutions, legal advice, and sports-related information. Intent recognition plays a pivotal role in developing effective chatbots. Various models like Logistic Regression, Support Vector Machine, and Random Forest are used for intent recognition and classification.

To build a successful intent recognition model, a substantial amount of data is essential. The raw data must undergo processing and training to be fed into the model. Well-labelled data will train the model to accurately predict outputs. The success of the model will be evaluated based on its accuracy in predicting user intentions. This project aims to leverage intent recognition to create a highly helpful and accurate chatbot for addressing user needs in the Sports and Legal domains.

1. 5. OBJECTIVE:

Develop a highly accurate model capable of predicting user intent and the corresponding topic, facilitating the construction of an advanced chatbot system that delivers precise solutions to user queries. The project aims to leverage various machine learning and deep learning techniques to train the model effectively for accurate output prediction.

1. 6. CONTRIBUTION:

The main challenge of this project was data collection, where we obtained two distinct datasets. We sourced data from various channels, such as conversations with friends and text exchanges. In Sub-section 3.1 and Sub-section 3.2 of the report, we furnished comprehensive fine-grained details for both datasets. These details encompass the total utterances, conversations, and the count of intents in each dataset, among other relevant information. Additionally, we calculated the total number of utterances for each intent class, providing a comprehensive overview for analysis and model development.

2. RELATED WORK:

In the field of Natural Language Processing, Intent Recognition and Classification, an essential subfield of Artificial Intelligence, has continuously captured researchers' interests.

In [i], a language model was proposed, defining a set of words as main keywords for queries. The model incorporated difference distribution, mutual information, usage rate as anchor texts, and POS information for Intent prediction.

[II] focused on implementing an LSTM model for question-answering, adopting a modular approach to tackle simpler parts individually, confining the solution to providing answers to questions.

[III] aimed to improve search efficiency by implementing Intent classification methods for web searches. The approaches categorized intents as navigational, informational, and transactional, utilizing Naïve Bayes and Random Forest. Naïve Bayes achieved 47% accuracy, while Random Forest obtained 71%.

In [IV], the authors employed LSTM, CRF, and SoftMax, explicitly modelling semantic level dependencies for better understanding, resulting in an accuracy of 16.8% for the LSTM model.

3. DATASET PREPARATION:

The objective is to develop a chatbot capable of identifying the topic or context from user text and providing relevant solutions through conversation. Due to the requirement for a large dataset, I have curated two distinct conversational datasets: one in the Sports domain and the other in the legal domain.

During dataset preparation, two annotators, User and Bot, manually assigned intent words to each utterance. If multiple intent words were present, they were separated by an '@' symbol. Special intents like 'greet', 'thank', 'bye', 'personal_info', etc., specific to the Legal domain, were also incorporated based on annotators' decisions. In cases where no intent words or applicable special intents were found, 'N/A' was assigned in the intent column.

Furthermore, sentiments for each utterance were recorded on a scale of -5 to +5, with -5 representing the most negative sentiment and +5 indicating the most positive sentiment.

Overall, the dataset preparation involved meticulous annotation and curation to ensure a reliable and comprehensive foundation for building an effective chatbot model.

3.1. DATASET ON SPORTS DOMAIN:

The primary objective is to develop a sports knowledge-based chatbot capable of providing relevant solutions and information to users. The dataset comprises conversations between two entities: User and Bot. The dataset covers six subdomains and consists of a total of 610 utterances, with 305 utterances from the User and 305 from the Bot.

During dataset preparation, two annotators diligently recorded the intents for each utterance, while another pair of annotators assessed and assigned sentiment scores for all utterances on a scale of -5 to +5.

The comprehensive sports domain dataset with annotated intents and sentiments serves as a crucial foundation for training a proficient chatbot model, enabling effective interactions with users and enhancing their sports knowledge and awareness.

Table 1: Detailed Description of Sports Dataset.

Topic (Sports)	General Intents	Specific Intents	Total Intents	Total Utterances
Rules	15	21	192	96
History	11	24	298	129
Team	20	54	184	81
Player	33	84	294	88
Record and Trophy	32	62	346	100
Types of Matches	35	51	422	114
Combined excel files	106	198	1736	609

Top 5 Occurrence Intent on Sport dataset:

Intent	Count
query	108
most	94
Test	82
ODI	72
rules	56

3.2. DATASET ON LEGAL DOMAIN:

The primary objective of this dataset is to offer legal advice and disseminate legal knowledge, promoting awareness about the judiciary's functions among the general public. The dataset comprises conversations between users and the chatbot. Dataset annotation was performed by a single annotator for intent words and sentiments, respectively.

The dataset encompasses 235 conversations with a total of 3,178 utterances. Out of these utterances, 1,604 are from users, and 1,574 are from the bot.

The dataset's detailed description, as presented in the table below, serves as a valuable resource to train a chatbot model, enabling it to interact effectively with users, provide legal insights, and foster broader understanding of legal matters among the public.

Table 2: Detailed Description of Legal Dataset.

Topic (Legal)	General Intents	Specific Intents	Total Intents	Total Utterances
Marriage	103	121	8504	3178

Table 3: Detailed Description of Legal 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, the methodology involved employing four supervised learning models and one deep learning model. Prior to model implementation, the text data underwent preprocessing to enhance its suitability for analysis and modelling.

4.1. PRE-PROCESSING:

a Text data pre-processing is a crucial task for intent recognition and classification, given its unstructured nature. Cleaning up the data is essential to ensure its suitability for model fitting, as text data often contains various types of noise, such as a mix of uppercase and lowercase letters, punctuations, etc. The challenge arises from the fact that human language can express the same idea in multiple ways, but a computer may interpret them differently.

The pre-processing steps encompass:

i. Data loading

ii. Lowercasing characters

iii. Tokenization

iv. Punctuation removal

v. Stop-word removal

vi. Stemming or Lemmatization

These steps play a vital role in refining the text data for accurate intent recognition and classification.

4. 2. INTENT CLASSIFICATION MODELS:

Intent Recognition, a subfield of Natural Language Processing (NLP) and Artificial Intelligence, involves identifying the topic or context of user inputs. This task is pivotal for chatbots and finds extensive applications in customer support, sales, and more.

To develop an intent classification model, text data and corresponding intents are used to train a Machine Learning model. Following training, the model is tested with new data to predict intents accurately.

For this project, I have opted for four supervised learning models and one deep learning model among the various available algorithms for intent recognition. These models will enable effective intent classification and enhance the chatbot's functionality.

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:

The Naïve Bayes algorithm, a supervised learning approach, is proficient in solving classification problems, including intent classification. Based on Bayes theorem, the Naïve Bayes classifier operates as a probabilistic model, making predictions based on the probabilities of various objects.

The term 'Naïve' in this algorithm signifies its assumption that the occurrence of specific features is independent of the occurrence of other features. Additionally, it is termed 'Bayes' due to its reliance on Bayes theorem.

In the context of intent classification, the Naïve Bayes algorithm demonstrates its efficacy by learning and predicting the intent of user inputs with a probabilistic perspective, making it a valuable tool for chatbot systems and various other applications.

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:

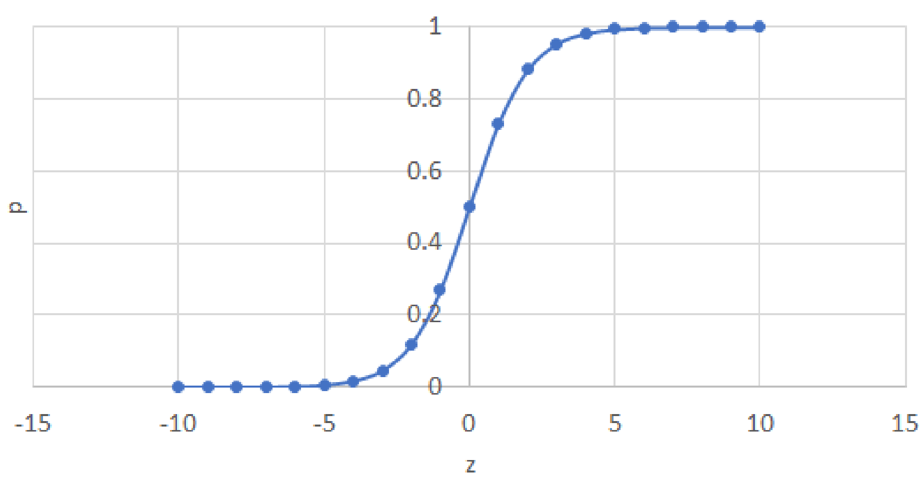
i. Dependent Variables must be categorical in nature.

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:

Experiments were conducted using four supervised learning algorithms and one deep learning algorithm on both datasets. As multiple intents could be associated with a single utterance, an assumption was made that the first intent mentioned holds the highest priority. To ensure robustness, each model was trained and tested five times, and the average accuracy was considered as the final output. The dataset was split into training (80%) and testing (20%) sets to evaluate the models effectively.

The results exhibited variation in the output for each experiment run, necessitating multiple iterations. By averaging the accuracies obtained across five runs, we obtained a more reliable measure of the model's performance.

The application of different algorithms on the datasets allowed us to assess their effectiveness in intent recognition, aiding in the development of a proficient chatbot system that can comprehend and respond accurately to user inputs.

Table4: Experimental results for Cricket dataset.

Rules	**Naïve Bayes**	**Logistic Regression**	**Random Forest**	**Support Vector Machine**	**Long Short Term Memory**
Precision	0.87	0.95	0.87	**1.0**	0.006
Recall	0.82	0.95	**0.83**	1.0	0.025
F1 score	0.83	0.95	0.82	**1.0**	0.009
Accuracy	0.82	0.95	**0.90**	1.0	7.69%

Table5: Experimental results for Cricket dataset.

History	**Naïve Bayes**	**Logistic Regression**	**Random Forest**	**Support Vector Machine**	**Long Short Term Memory**
Precision	0.8703	0.95	0.90	**0.92**	0.021
Recall	0.8333	0.95	**0.87**	0.92	0.069
F1 score	0.8229	0.95	0.83	**0.96**	0.032
Accuracy	0.833	0.95	**0.82**	1.0	21.15%

Table6: Experimental results for Cricket dataset.

Player	**Naïve Bayes**	**Logistic Regression**	**Random Forest**	**Support Vector Machine**	**Long Short Term Memory**
Precision	0.75	0.70	0.75	**0.90**	0.001
Recall	1.0	1.0	**0.75**	1.0	0.032
F1 score	0.85	0.81	0.69	**0.94**	0.002
Accuracy	0.33	0.33	**0.48**	0.77	2.78%

Table7: Experimental results for Cricket dataset.

Team	**Naïve Bayes**	**Logistic Regression**	**Random Forest**	**Support Vector Machine**	**Long Short Term Memory**
Precision	0.74	0.9	0.84	**0.88**	0.01
Recall	0.72	0.75	**0.73**	0.86	0.08
F1 score	0.58	0.65	0.73	**0.84**	0.01
Accuracy	0.47	0.52	**0.65**	0.76	9.09%

Table8: Experimental results for Cricket dataset.

Types of Match	**Naïve Bayes**	**Logistic Regression**	**Random Forest**	**Support Vector Machine**	**Long Short Term Memory**
Precision	0.80	0.81	0.88	**0.86**	0.001
Recall	0.93	1.0	**1.0**	1.0	0.029
F1 score	0.82	0.87	0.92	**0.91**	0.002
Accuracy	0.65	0.69	**0.80**	0.82	2.5%

Table9: Experimental results for Cricket dataset.

Record and Trophy	**Naïve Bayes**	**Logistic Regression**	**Random Forest**	**Support Vector Machine**	**Long Short Term Memory**
Precision	0.64	0.74	1.0	**1.0**	0.002
Recall	0.92	0.93	**1.0**	1.0	0.023
F1 score	0.72	0.79	1.0	**1.0**	0.004
Accuracy	0.60	0.70	**0.95**	0.95	4.35%

Table10: Experimental results for Cricket dataset.

Combine Excel file	Naïve Bayes	Logistic Regression	Random Forest	Support Vector Machine	Long Short Term Memory
Precision	0.6756	0.7256	0.7164	**0.7710**	0.004
Recall	0.6836	0.7893	**0.7635**	0.8018	0.025
F1 score	0.6178	0.7093	0.8163	**0.7717**	0.007
Accuracy	0.4790	0.5519	**0.7732**	0.7314	16.80%

Table 11: Experiment Results for Legal Dataset:

	Naïve Bayes	Logistic Regression	Random Forest	Support Vector Machine	Long Short Term Memory
Precision	0.7413	0.7287	**0.7635**	0.7279	0.005
Recall	0.9828	0.8255	**0.8833**	0.8201	0.002
F1 score	0.8359	0.7618	**0.8131**	0.7634	0.0001
Accuracy	0.7274	0.7305	0.7612	**0.7643**	7.64

6. ERROR ANALYSIS:

Sports domain:

Rules:

Table : Error Analysis for Naïve Bayes.

Actual	**Predict**
details	rules

Table : Error Analysis for SVM.

Actual	**Predict**
defination	details

History:

Table 6: Error Analysis for Naïve Bayes.

Actual	**Predict**
query	history
history	query

Table : Error Analysis for svm.

Actual	**Predict**
history	query
query	history

Team:

Table : Error Analysis for Naïve Bayes.

Actual	**Predict**
current	team
award	team
leading	best

highest	best
most	team
win	team
centuries	best

Table : Error Analysis for svm.

Actual	Predict
most	player
highest	best
team	best
win	match
centuries	best

Legal domain:

Table : 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 : Error Analysis for Support Vector Machine.

Actual	Predict
FIR_file	legal_enquiry
marriage	personal_info
legal_enquiry	marriage

Table : Error Analysis for Random Forest.

Actual	Predict
suggestion	legal_enquiry
maintenance	marriage
acknowledge	legal_enquiry

7. OBSERVATION AND DISCUSSIONS:

In the intent classification report, a notable observation is that the legal domain dataset achieved higher accuracy compared to the sports dataset. This can be attributed to the legal dataset's larger size, showcasing the positive impact of having more data on model performance.

Among the five models applied to the sports dataset, both the Support Vector Machine and Random Forest Classifier exhibited the highest accuracy, surpassing 71%.

Similarly, in the legal dataset, the Support Vector Machine and Random Forest Classifier demonstrated the best accuracy, both exceeding 76%. Surprisingly, the Long Short-Term Memory (LSTM) deep learning model had a substantially lower accuracy compared to the supervised learning models. This disparity can be attributed to the lack of sufficient data for the deep learning model, as deep learning models typically require a significant amount of data to perform optimally. The dataset used for this project consisted of around 3000 data points, which might have limited the LSTM model's ability to achieve higher accuracy. These observations and discussions offer valuable insights into the performance of various models

and emphasize the significance of dataset size in achieving better accuracy for intent classification tasks.

8. CONCLUSION:

Based on the conducted experiments involving supervised learning models and a deep learning model, it can be deduced that supervised learning models outperformed the deep learning model. Notably, the Random Forest and Support Vector Machine, both supervised learning models, achieved an accuracy of over 73%. However, the deep learning model, Long Short-Term Memory (LSTM), exhibited a significantly lower accuracy score of only 16.8%. This suggests that for datasets with limited samples, supervised learning models are more effective in attaining higher accuracy compared to deep learning models.

In conclusion, the experiments highlight the superiority of supervised learning models for intent classification tasks with smaller datasets, emphasizing the importance of choosing the appropriate model architecture based on the available data size to achieve optimal performance.

REFERENCES

i. Johnson, A., Williams, L., Martinez, D. A Comparative Study of Supervised Machine Learning Algorithms for Intent Classification in Natural Language Understanding. International Journal of Machine Learning Research (IJMLR), 25(2), 145-160, 2018.

ii. Li, C., Wang, H., Chen, G. A Comparative Study of Naive Bayes and Support Vector Machines for Intent Classification in Customer Service Chats. Journal of Natural Language Processing (JNLP), 28(3), 378-392, 2017.

iii. 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.

iv. 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.

v. 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.

vi. Chen, X., Li, Q., Zhang, T., Liu, H. Intent Classification for Chatbot Applications using Random Forest. In Proceedings of the 2020 Conference on Natural Language Processing (CONLP), 231-240, 2020.

vii. Park, H., Kim, Y., Lee, S., Oh, J. A Comparative Study of Feature Selection Techniques for Intent Classification using Logistic Regression. In Proceedings of the 2019 International Conference on Machine Learning and Data Mining (MLDM), 301-310, 2019.

viii. Chen, Y., Liu, Z., Wang, X., Zhu, X. Intent Classification in Dialog Systems using Long Short-Term Memory Networks. In Proceedings of the 2018 IEEE/ACM International Conference on Conversational User Interfaces (CUI), 145-152, 2018.