**Natural Language Understanding for**

**Dialog Systems Assignment**

**Documentation and Reporting**

**1. Introduction**

In this project, we aim to develop a conversational AI system capable of understanding user intents, extracting relevant entities, and providing appropriate responses. The system consists of several components:

Intent Recognition: Identifying the intent behind user queries, such as booking a flight or getting weather information.

Entity Extraction: Identifying important entities from user queries, such as locations and dates.

Slot Filling: Filling slots with extracted entities to maintain context and understand user requests fully.

Dialog Management: Generating responses based on recognized intents and filled slots.

Model Evaluation: Assessing the performance of intent recognition models, including traditional ML models and a finetuned BERT model.

**2. Data Preprocessing**

**Data Acquisition:**

We start by loading the dataset containing flight related queries and their corresponding class types.

**Intent Recognition:**

The 'Airline' column is used as input features (X), and the 'class\_type' column is used as target labels (y).

The data is split into training and testing sets using a 8020 split.

**3. Model Development**

**Intent Recognition:**

We use a pipeline consisting of TFIDF Vectorizer and a Random Forest Classifier for intent recognition.

The pipeline is trained on the training data and evaluated using accuracy, precision, recall, and F1score metrics.

**Entity Extraction:**

Named Entity Recognition (NER) is performed using the spaCy library to extract entities like dates and locations from user queries.

**Slot Filling and Context Handling:**

A slot filling function is defined to map extracted entities to specific slots, such as 'date' and 'location'.

Context management function is defined to update conversation context based on previous context, current intent, and slots.

**Dialog Management:**

A rule-based dialog management system is implemented to generate responses based on recognized intents and filled slots.

**Model Evaluation:**

The intent recognition model is evaluated using accuracy, precision, recall, and F1score.

BERT Finetuning for Intent Recognition:

BERT model is finetuned for intent recognition using the 'bertbaseuncased' model.

Data is pre-processed, tokenized, and encoded using the BERT tokenizer.

The model is trained on the training dataset, and evaluation is performed on the test dataset.

**4. Evaluation Results**

**Intent Recognition:**

**Traditional ML Model:**

Accuracy: 0.82

Precision: 0.82

Recall: 0.82

F1 Score: 0.82

**BERT Model:**

Accuracy: 0.85

Precision: 0.85

Recall: 0.85

F1 Score: 0.85

**5. Visualizations**

**NLU Module Operation:**

Intent Recognition: Visual representation of the performance metrics (accuracy, precision, recall, F1score) for both the traditional ML model and the BERT model.

Entity Extraction: Visualization of named entities extracted from sample user queries.

Slot Filling: Illustration of how extracted entities are mapped to slots.

Dialog Management: Flowchart demonstrating how intents and slots influence response generation.

**6. Conclusion**

In this project, we successfully developed a conversational AI system capable of understanding user intents, extracting entities, and generating appropriate responses. The system utilizes both traditional machine learning and stateoftheart BERT models for intent recognition, achieving satisfactory performance. Additionally, the system demonstrates effective entity extraction, slot filling, and context handling to maintain meaningful conversations with users.