Record-keeping and Reporting   
1. Overview  
This article describes how Natural Language Understanding (NLU) models and machine learning techniques were used to construct a dialog system for intent identification and entity extraction. In order to satisfy user demands, the system is made to recognize human intents from text input and retrieve pertinent items.   
  
  
2. Preprocessing Data   
Data Acquisition: Flight data that was scraped from the internet was utilized as the training and testing dataset for the model.   
The collection includes data on class types (class\_type) and airline names (Airline).

Recognition of Intent:

Classifying user intents according to the airline that is specified in the text is the task of intent recognition.  
The input feature (X) for intent recognition in the dataset is the Airline column.   
  
The target label (y) is found in the class\_type column.   
A train-test split ratio of 80:20 is used to divide the dataset into training and testing sets.   
Extracting Entities:   
In order to extract entities from user queries, such as locations and dates, Named Entity Recognition (NER) is utilized.   
For entity extraction, Spacy's English language model (en\_core\_web\_sm) is utilized

3. Development of Models   
Intent Recognition: To recognize intent, a Random Forest Classifier and TF-IDF Vectorizer are combined into a pipeline.   
  
Text input is transformed into numerical features using the TF-IDF Vectorizer.   
Based on the TF-IDF characteristics, the Random Forest Classifier is trained to categorize intents.   
Metrics such as accuracy, precision, recall, and F1-score are used to assess the model.

Slot Filling and Context Handling: In slot filling, pertinent data (such as date and location) is extracted from user requests.  
By updating the prior context with the current intent and slots, context management preserves the conversation's context

5. Evaluation Outcomes   
Intent Recognition Model: On the test set, the Random Forest Classifier achieves measures such as accuracy, precision, recall, and F1-score.

The BERT system, which stands for Bidirectional Encoder Representations from Transformers, has been fine-tuned to recognize intent.  
Input text is encoded and intents are categorized using the BERT tokenizer and model.   
  
Metrics such as accuracy, precision, recall, and F1-score are used to assess the model.

6. NLU Module Visualization and Explanation Intent Classification: In order to ascertain the intent of the user, the NLU module analyzes user queries.   
  
TF-IDF Text is converted by a vectorizer into numerical vectors that indicate the significance of each word in the document.   
Based on the vectorized features, the Random Forest Classifier forecasts the intent.

Entity Extraction: To extract entities from user queries, like locations and dates, the NLU module uses Spacy's NER model.  
Based on established labels (e.g., 'LOC' for location, 'DATE' for date), entities are detected.