

Task 02 Report

**Jayani Adikari | Nisal Liyanage | Tarinya
Perera | Satheesha Fernando**

05/04/2024

—

IntelliHack

—

IEEE Event

Contents

Chapter 01 Model Selection	3
Chapter 02 Preprocessing	3
Chapter 03 Rationale behind the Confidence Threshold	3

Chapter 01| Model Selection

For this task, a simple Naive Bayes classifier, specifically the Multinomial Naive Bayes, was chosen. Naive Bayes classifiers are known for their simplicity, efficiency, and effectiveness in text classification tasks, especially when dealing with relatively small datasets like the one that has been utilized.

```
model = make_pipeline(CountVectorizer(), MultinomialNB())
```

Chapter 02| Preprocessing

The text data is preprocessed using a CountVectorizer from scikit-learn, which converts the text inputs into numerical feature vectors by counting the occurrences of each word (token) in the text. This converts the textual data into a format that can be used by machine learning algorithms.

No further preprocessing steps such as stop word removal or stemming were applied in this case, as the dataset is relatively small, and the intent classification task is straightforward.

```
model = make_pipeline(CountVectorizer(), MultinomialNB())
```

Chapter 03| Rationale behind the Confidence Threshold

The confidence threshold, set at 0.28, serves as a critical benchmark in our system. We experimented with various threshold values to strike a balance between precision and recall. Higher thresholds didn't predict intents confidently, while lower ones led to incorrect predictions. Our chosen threshold ensures reliable predictions while tolerating some false positives. If a prediction's confidence score falls below this threshold, the system triggers a fallback response, indicating uncertainty in determining the intent.

```
C:\Users\SARANGA\AppData\Local\Programs\Python\Python38\python.exe C:/Users
Enter a text: Hi
Predicted Intent: Greet, Confidence: 0.30
Enter a text: LoL
NLU fallback: Intent could not be confidently determined
Enter a text: Go
NLU fallback: Intent could not be confidently determined
Enter a text: You are a fool
Predicted Intent: Condemnation, Confidence: 0.33
Enter a text: What is going on
Predicted Intent: Inquiry, Confidence: 0.34
Enter a text: I'm really sorry
Predicted Intent: Apology, Confidence: 0.30
Enter a text: Thanks so much brother
Predicted Intent: Thank, Confidence: 0.30
Enter a text: lmao
NLU fallback: Intent could not be confidently determined
Enter a text: something
NLU fallback: Intent could not be confidently determined
Enter a text: dude what is going on?
Predicted Intent: Inquiry, Confidence: 0.34
```