CS 445 NLP Course Project Milestone Report

Erdem Özcan - 29225 Mert Coşkuner - 29120 Kanat Özgen - 30973
Gürsel Yiğit Pekgöz - 29384

1. Introduction

This project aims to develop an intent detection mechanism using the ATIS (Airline Travel Information System) database. Intent detection is one of the fundamental components used in virtual assistants and conversational AI in natural language processing. Currently, we will use feature extraction and neural network-based classification as the main methodology. Furthermore, we are searching for various approaches for feature extraction, using Convolutional Neural Networks (CNN) or other models. Following feature extraction, we plan to implement a classifier, possibly using a neural network or an LSTM-based solution, to accurately classify the detected intent.

---

2. Dataset Selection

The selected dataset is ATIS for this task. This dataset is quite a popular dataset used in intent detection which consists of data which is related to airline travel. Furthermore, it is accurate to use this dataset because of the diverse intents and structured format which allows to efficiently perform the model training and evaluation.

The dataset is split into three parts:

- 70% for training
- 15% for validation
- 15% for testing

This split ensures a robust evaluation of the model's performance and prevents overfitting.

3. Approach Plan

Our approach focuses on feature extraction and classification. For feature extraction, we are exploring CNN-based methods to capture patterns in text. This approach is based on the studies which are demonstrated by Yoon Kim's study "Convolutional Neural Networks for Sentence Classification," which helped us to understand that CNNs could achieve high accuracy in sentiment analysis. Moreover, we are also considering alternatives like embeddings and pre-trained models such as BERT to better understand context and meaning which is inspired by the study “Self-attention networks for intent detection”.

For classification, we are considering a feedforward neural network because it is simple and efficient. We are also exploring an LSTM-based model, which is good at understanding the order of words in data and works well for intent detection. The key study shows the capability of LSTMs is the "Deep Bi-Directional LSTM Network for Query Intent Detection".

Our approach is inspired by key papers mentioned in the references part and studies in the field, including research on CNNs for sentence classification and the application of LSTMs in similar tasks. While we have not yet finalized our methodology, we are actively exploring the most effective ways to utilize extracted features and plan to make a final decision based on further experimentation and an in-depth review of the literature.

4. Next Steps

The intermediate steps will focus on the following. Firstly, understanding the data and then selecting different features to extract, and testing them out. After, carefully selecting the features we will choose the most effective features to continue with the ATIS dataset. Secondly, we will implement both neural network and LSTM-based classifiers and then evaluate them.

Thirdly, we plan to optimize our models by conducting experiments on hyperparameters and using various algorithms to optimize them.

While all members of the group will concentrate on all steps, one specific member will be selected for each task to mainly focus on that objective and understand how can we develop a further optimized technique. Moreover, these tasks include the implementation of the CNN-based feature extraction methods, investigating and evaluating alternative methods for feature extraction such as embeddings and pre-trained models like BERT, developing and testing neural networks and LSTM-based classifiers, and lastly hyperparameter optimization, overall performance evaluation to ensure the integration of feature extraction and classification stages is optimized.

---

5. References

1. Kim, Y. (2014). Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882.* https://doi.org/10.48550/arXiv.1408.5882
2. Zhang, Y., & Wallace, B. C. (2017). A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. *Proceedings of the 8th International Joint Conference on Natural Language Processing (Volume 1: Long Papers),* 253–263. https://aclanthology.org/I17-1026/
3. Xu, H., Wu, F., Wu, C., & Yuan, Y. (2020). Self-attention networks for intent detection. *arXiv preprint arXiv:2006.15585.* https://doi.org/10.48550/arXiv.2006.15585
4. Sreelakshmi, K., Rafeeque, P. C., Sreetha, S., & Gayathri, E. S. (2018). Deep Bi-Directional LSTM Network for Query Intent Detection. *Procedia Computer Science, 143*, 939–946. https://doi.org/10.1016/j.procs.2018.10.394