Intent Detection System Based on BERT

Abhisht Makarand Joshi – 300311048
Nidhi Kumari Chauhan – 300279256
CSI 5180 - Topics in AI: Virtual Assistant
University of Ottawa
Project Proposal

1. Project Introduction:

In natural language processing (NLP), intent detection is a crucial task that aids in determining the underlying goal or purpose of a user's input, such as a text message or voice command, to offer the right response. It helps machines to interpret human language and respond accordingly. Intent detection, for instance, can assist in a chatbot or virtual assistant in determining what the user wants or asks for and providing the appropriate response.
Our project's objective is to develop a reliable intent detection system employing an optimized model, a cutting-edge language model that has showed promising results on a variety of natural language processing tasks. We will first gather a large dataset of texts and the corresponding intents to build our intent detection system. We propose to start with Multinomial Naive Bayes as a baseline model for intent classification, and then incorporate advance language models like BERT and its variants to boost performance measure such as accuracy/loss. Overall, the experiment demonstrates how effective the BERT model is at determining intentions. Additionally, it will demonstrate how crucial it is to fine-tune BERT model for the most effective performance.

2. Goals and Limitation

- **What will you achieve (learn + produce) and what is your prior knowledge of this task?**
  This project aims to tackle the intent detection problem using Natural Language Processing and Deep Learning Methods. It will help us learn more about how the intent detection phase of a virtual assistant works. It's a crucial phase that takes place after the speech has been transformed into the text by the VA, it will reveal what the user genuinely wants by making the request to VA. The project will improve our knowledge, particularly in natural language processing (NLP).

- **What algorithms/approaches will you be testing/developing?**
  Using a labelled dataset of input and the matching intent labels, we will first implement our baseline model i.e. Multinomial Naive Bayes. Afterwards, we will implement different variants of BERT Model and fine-tune the BERT model. Also, we will try exploring several different BERT models, such as RoBERTa, DistilBERT etc. and evaluate them using the 4 different datasets.

- **What will be the final deliverable and by whom could it be used, or what would be its contribution to the field?**
  A BERT-based intent detection project's end deliverable would be a trained model that can categorize user input into corresponding intent categories, a documented project report of the work completed, and the source code in a Jupyter Notebook (IPYNB). Businesses, developers, and researchers can employ BERT-based intent detection to enhance customer service, chatbots, and natural language processing applications.

- **What are the project boundaries that you are setting to be able to achieve your project within 50 hours (2 number of people in the group)? In other words, what is included/excluded to make your achievement realistic?**
  Our project dataset has been expanded to include four datasets, and further analyzing it and preparing the model for various intents. We will evaluate four distinct datasets using several BERT variations, including RoBERTa, DistilBERT, etc., fine-tune the BERT model, and contrast the various BERT Model performances using various metrics like accuracy/Loss.

3. Software Platform, Programming, and Dataset Description:

- **Software Platform**
  GoogleColab, macOS, Github, Kaggle
- **Programming Environment**
  Python, Huggingface Transformers, TensorFlow and Keras
- **Datasets**
  1. The first dataset, which we termed as **Snips NLU Corpus** is hosted on GitHub and is first presented in the paper [3] which comprised of 7 intents:
     a. SearchCreativeWork (e.g. Find me the I, Robot television show)
     b. GetWeather (e.g. Is it windy in Boston, MA right now?)
     c. BookRestaurant (e.g. I want to book a highly-rated restaurant for me and my boyfriend tomorrow night)
     d. PlayMusic (e.g. Play the last track from Beyoncé off Spotify)
     e. AddToPlaylist (e.g. Add Diamonds to my roadtrip playlist)
     f. RateBook (e.g. Give 6 stars to Of Mice and Men)
     g. SearchScreeningEvent (e.g. Check the showtimes for Wonder Woman in Paris)
  2. The other three Dataset is taken from the paper [4].
     a. AskUbuntu Corpus
     b. Web Applications Corpus
     c. Chatbot Corpus

4. Activity Table:

Activity	Why	Time Planned	Deliverable
Research and study current articles and achievements, and building background knowledge	Gather knowledge about intent detection and understanding the problem statement thoroughly.	4h	Gathered good online resources and free video tutorial
Finalizing the 4 datasets and NLP models to work upon	Comparing different datasets and evaluating which dataset is the most suitable for the project. To work with the best algorithms which would further create best and suitable trained models for intent detection.	3h	List of dataset and NLP model to work upon
Configuring and preparing the environment	Installing huggingface transformer library for the Bert model and installing other necessary libraries.	2h	Environment for the Model
Exploratory Data Analysis	Exploring the 4 datasets for the complete understanding using visualization	4h	Synopsis of the dataset and it's labels and distribution.
Preprocessing of 4 datasets	For models to work, we need a preprocessed and cleaned dataset, and analysis of datasets gives good insights.	5h	Cleaned and preprocessed datasets

Developing a Baseline Model and training the dataset on the Baseline Model	The baseline model such as Multinomial Naive Bayes will be trained.	8h	The baseline model will be trained on four different datasets.
Fine-Tunning the BERT Model and implementing Advance model and training to improve result.	First fine-tuning the models using BERT and applying advance algorithm (Variants of BERT) such as RoBERTa, DistilBERT.	11h	A fine-tuned model trained on four different datasets.
Comparative Analysis	Compare different performance measure for all the trained models and select the most efficient model.	3h	The model having best performance measure (Loss/accuracy) and outperforming other models.
Extracting, visualizing, and understanding the result	To get the full understanding of the models, comparing results using graphical representation.	7h	Results and graphs
Preparing PPT and Writing Report	Documenting full project, and summarizing the whole project	3h	Final Project Report

References:

[1] https://github.com/sonos/nlu-benchmark/tree/master/2017-06-custom-intent-engines

[2] https://github.com/sebischair/NLU-Evaluation-Corpora

[3] Coucke, A., Saade, A., Ball, A., Bluche, T., Caulier, A., Leroy, D., Doumouro, C., Gisselbrecht, T., Caltagirone, F., Lavril, T., Primet, M., & Dureau, J. (2018). Snips Voice Platform: An embedded Spoken Language Understanding system for private-by-design voice interfaces. *ArXiv*. https://doi.org/10.48550/arXiv.1805.10190

[4]Daniel Braun, Adrian Hernandez Mendez, Florian Matthes, and Manfred Langen. 2017. Evaluating Natural Language Understanding Services for Conversational Question Answering Systems. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pages 174–185, Saarbrücken, Germany. Association for Computational Linguistics.

[5] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pretraining of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

[6] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, et al.. 2020. Transformers: State-of-the-Art Natural Language Processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online. Association for Computational Linguistics.