**Name Entity Recognition using BERT**

**PROBLEM STAEMENT**

In this project we will be performing one of the most famous task in the field of natural language processing i,e Name Entity Recognition.

**DESCRIPTION OVERVIEW**

Named Entities Recognition (NER) is a basic task of Natural Language Processing (NLP). The purpose is to identify named entities such as person names, place names, and organization names in the corpus. Due to the increasing number of these named entities, it is usually impossible to exhaustively list them in the dictionary, and their constituent methods have some regularities. Therefore, the recognition of these words is usually included in the task of morphological processing (such as Chinese segmentation). Independent processing, called named entity recognition.

Named entity recognition technology is an indispensable part of many natural language processing technologies such as information extraction, information retrieval, machine translation, and question answering systems.

Named entities are the research subjects for named entity recognition. Generally, named entities include 3 categories (entity, time, and number) and 7 categories (person, place, institution, time, date, currency, and percentage). Judging whether a named entity is correctly identified includes two aspects: whether the boundary of the entity is correct; and whether the type of the entity is correctly labeled.

The main types of errors include correct text, which may be of the wrong type; conversely, text boundaries are incorrect, and the main entity words and part-of-speech tokens it contains may be correct.

**TECHNOLOGY USE**

Here we will be using **Anaconda Python 3.6 , Pytorch 1.4 with GPU support CUDA 10 with CuDNN 10.**

**INSTALLATION**

Installation of this project is pretty easy. Please do follow the following steps to create a virtual environment and then install the necessary packages in the following environment.

**In Pycharm it’s easy**

1. Create a new project.

2. Navigate to the directory of the project

3. Select the option to create a new new virtual environment using conda with python3.6

4. Finally create the project using used resources.

5. After the project has been created, install the necessary packages from requirements.txt file using the command pip install -r requirements.txt

**In Conda also it’s easy**

1. Create a new virtual environment using the command

conda create -n your\_env\_name python=3.6

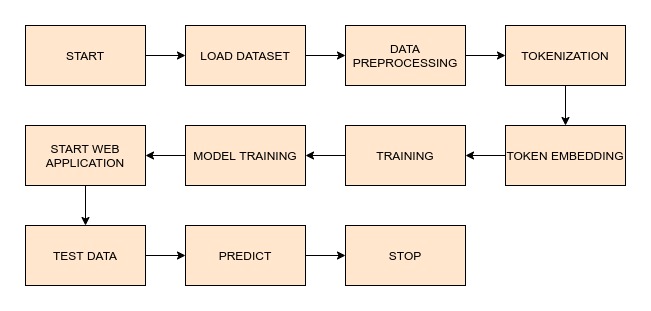
2. Navigate to the project directory.

3. Install the necessary packages from requirements.txt file using the command

pip install -r requirements.txt

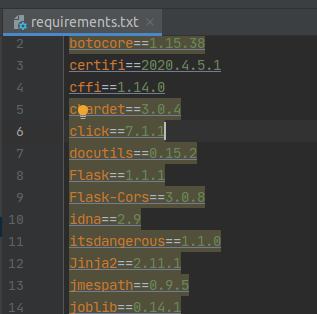
1. Download the model bin file , from <https://huggingface.co/dslim/bert-base-NER/blob/main/pytorch_model.bin> and add it to out\_base folder in flask-backend
2. bert-base-NER is a fine-tuned BERT model that is ready to use for Named Entity Recognition and achieves state-of-the-art performance for the NER task. It has been trained to recognize four types of entities: location (LOC), organizations (ORG), person (PER) and Miscellaneous (MISC).

**WORKFLOW DIAGRAM**

****

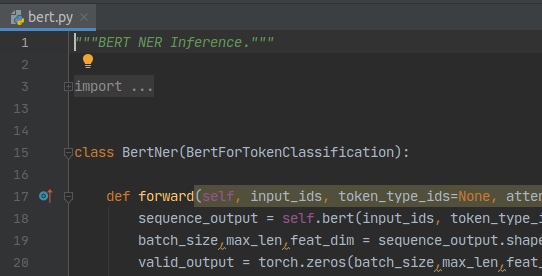
**IMPLEMENTATION**

**1. requirements.txt**

****

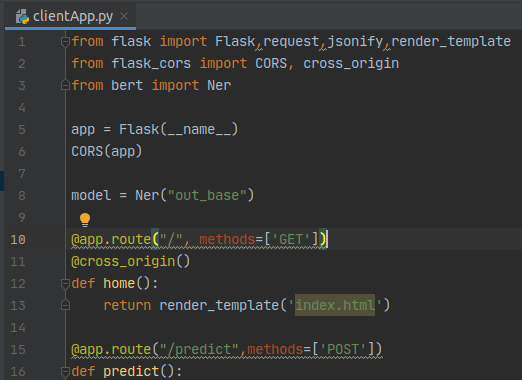
requirements.txt file consists of all the packages that we need to run with project.

**2. bert.py**

****

This is the most important file in this project which we will be using for training and prediction.

**3. App.py**

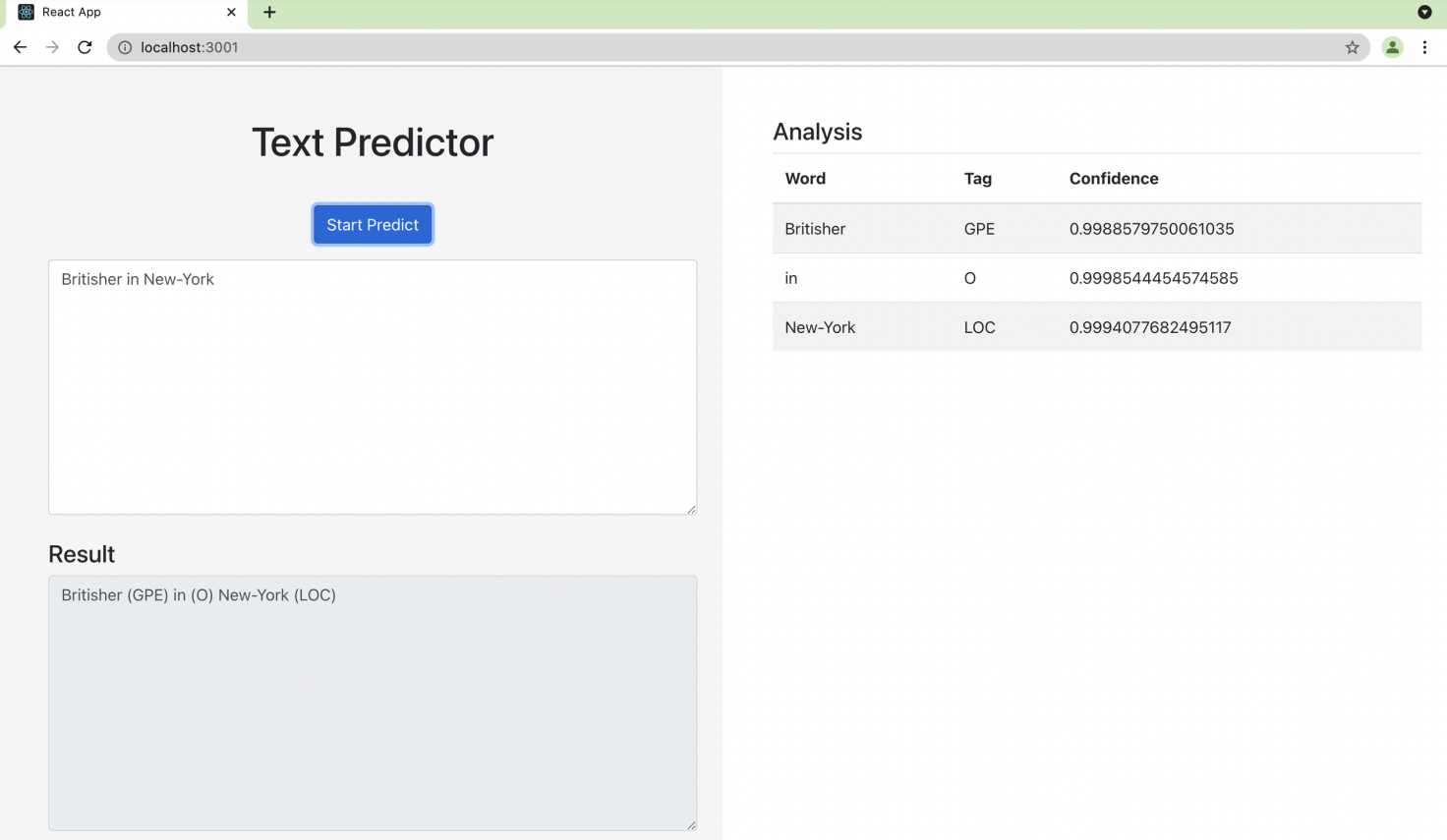
****

This is the flask server file and the entry point of application.

**RUN**

1. yarn install while pwd is react-frontend
2. yarn start
3. python App.py while pwd is flask-backend

**CONCLUSION**

Here we successfully performed Named Entity Recognition on the given dataset.

**COMPARISION**

More data or better larger dataset can be used to build a better model. We can also try out better pre trained model with fine tuning to increase the performance.