READ ME

共有四份code：

1. data-cleaning.ipynb

此code可以對原始data做Word Contractions Expanding(" you're " is expanded in to two words: you are)，並且做一部分的data preprocessing。需要先安裝pycontractions 2.0.0 (<https://pypi.org/project/pycontractions/>)，最後會對training data和testing data分別產生對應的clean版本的CSV檔案，在後面的code中需要用到。

DATA\_SET\_DIR = './project2\_data/' 為training data的路徑，需要更改。

1. NLP-Final-bert.ipynb

此code為利用bert Embeddings的訓練模型程式，需要在google colab執行，File Path皆為Google Drive的Path，需要掛載Google Drive

1. NLP-Final-Glov.ipynb

此code為利用Glov Embeddings的訓練模型程式，需要在google colab執行，File Path皆為Google Drive的Path，需要掛載Google Drive

1. NLP-Final-Plot.ipynb

此code為繪圖程式，繪製多個最佳模型的confusion matrix並計算testing set的F1-score和accuracy，需要在google colab執行，File Path皆為Google Drive的Path，需要掛載Google Drive