Evaluating and Developing Methods of Generating Code-Switched Data
07-400, Spring 2022

Abhishek Vijayakumar,
under Prof. Alan W. Black, Language Technologies Institute
https://inkyubeytor.github.io/category/code-switching.html

April 18, 2022

1 Major Changes

The primary goal of this project is to evaluate the usefulness of the GLUECoS benchmark as an evaluation for the quality of generated code-switched data.

2 What You Have Accomplished Since Your Last Meeting

I tested different loss functions (smoothed F1 and precision vs NLL), hyperparameters (epochs and batch size), and datasets (English IMDB vs code-switched) on the sentiment analysis task. Through this tuning, I developed a version of the sentiment analysis task that produces measurable differences in performance between different models.

3 Meeting Your Milestone

I achieved my goal of developing a measure that can differentiate between the performance of different models by making an easier sentiment analysis task. This will be the focus of my final evaluations and writeup for my project, as this is the only task in the GLUECoS benchmark for which such statistically significant deviations have been achieved.

4 Surprises

A particularly surprising result was the sheer difference in performance between the bert-base-multilingual-cased and xlm-roberta-base pretrained multilingual models on this modified sentiment analysis task. It is possible that the task data happens to skew in the direction of the mBERT model, and this is something I should explore more thoroughly.

5 Looking Ahead

I plan to further explore this new version of the sentiment analysis task and attempt to use it as a differentiator of performance for code-switched pretraining schemes.

6 Revisions to Your Future Milestones

The primary focus of my project until the Meeting of the Minds will be on evaluating this adapted task and writing up my findings over the semester.

7 Resources Needed

No further resources are needed for this project at this time.