Group Proposal

By Jialiang Wu(G20217024) and Zhongyang Wu(G23209468)

We are going to select the Kaggle competition called "Jigsaw Unintended Bias in Toxicity Classification" as our final project. The purpose of this competition is to build a model to detect toxicity across a diverse range of conversations, which recognizes toxicity and minimizes this type of unintended bias with respect to mentions of identities. We select it not only because it is an interesting topic, but also it is an NLP problem, which is suitable for DNN models. Also, we believe now a days, we have witness more and more cyberbullying, by participating this competition, we would like to change this situation.

Data set:

The data set is provided by the Kaggle platform, which has 1804874 lines of training data with 43 features and 97320 lines of test data. The size of the data set is more than enough for us to train a deep network.

Models:

The deep network models we are going to use maybe some of CNN, LSTM and GRU, and some pretrained models like BERT. As we are going to take part in the competition, it should be customized to get a better performance.

Framework:

The framework we are going to use to implement the network is Keras, which allows for easy and fast prototyping (through user friendliness, modularity, and extensibility). And it also supports both convolutional networks and recurrent networks, as well as combinations of the two. What's more, it runs seamlessly on CPU and GPU[1].

Materials:

The reference materials we are going to use to obtain sufficient background on applying the chosen network to the problem are public kernels which are shared by competitors for the this and other competitions. And we also will get inspired from the discussions in this platform.

Judgement:

[1] Keras.io

We will judge the performance of the network through the Kaggle competition platform by submitting our prediction of the test set. And the platform will give us a score to reflect the performance of out model. We will use the AUC as our metric.

Schedule:

08/04 – 14/04: Learning the basic knowledge of the basic NLP and the framework keras.

15/04 – 22/04: Building the baseline of the model and improving it.

23/04 – 29/04: Writing the final report, and making preparation for the Group Presentation.