Display name : ximing

Name ; Ximing Wen

For this assignment, I used three features besides the edited word in the baseline model. There are the incongruity of the edited word and the original word in each joke, the position of the edited words in each joke and the length of each joke.

To obtain the distance between the edited word and the original word, I use the trained google word2vec model to get the vector for each word. The assumption is the longer the distance is, the funnier the edited joke will be. And then I use cosine distance to describe the difference between those word.

I have tried standardizing these features because they have different scales. But the result showed better when I don’t standardize them.

For the model itself, I tried ridge regression model, linear regression model and Lasso regression model. I found ridge regression model perform the best. I tried several values for the parameter alpha, and found 4 is the best choice.

I obtain more data and use those data to train the model and this also result in the improvement.

Thanks to the essay *“President Vows to Cut Taxes Hair”: Dataset and Analysis of Creative Text Editing for Humorous Headlines，*the information in this essay helped me a lot.