CS 579: Fake News Classification

Sam Golden, Theo Guidroz

Introduction

Given two news titles, the goal of this project is to classify whether the titles agree, disagree, or are unrelated. This task is often used to perform fake news classification.

Dataset

The dataset provided consists of 256,442 pairs of titles and their assigned label. This large dataset has an imbalance in labels which could affect the performance of any models trained on it.

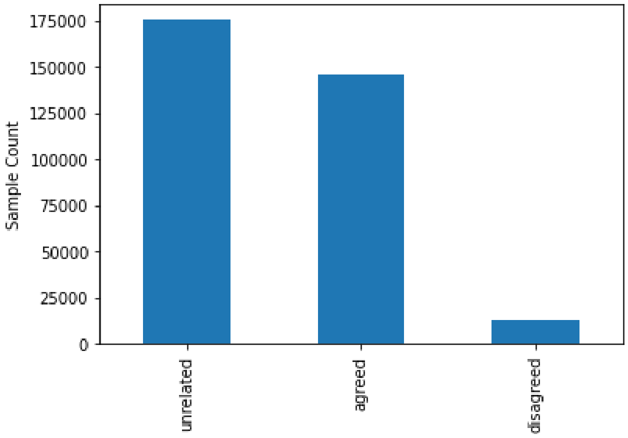

Fig 1: Barchart of label distribution

Data Augmentation

To overcome the problem of the unbalanced dataset, each sentence from the dataset was translated to French, then back to English using Google Translate. If the double-translated sentence differs from the original one, it would be added to the dataset. It is worth mentioning that only sentences labeled as agreed or disagreed were augmented.

Pick one sentence	Translate to French	Translate back to English
Is beer belly because of beer?	Le ventre de bière est-il causé par la bière?	Is beer belly caused by beer?

This data augmentation technique doubled the amount of sample count of agreed and disagreed labels.

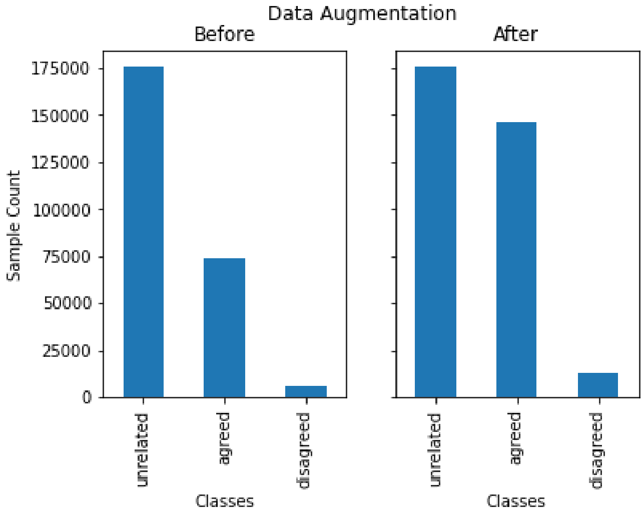

Fig 2: Results of Data Augmentation

Data Preprocessing

Every sentence from the dataset was processed before training. Here are the steps taken during the preprocessing stage:

1. All the letters were converted to lowercase
2. Removed all links in the sentences
3. Removed any hashtags but kept the actual word
4. Removed non-alphanumeric characters
5. Lemmatized each word

Lemmatizing plays an important role here since it groups the inflected forms of a word so they can be analyzed as a single item. For example, after lemmatizing, the words ‘balls’ and ‘ball’ are seen as the same word.

Before:

Hello, my name is #Theo and I am a student taking CS579 www.iit.edu. and I like training models.

After:

hello my name is theo and i am a student taking cs579 and i like training model

Model Architecture

We trained a variety of architectures but settled on a model with the following structure: two symmetric convolutional LSTM models, one for each sentence, which are concatenated and fed to a small dense network for the final prediction. See Figure 3 for an illustration.

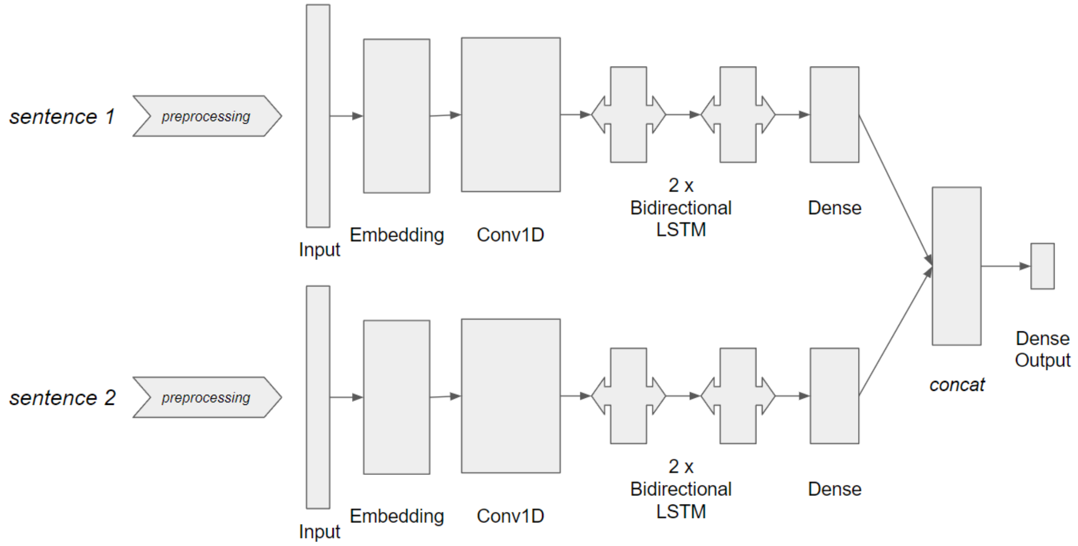

Fig 3: Illustration of the model architecture

The input layer is a 1-dimensional tensor that has only the result of tokenizing the pre-processed input sequence. The embedding layer makes that 1D tensor dense, and feeds it to the convolutional layer. The bidirectional LSTM layers are the last step in the symmetric model; they have the lowest unit count. Finally, the LSTMs feed to a dense layer and the dense layers from each symmetric model are concatenated to a smaller dense network. The final layer in the network is a 3-unit dense layer with softmax activation to yield the class decision. Some variations on that above model include

- Deeper LSTM layers and no CNN
- A pre-trained Embedding layer with Stanford's GloVe 100d weights
- A deeper dense network in the post-concatenation section of the model

Please see the associated accuracy of each variation in the next section.

Results

The table below shows the results of the different model architectures trained.

Model Architecture	Validation Accuracy (%)
Bidirectional LSTM	69.23

Bidirectional LSTM + pre-trained Embedding Layer	69.74
Bidirectional LSTM + post-concatenation fully connected section	71.16
Bidirectional LSTM + CNN	74.14

Here is some prediction example from the best model:

Unrelated:

- *The great coat brother Zhu Zhu Wen, in the mandarin love song to sing the song is really the lanca- talent is very sweet!*
- *Lin xinsheng after the birth of "hard milking," Huo jianhua is not seen, "forced marriage" is real?*

Agreed:

- *Game of Thrones has finally confirmed that season 8 will return in 2019*
- *HBO "Game of Thrones," the final season of season 8 will be back in 2019!*

Disagreed:

- *liu mingwei has lost his examination certificate*
- *Liu Mingwei's test card has been lost? It's a rumor. Don't be fooled again!*

Conclusion

In this project, we successfully trained a model to classify 2 given titles as agreed, disagreed, or unrelated with an accuracy of 74.14%.

References

Keras Documentation and Tutorials

https://keras.io/

https://keras.io/examples/nlp/text_extraction_with_bert/

A Detailed Explanation of Keras Embedding Layer

https://www.kaggle.com/code/rajmehra03/a-detailed-explanation-of-keras-embedding-layer/notebook