Trust or Suspect? An Empirical Ensemble Framework for Fake News Classification

Shuaipeng Liu, Shuo Liu, Lei Ren

Meituan-Dianping Group

{liushuaipeng,liushuo19,renlei04}@meituan.com

Challenge

In this paper, we proposed an ensemble framework to address the fake news classification challenge in ACM WSDM Cup 2019 . In our solution, we regarded this problem as the Natural Language Inference (NLI) task and proposed a novel empirical ensemble framework and finally our team Travel won the 2nd place with a weighted accuracy score of 0.88156 on the private leaderboard.

Our Solution

An overall framework and processing pipeline of our solution is showed in Figure 3.

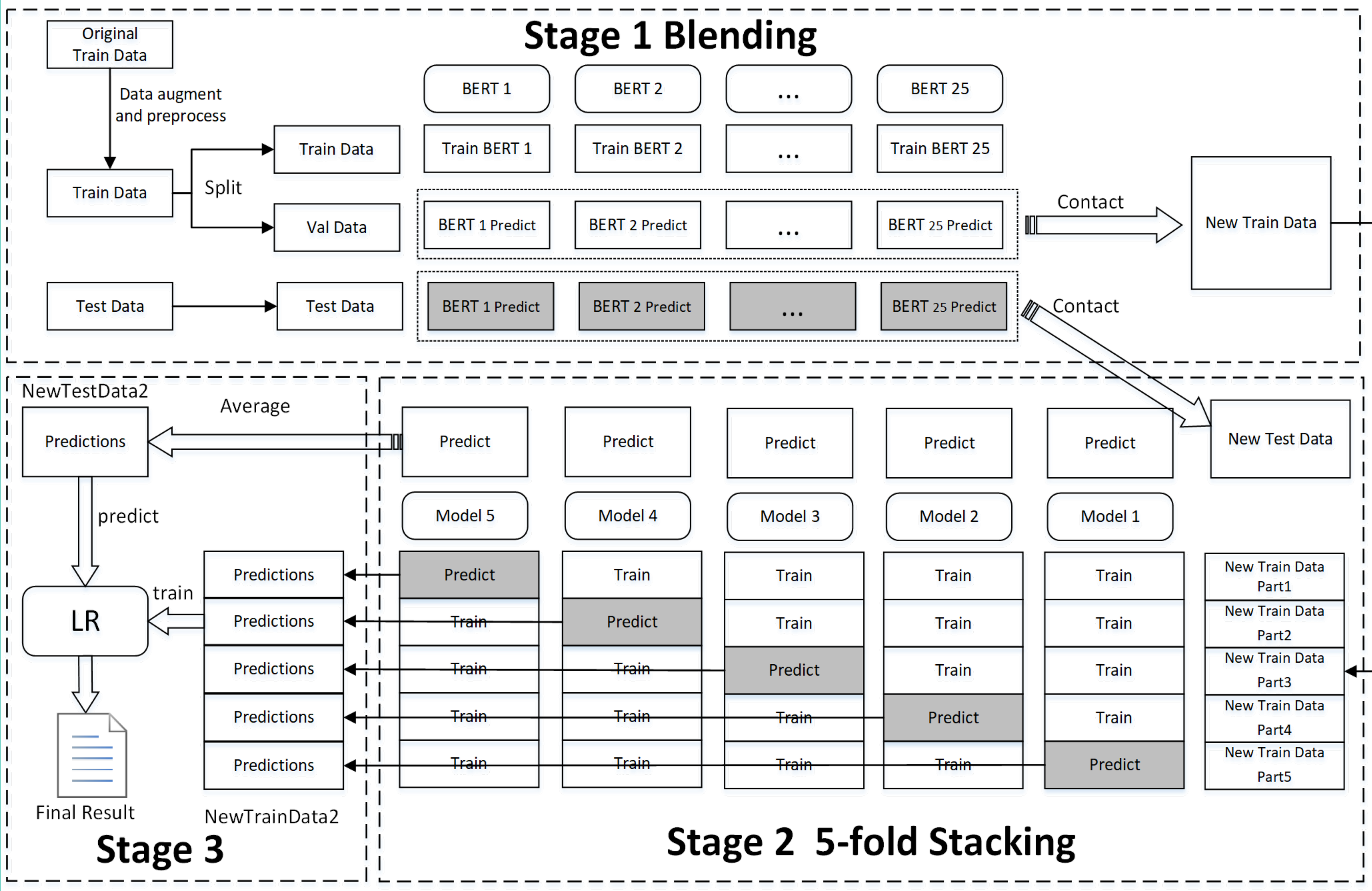

Figure 1: An overall framework and pipeline of our solution for fake news classification

1) Data Augment

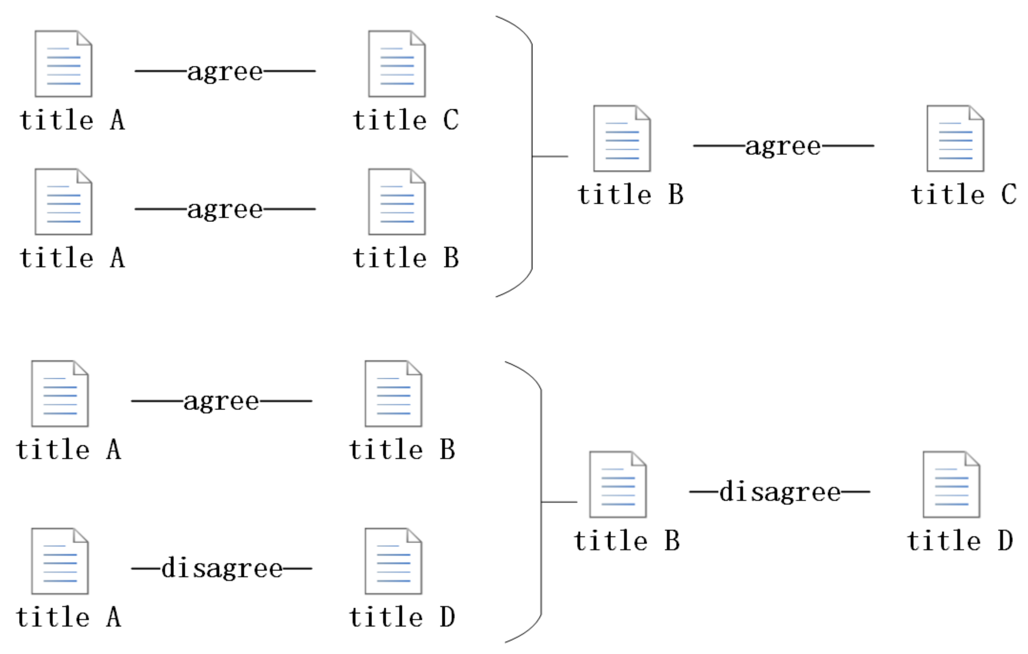

Figure 2: Data augment strategy for generating news pairs

We proposed a simple yet effective method for data augment, by using the semantic transitivity.

2) Data Preprocess

In the data preprocessing part, we first made a transformation between Traditional Chinese and Simple Chinese for the news titles, and then removed the stop words in the dataset.

3) Base Model

We used the BERT as the base model.

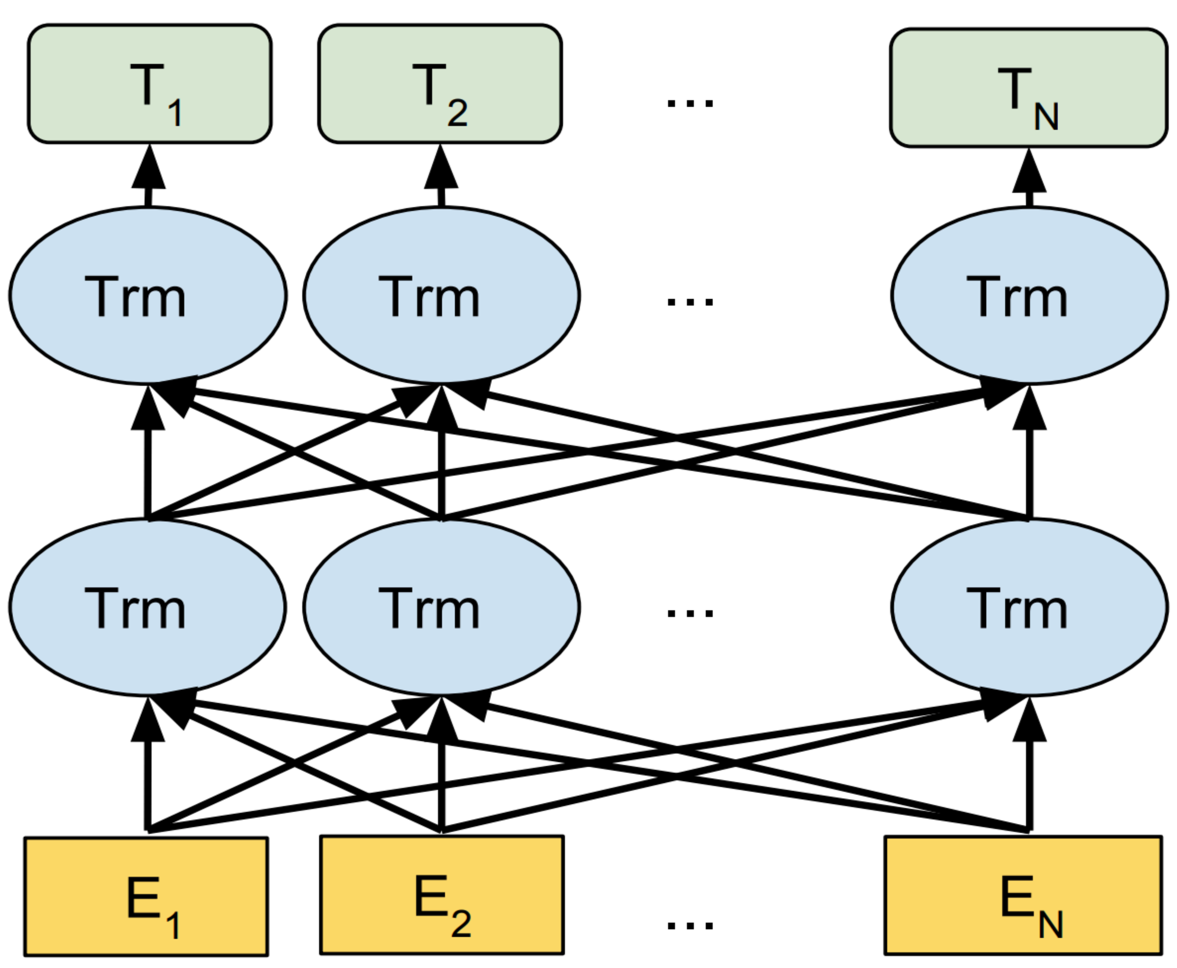

Figure 3: Architecture of BERT for pre-training with a bidirectional Transformer

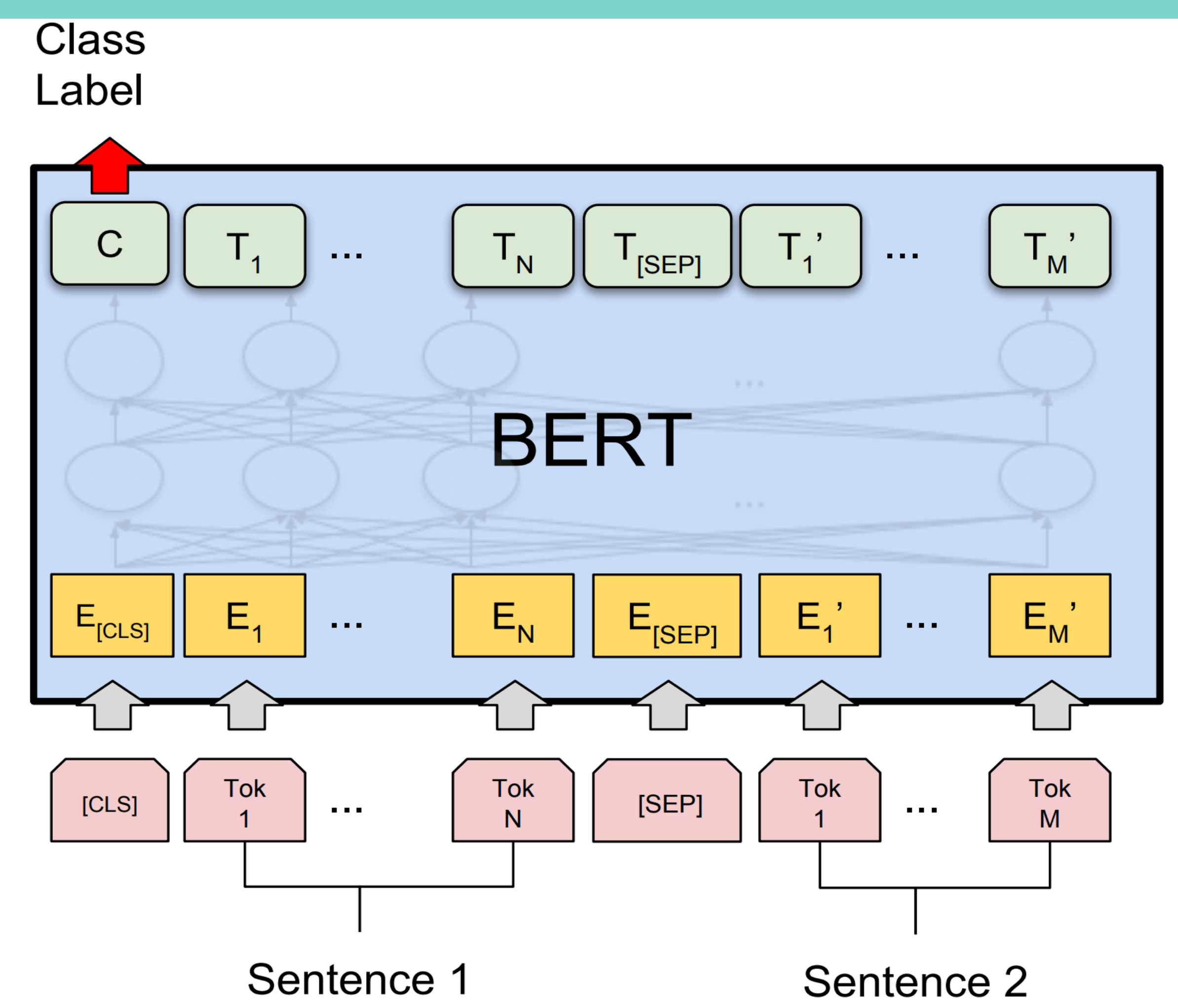

Figure 4: Fake news classification model by incorporating BERT with one additional output layer

4) Model Ensemble

As shown in Figure 7, we build a three-level architecture as the ensemble model to perform the fake news classification.

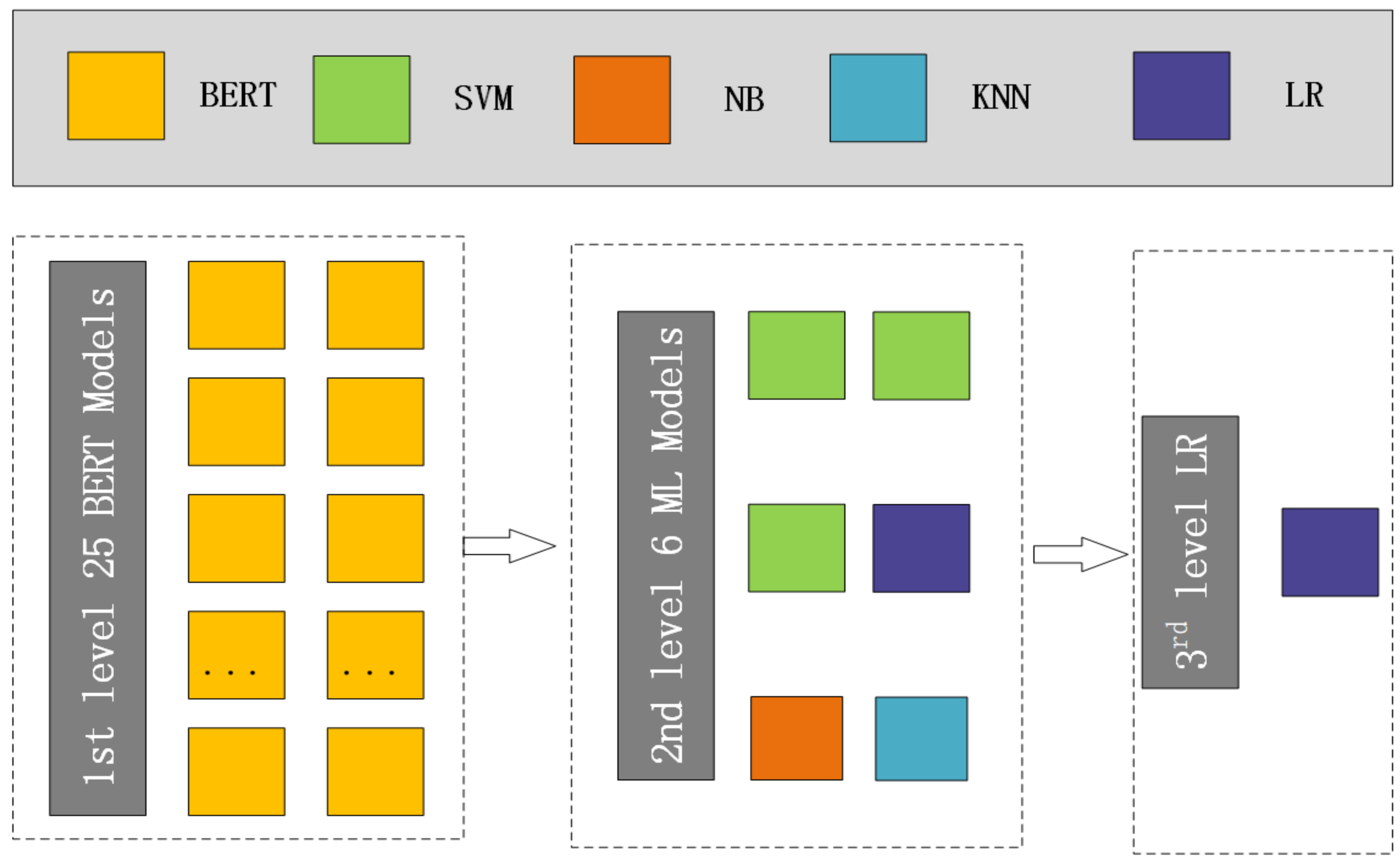

Figure 5: Model Ensemble Architecture

Experiment and Result

1) Data Set

The training dataset consists of 320,767 news pairs with 3 class labels (agreed, disagreed and unrelated), and the testing dataset contains 80,126 news pairs without labels.

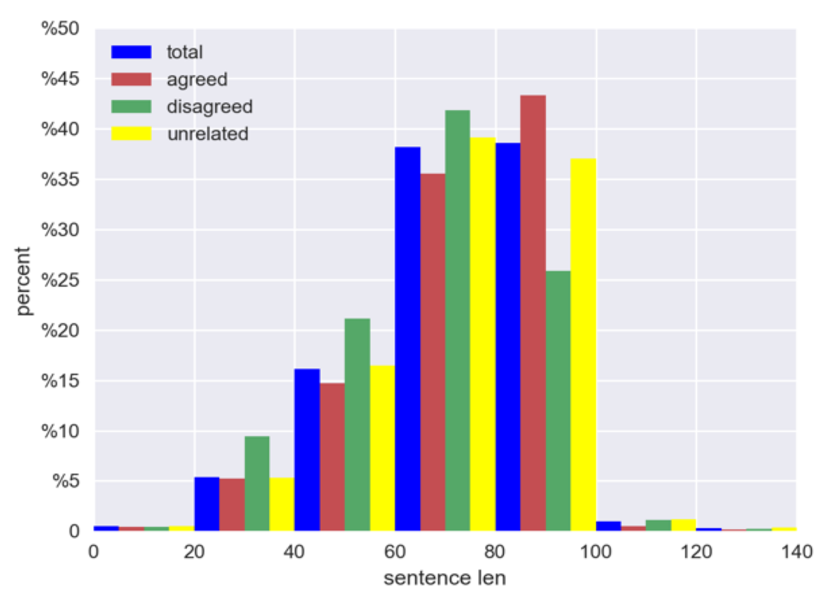

Figure 6: Distribution of news title length in Chinese

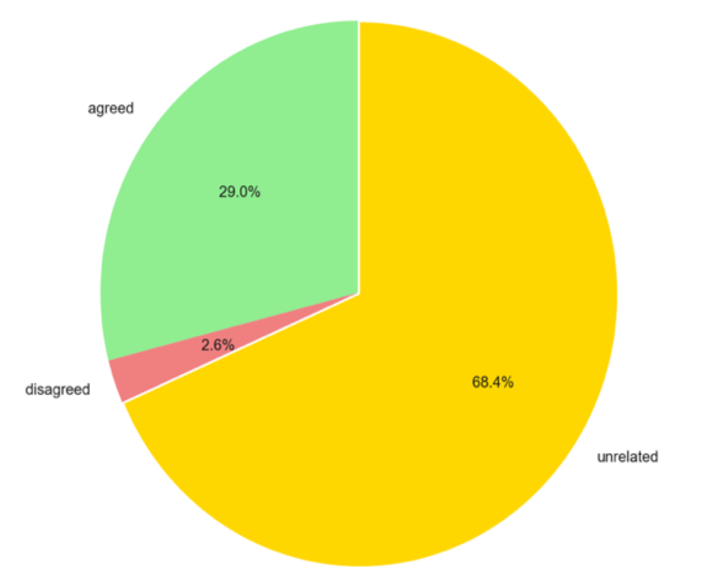

Figure 7: Distribution of class labels in the training dataset

2) Metrics

$$\textbf{WeightedAccuracy}(y, \hat{y}, \omega) = \frac{1}{n}\sum_{i=1}^{n}\frac{\omega_i(y_i = \hat{y}_i)}{\sum \omega_i}$$

Where y are the ground truth, $\hat{y}$ are the predicted results, and ω_i is the weight associated with the i-th item in the dataset. The weights of the three categories, agreed, disagreed and unrelated are 1/15, 1/5, 1/16 respectively.

3) Results

Table 1: performance of various models

Model	Weighted Acc on Private LB
Best Single base model	0.86750
Averaging of 25 BERT	0.87700
Weighted Averaging of 25 BERT	0.87702
Our Empirical Ensemble Model	0.88156