

Fake News Detection Competition

Technical Approach

Since the BERT model is trained on a large amount of data with the specific goal of understanding the context, Hence I used a pre-trained BERT model from Hugging Face for this task. (<https://huggingface.co/bert-base-cased>)

Experimental Settings

I performed a few experiments with changes in the training dataset and encoder setting.

Experiment 1:

Dataset: Use only the statement text feature as input

Encoder: padding = true, truncation = true, padding = {left, right}

Experiment 2:

Dataset: Use all the features as input. I concatenated all the features with the '<col_name>' token before the feature. For example my input sentence would look something like below

'<context> a floor speech. When did the decline of coal start? It started when natural gas took off that started to begin in (President George W.) Bushs administration. <subject> energy,history,job-accomplishments <job_title> State delegate <party> democrat'

Experiment 3:

I also tried a few other variants of the bert model like '*microsoft/deberta-v3-base*' but the base model performed the best.

The experiment 1 (with padding = right) performs the best on the final test set

Implementation details

Dataset: I preprocessed the dataset to convert labels from categorical to numeric. All the missing values were replace with empty string ""

Training: As mentioned earlier, I used the pre-trained BERT model from Hugging Face (<https://huggingface.co/bert-base-cased>) with 6 labels. I used a batch size of 8 for training. Other parameters were kept to default transformer training arguments provided by hugging face. The model was trained for 10 epochs on Nvidia GTX 1080i gpu. It took ~2 mins for one epoch

Inference: Inference is done using the saved checkpoints from training with a batch size of 24.