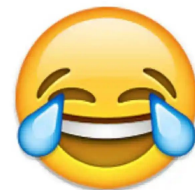


Predicting Humor in Headlines



Christian Kavouras, James Blackburn, Joseph Nollette, Joshua Valdez



Task Description

- Implement a regression based model to predict humor in a dataset of headlines
- Dataset consists of headlines taken from the Humicroedit dataset
- https://competitions.codalab.org/competitions/20970#learn_the_details-overview



Task Description

- Single word edits to headlines for comedic effect
- Labeled scores on a 0-3 scale of “funniness”
- Averaged across 5 human scores



Approach

- We use a single regression pre-trained BERT model
- BERT then fine-tuned on our headline data
- Evaluated as a single-regression task with RMSE loss
- Note: Due to access to external GPUs, we are able to fine-tune BERT itself



Approach

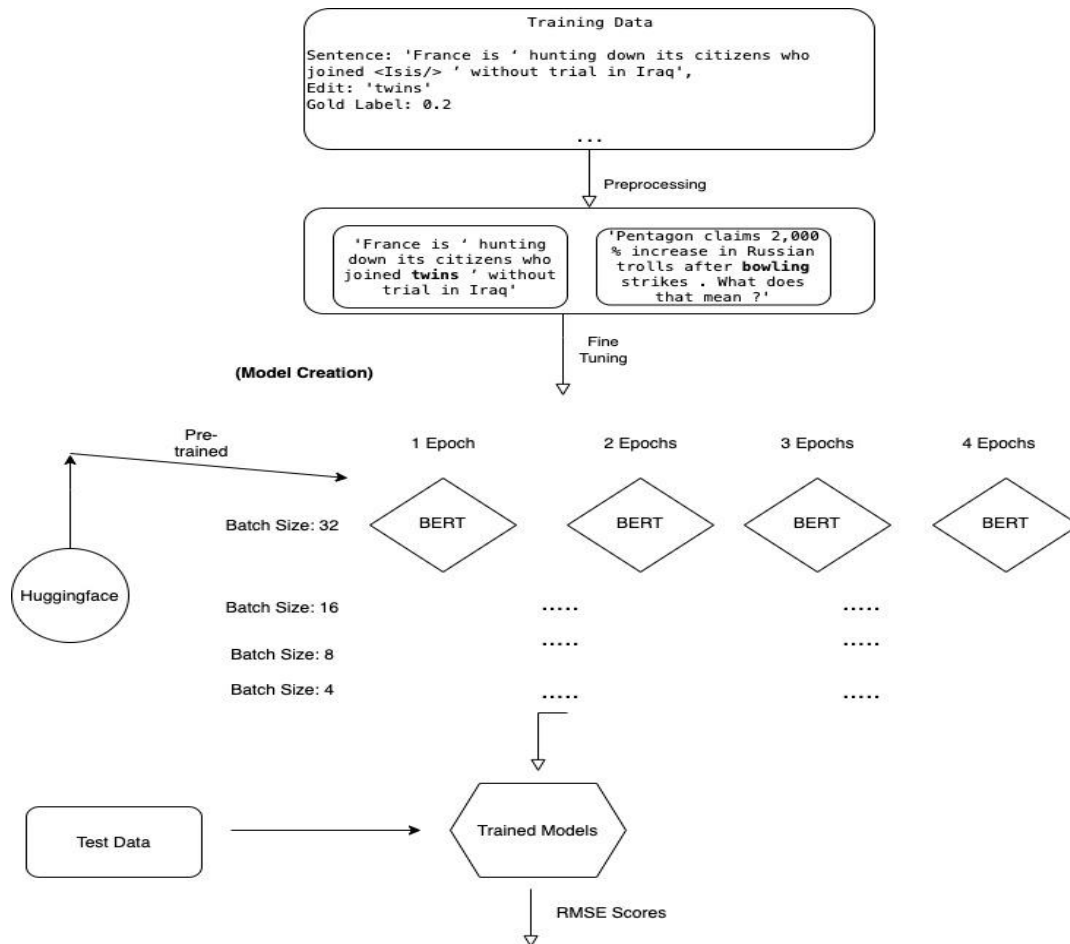
Features

- Inputs to the model are raw sentence embeddings
- Tokenized using vanilla BertTokenizer

Hyperparameters

- Tuned models across varying epochs:
 - 1, 2, 3, 4
- Tuned with various batch-sizes:
 - 32, 16, 8, 4
- Adam Optimizer

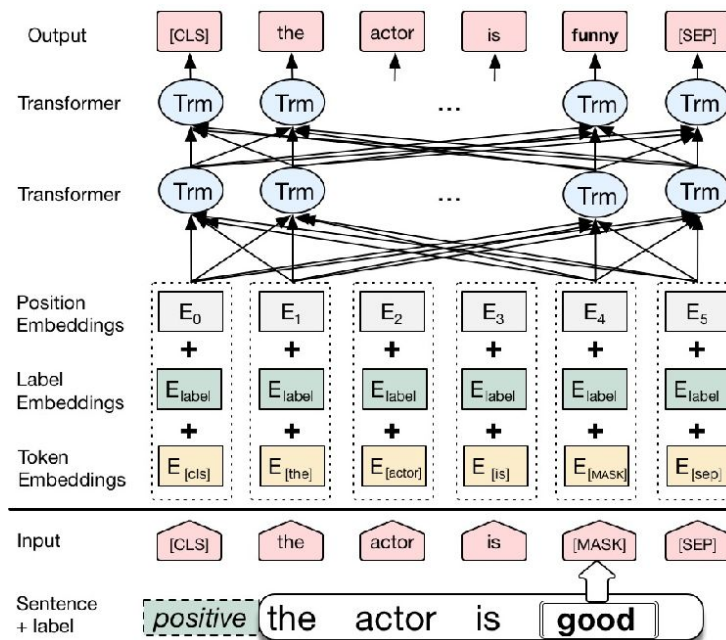
System Design



System Design

(Schematic of BERT architecture:)

In our case, labels are decimal values in [0, 3]





Results

results for subtask1:

baseline: 0.57471

BERT single regressor:

task A RMSE results:

epochs	batch=32	batch=16	batch=8	batch=4
1epochs	0.54199	0.54312	0.53396	0.53438
2epochs	0.53324	0.54381	0.5383	0.54047
3epochs	0.54728	0.54737	0.56014	0.55393
4epochs	0.55964	0.55747	0.56742	0.56569



Discussion

- For our primary task, our BERT model performs impressively well
 - (Butts up against the baseline)
- Can expand from a single regression
- Can further fine tune a downstream, feature based model on our input



Related Reading

BERT paper: (Devlin et al., 2018) <https://arxiv.org/pdf/1810.04805.pdf>

Aspect-based sentiment analysis - ideas from a similar task domain (Xu, et al., 2019)
<https://arxiv.org/pdf/1904.02232.pdf>

Winning Model for same task (Spanish Dataset) http://ceur-ws.org/Vol-2421/HAHA_paper_3.pdf