FakeNews Engine

Table of Contents

How to read

Assignee:

Engine Detail
- **Description**:
- **Input/output**: (+)
- **Structure**: diagram
- **Interface/usage**:
- **Performance**:
- **References**:

Release Notes: ,

Assignee

@

Engine Detail

Description

- links: Bitbucket | Docker
- FakeNews Classifier
- .
- BERT Representation classification layer .
- (,) BERT classification .

Input/output data

Input	Output

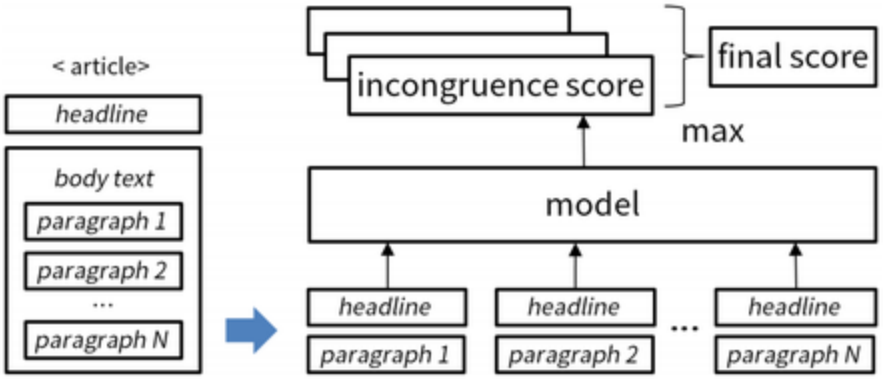
• text
  • "[] [EOS] []"
  • text [EOS] .
  •)
  " \" \" LA [EOS] . (19) . 19 ( ) . " ,
  " ."
	• category
	  • "none" or "clickbait"
	• label
	  • 0 if "none", 1 if "clickbait"
	• probability of being clickbait
	  • calculated by taking the max predicted probability of all (title, sentence) pairs in the article

FakeNews.proto

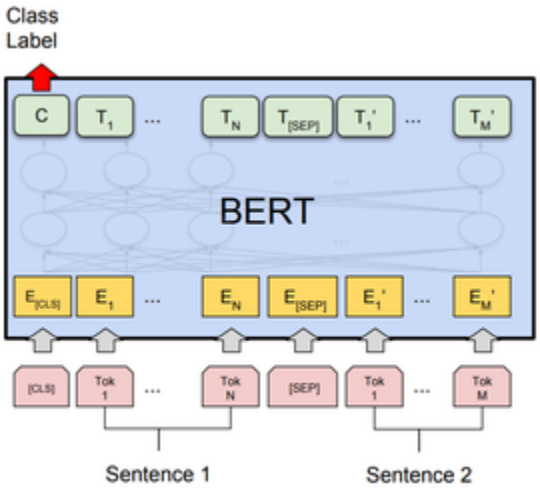
FakeNews.proto

Structure

- Inference structure



- Model structure (BERT with sentence pair input)



Interface/usage

.proto	FakeNews.proto
demo client	FakeNews_client.py
how to run	**docker run -it --gpus '"device=0"' -p { }:8080 docker.maum.ai:443/brain/fakenews:v1.0.0**

Performance

Dataset	Environment	Configuration	Accuracy	AUROC
KAIST AAAI-2019 Validation Data	V100	• train batch: 32 • learning rate: 2e-5 • linear schedule with warmup (warmup = 0) • epochs: 1 (saved checkpoint at every 1000 steps) • train time: 4 days	97.7%	0.997
JIPS 2018 Irrelevant Data	V100	• train batch: 32 • learning rate: 2e-5 • linear schedule with warmup (warmup = 0) • epochs: 1 (saved checkpoint at every 1000 steps) • train time: 4 days	79.0%	0.851

References

- +
- KAIST AAAI-2019 | Github
- 2018 JIPS | Github
- How to fine tune BERT for Text Classification?
- Hierarchical Transformers for Long Document Classification
- KoELECTRA Github
- HanBERT Github

Release Notes

Date	version (links)	What's new?
26 Feb 2021	v1.0.0 (commit | docker)	• Initial release • BERT fine-tuning • 3 v2.0.0 • Enhanced BERT , , • Bitbucket repo