Fake News Detection

Alin-Florin Mihăilă, ISI2
Constantin-Cătălin Crăiță, ISI2

Introduction

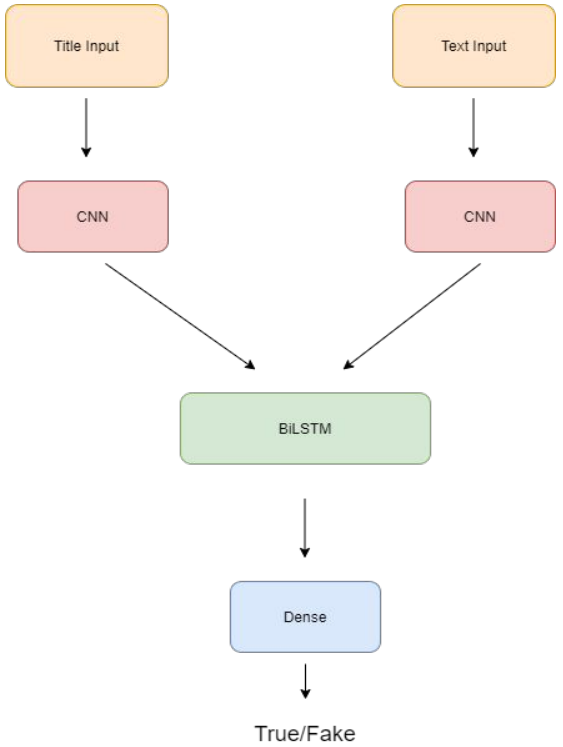
- Worldwide phenomenon
- Abuse of social media (fake news/cyber propaganda)
- Automated classification of articles as misinformation/disinformation
- Challenging task in NLP (Natural Language Processing)

Dataset

- Crawling from various romanian news sites
- Crawled sites : libertatea.ro, evz.ro, puterea.ro, notabn.ro, gandul.ro, news.ro, b1.ro, zf.ro, ziare.ro, stiripesurse.ro, dcnews.ro, adevarul.ro, digi24.ro, agerpres.ro, informateca.ro, mediafax.ro
- Total: 19626 news
- Experiment: news from b1.ro and puterea.ro were labeled as fake news
- ~17.000 true and ~3.000 fake

Method

- SVM
  - baseline
  - TF-IDF
  - NGRAMS - bi & trigrams
- Deep Learning - CNN & LSTM
- Preprocessing
  - Punctuation removal
  - Transliterate to latin
  - Stemming



Model details

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 100, 10)]	0	
input_2 (InputLayer)	[(None, 100, 10)]	0	
time_distributed (TimeDistribut	(None, 100, 10, 30)	810	input_1[0][0]
time_distributed_4 (TimeDistrib	(None, 100, 10, 30)	810	input_2[0][0]
dropout (Dropout)	(None, 100, 10, 30)	0	time_distributed[0][0]
dropout_1 (Dropout)	(None, 100, 10, 30)	0	time_distributed_4[0][0]
time_distributed_1 (TimeDistrib	(None, 100, 10, 32)	2912	dropout[0][0]
time_distributed_5 (TimeDistrib	(None, 100, 10, 32)	2912	dropout_1[0][0]
time_distributed_2 (TimeDistrib	(None, 100, 1, 32)	0	time_distributed_1[0][0]
time_distributed_6 (TimeDistrib	(None, 100, 1, 32)	0	time_distributed_5[0][0]
time_distributed_3 (TimeDistrib	(None, 100, 32)	0	time_distributed_2[0][0]
time_distributed_7 (TimeDistrib	(None, 100, 32)	0	time_distributed_6[0][0]
concatenate (Concatenate)	(None, 100, 64)	0	time_distributed_3[0][0] time_distributed_7[0][0]
spatial_dropout1d (SpatialDropo	(None, 100, 64)	0	concatenate[0][0]
bidirectional (Bidirectional)	(None, 200)	132000	spatial_dropout1d[0][0]
dense (Dense)	(None, 1)	201	bidirectional[0][0]

Total params: 139,645
Trainable params: 139,645
Non-trainable params: 0

Updated Model

- Change char embedding and CNN with only word embedding
- Use the best values found from grid search

Layer (type)	Output Shape	Param #	Connected to
input_70 (InputLayer)	[(None, 50)]	0	
embedding_69 (Embedding)	(None, 50, 50)	724600	input_70[0][0]
embedding_70 (Embedding)	(None, 50, 50)	3204750	input_70[0][0]
concatenate_32 (Concatenate)	(None, 50, 100)	0	embedding_69[0][0] embedding_70[0][0]
bidirectional_36 (Bidirectional	(None, 200)	160800	concatenate_32[0][0]
dense_34 (Dense)	(None, 1)	201	bidirectional_36[0][0]

Total params: 4,090,351
Trainable params: 4,090,351
Non-trainable params: 0

Grid Search Results

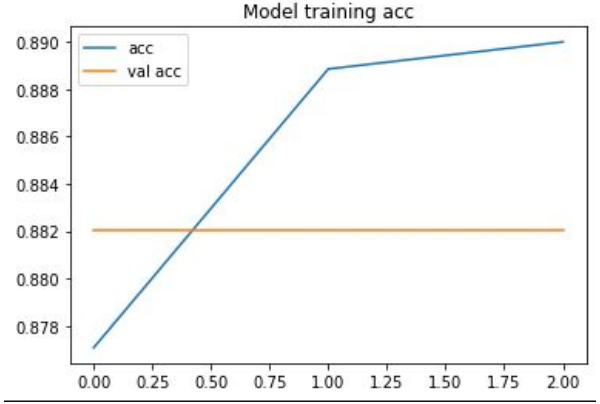
- Grid search on f1-score maximization
- Grid search with following values:
  - Sentence sizes - 50, 75, 100
  - Embedding sizes - 50, 100
  - # LSTM Units - 100, 200
  - Batch sizes - 32, 64
- Best f1 value - **78.8**
- Overfitting with all params

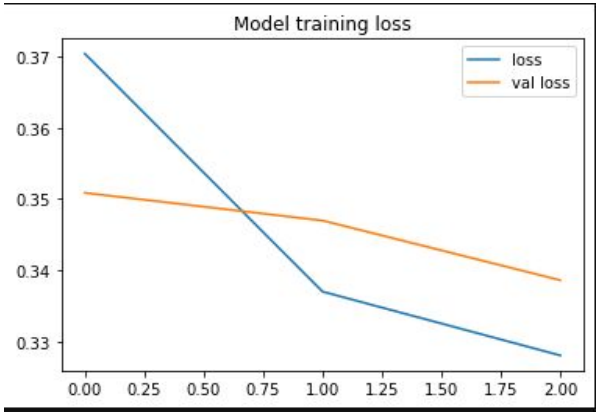
BEST CONFIG

Param	**Val**
Sentence size	50
Word Embedding Size	50
# LSTM Units	100
Batch size	32

Training details

- 100 characters sentence size
- 10 characters word size
- 32 batch size
- 3 epochs
- Adam optimizer





Results

Model	Precision	Recall
Neural Network	99.31%	20.9%
SVM	99.84%	22.85%

Results Comparison

Conclusion

We lost ~15% in precision but gained ~65% in recall

Model	Precision	Recall
Neural Network	99.31%	20.9%
New Neural Network	84.76%	85.73%
SVM	99.84%	22.85%

Technology stack

Android Studio

TensorFlow

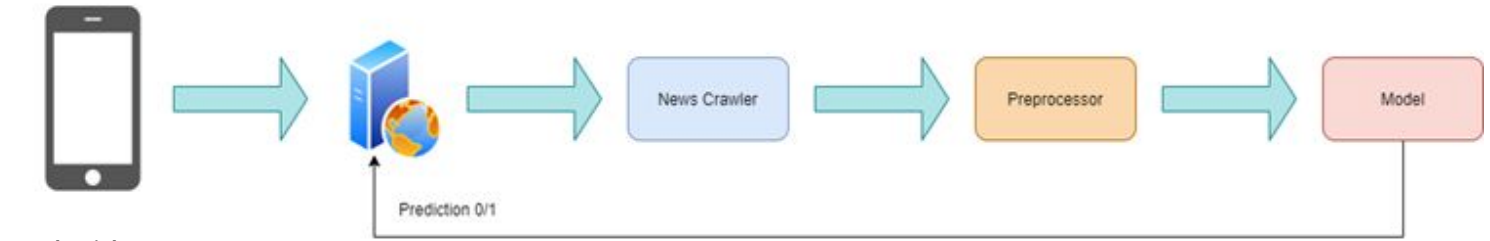
news-please

Beautiful Soup

FastAPI

docker

Application architecture



Android App

Demo time!

Conclusions

- Need manual labelling for proper results
- Big size of text was difficult to fit a NN model
- Imbalanced dataset created problems for training
- Obtained very good precision but bad recall - missing 80% of fake news
- Difficult to asses fake news with content only, external context needed

Questions?