Identifying Political Spectrum in News Articles

Markus Zechner and Halldora Gudmundsdottir

Introduction and Motivation

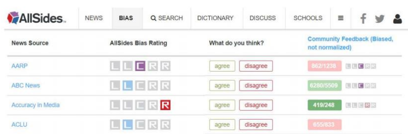
- The main goal of this project is to automatically classify news articles based on their political spectrum.
- The political spectrum ranges from liberal (left) to conservative (right) and the classification is performed on the text body of the articles.
- But what are the direct implications of ranking news article based on the political spectrum?
  - Automatically identifying the political spectrum of an article, the recommendation engine for users could be improved.
  - A variety of articles could be offered to users to provide them with different angles of a story.

Data, Labels and Cleaning

- The dataset is a collection of news articles (deepnews.ai) that originate from different news organizations ranging from very liberal to very conservative.
- The articles are labeled **liberal**, **conservative** or **neutral** using the website *www.allsides.com*. It is important to point out that articles are labeled based on publisher and no manual labeling is involved.
- The train/dev/test set split is 70/15/15.
- The text body of each article is cleaned by (1) removing special characters, (2) converting letters to lower case, (3) splitting each sentence into words.

Publisher	Dataset 1	Dataset 2
New York Times	3,340	7,799
Guardian	3,340	25,914
Washington Post	3,340	11,002
Atlantic	3,340	7,113
Breitbart	3,340	23,375
Fox News	3,340	4,324
National Review	3,340	6,144
New York Post	3,340	17,466
Reuters	3,340	10,710
Quartz	3,340	17,132
Financial Times	3,340	17,134
Business Insider	3,340	6,332
Total	40,080	154,445

Liberal
Conservative
Center



"During Donald Trump's many decades as a famous American tycoon, even his most disgraceful antics never provoked a nationwide resistance movement. Most progressives shrugged, went about their lives, and let the Donald be the Donald. Only when he obtained the power of the presidency did progressives put Princess Leia stickers on their Priuses and rose up to resist."

Example from a conservative article (National Review)

"It was not just showing people who do not understand her and who do not trust her who she is as a person, or laying out her policy proposals, but also demonstrating that when she represents them on the world stage, she would do so with that aura of leadership and power. And she did. In her white suit, with her white crew neck underneath, Mrs. Clinton looked supremely unflappable: perfectly tailored and in control."

Example from a liberal article (New York Times)

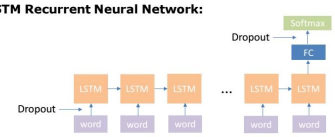
Models

Pre-training: GloVe 50 dimensional word embeddings are used for word representation.

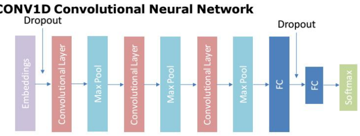
Baseline: For each article word embeddings are averaged to a single vector which is fed into a softmax activation function.

1. **Fully Connected Neural Network:** For each article word embeddings are averaged to a single vector which is fed into a 3 layered network with 500 neurons (relu activation), followed by a softmax output layer.

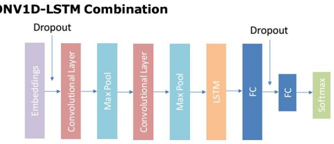
2. **LSTM Recurrent Neural Network:**



3. **CONV1D Convolutional Neural Network**

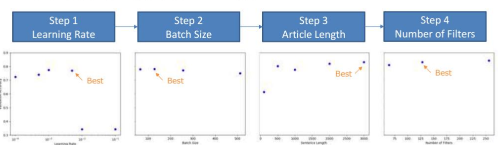


4. **CONV1D-LSTM Combination**



Hyperparameter Tuning

- Because of computational limitations, tuning the models is performed in a sequential manner: (1) learning rate, (2) batch size, (3) article length, (4) number of neurons/filters and (5) regularization/dropout.
- The most sensitive parameters are the learning rate and article length.
- An example of tuning the CONV1D model:

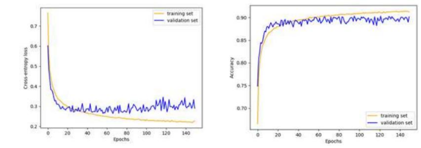


Results

- LSTM generally works well but long training time limited hyperparameter tuning.
- Most models experienced overfitting – dropout was used as a regularization.
- The larger training set (Dataset 2) shows a better performance – less overfitting and higher accuracy – but had very long training times (>30hrs).
- The best performance was achieved by the Conv Net, the LSTM does not work well for long sequences.
- A hybrid of Conv Net and LSTM showed promise in reducing number of parameters but retaining the capabilities of the LSTM.

Model	Training set		Validation set		Test set		Training Time	Dataset	Tuning
	Loss	Accuracy	Loss	Accuracy	Loss	Accuracy			
BASELINE	0.88	58.8%	0.89	58.5%	0.88	58.0%	1.6 min	1	Manually
FC	0.72	68.7%	0.73	67.6%	0.73	67.7%	10 min	1	Sequentially
LSTM	0.13	95.2%	0.61	84.4%	0.63	84.1%	15.8 hrs	1	Manually
CONV1D	0.16	93.8%	0.59	84.3%	0.60	83.8%	1.4 hrs	1	Sequentially
	0.23	**91.3%**	**0.29**	**90.2%**	**0.29**	**90.3%**	**5.5 hrs**	**2**	**Not tuned**
CONV1D-LSTM	0.29	88.6%	0.45	83.0%	0.46	82.1%	5.6 hrs	1	Sequentially
	0.18	93.1%	0.29	89.4%	0.29	89.9%	18.9 hrs	2	Not tuned

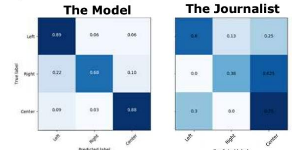
Loss: categorical cross entropy, metrics: accuracy, optimizer: Adam



Error Analysis

- It is harder for our model to quantify Left/Right than Left/Center or Right/Center.
- The Journalists correctly classifies most extremes (Left vs. Right) but has trouble identifying Left/Center and Right/Center.
- Assuming the Journalists provides the ground truth (current labels are based on publishers - the upper bound for accuracy is 65%

The Model **The Journalist**



Publisher	True Label	Predicted Label	Journalist Label	Journalist Comment
Atlantic	Liberal	Conservative	Center	factual article
Breitbart	Conservative	Liberal	Center	factual article
Washington Post	Liberal	Conservative	Center	factual article
★ National Review	Conservative	Liberal	Conservative	
Guardian	Liberal	Conservative	Conservative	
★ New York Times	Liberal	Conservative	Liberal	
★ New York Times	Liberal	Conservative	Liberal	this is a book review
Fox News	Conservative	Liberal	Center	factual article
New York Post	Conservative	Liberal	Center	factual article
New York Post	Conservative	Liberal	Center	factual article
★ National Review	Conservative	Liberal	Conservative	
★ Washington Post	Liberal	Conservative	Liberal	
Fox News	Conservative	Liberal	Center	factual article
★ Atlantic	Liberal	Conservative	Liberal	
★ Breitbart	Conservative	Liberal	Conservative	
★ Guardian	Liberal	Conservative	Liberal	
Reuters	Center	Liberal	Liberal	factual article
★ Reuters	Center	Conservative	Center	factual article
★ Business Insider	Center	Conservative	Center	factual article
★ Business Insider	Center	Liberal	Center	factual article
★ Finanical Times	Center	Liberal	Center	factual article
★ Finanical Times	Center	Conservative	Center	factual article
Quartz	Center	Liberal	Liberal	just slightly liberal
★ Quartz	Center	Conservative	Center	factual article

Future Work

- Investigate limitations of current labeling - algorithm might be learning style of writing rather than actual bias – ultimately get human labeled articles.
- Continue the error analysis
  - Can the misclassified examples be used to get new labels?
  - Let a journalist mark phrases that were critical for the classification.
- Use an Attention model to assess the phrases that were critical for the algorithm in the classification task, use additional features (quotes)

References

[1] Pennington, J., Socher, R. and Manning C.D. 2014. Glove: Global vectors for word representation. In *Empirical Methods for Natural Language Processing (EMNLP)*, pp. 1532-1543.
[2] Yang, Z., Yang, D., Dyer C., He, X., Smola, A. and Hovy E. 2016. Hierarchical attention networks for document classification. In *Proceedings of the 2016 Conference of the North America Chapter of the Association for Computational Linguistics: Human Language Technologies*.