Testing Computer Generated Disinformation with Human Written Fake News Using the Neural Model Grover

Aaron Chu-Carroll

Fake News

- Denotes news written to intentionally deceive or sway public opinion

- False news stories are 70% more likely to be shared than stories from reliable sources

Human fact-checking websites exist...
→ **cannot keep up with the rapid output of fake news**
→ **computer models rapidly generate these articles**
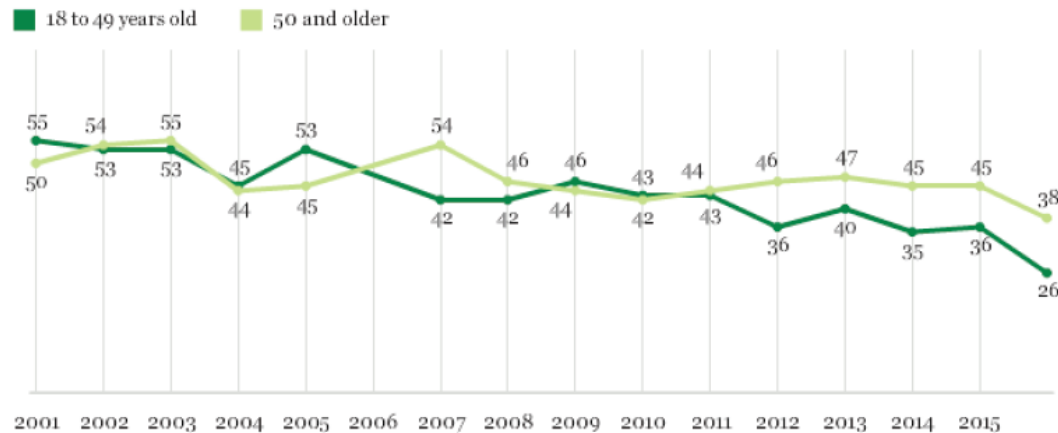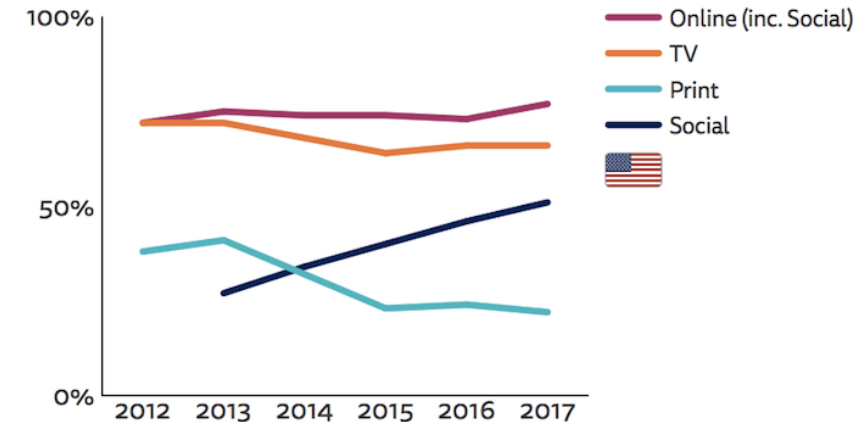

Total Facebook Engagements for Top 20 Election Stories

MAINSTREAM NEWS

FAKE NEWS

8.7 million

7.3 million

15 million / 12 million / 9 million / 6 million / 3 million

Feb.–April   May–July   Aug—Election Day

ENGAGEMENT REFERS TO THE TOTAL NUMBER OF SHARES, REACTIONS, AND COMMENTS FOR A PIECE OF CONTENT ON FACEBOOK SOURCE: FACEBOOK DATA VIA BUZZSUMO

https://www.vox.com/new-money/2016/11/16/13659840/facebook-fake-news-chart

Figure 1.1 Trust in Mass Media, by Age
% Great deal/Fair amount of trust

■ 18 to 49 years old   ■ 50 and older

55 54 55   53   54
50   53 53 45   46 46   43 44 46 47 45 45
            44 45 42 42 44 42 43   40   38
                              36   35 36
                                              26

2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015

GALLUP

https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook#.qqE7PoA2QI

SOURCES OF NEWS 2012-17 USA

— Online (inc. Social)
— TV
— Print
— Social

100%   50%   0%
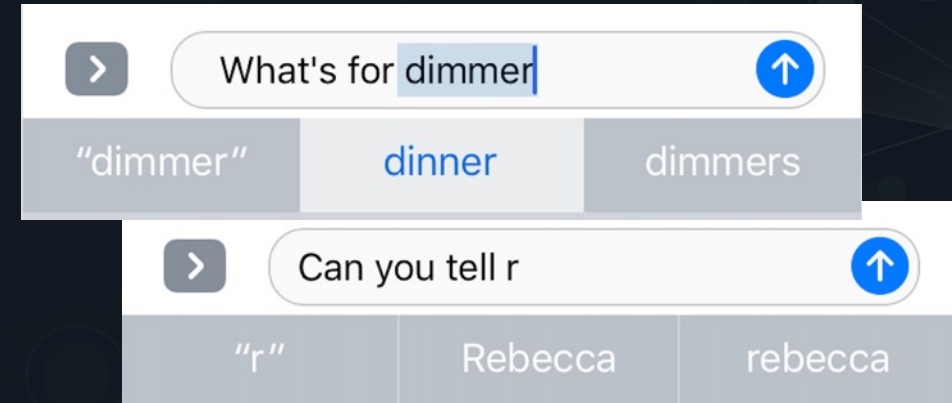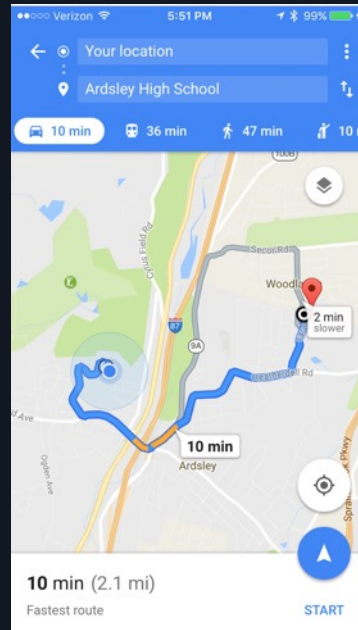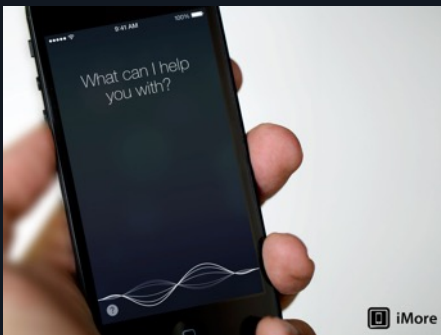
2012 2013 2014 2015 2016 2017

https://www.businessofapps.com/news/social-media-apps-are-the-go-to-news-source-for-51-of-people/

1

Natural Language Processing

- The study of the interaction between computers and human language
- Focuses on deciphering and interpreting language

# Using Machine Learning for Language Processing

- NLP research uses machine learning models to process data

- Uses algorithms to find correlations in large datasets

- Models can identify features, or characteristics, that separate one type of datapoint from another
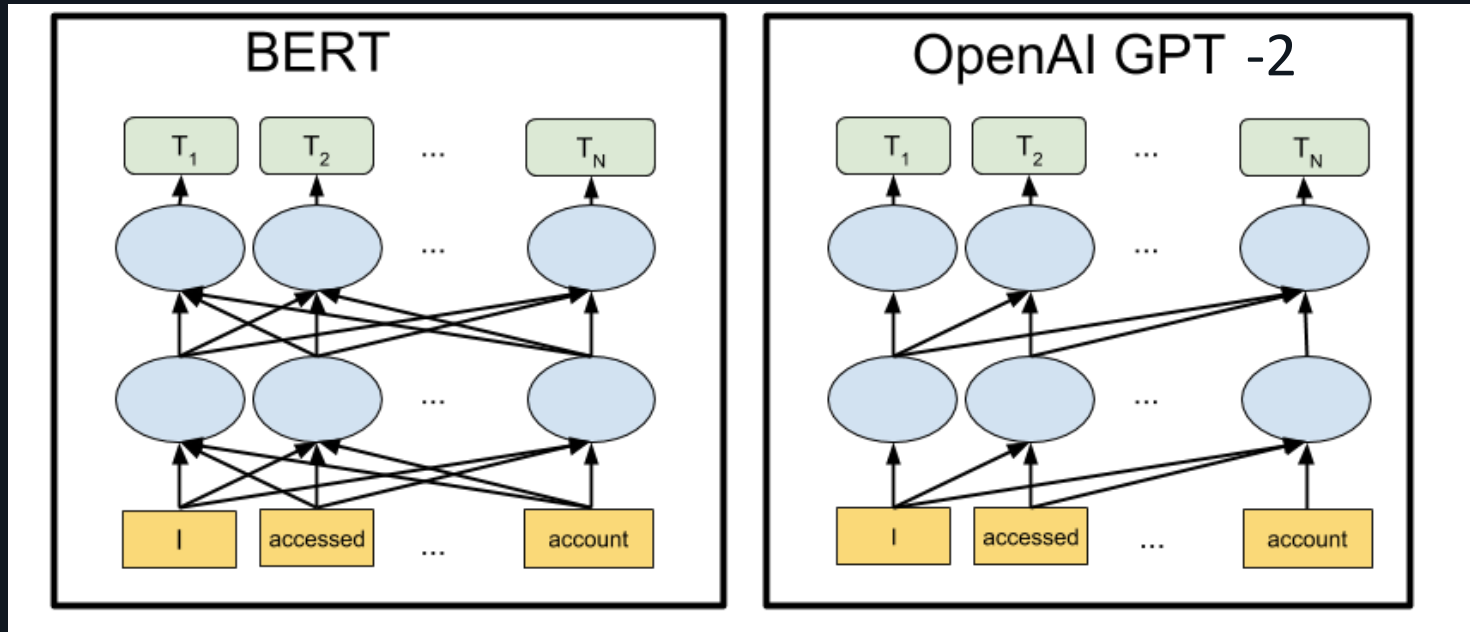

https://www.eventshigh.com/detail/bangalore/b80f5130003a636ad1c32bfe0e3d35e6-bootcamp-on-natural-language-processing
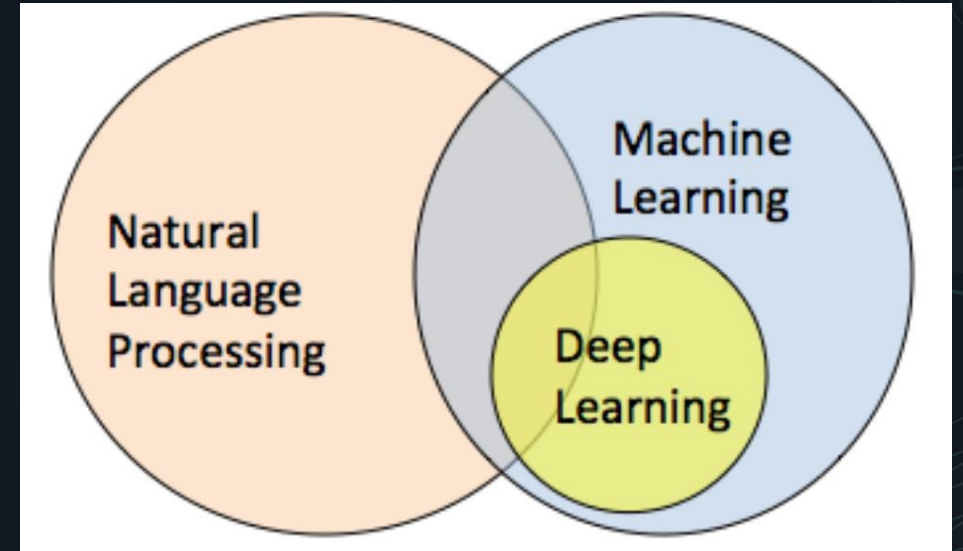
Deep Learning: Layered machine learning models creating a larger neural network

- A greater number of layers makes a more precise and accurate model

Dataset

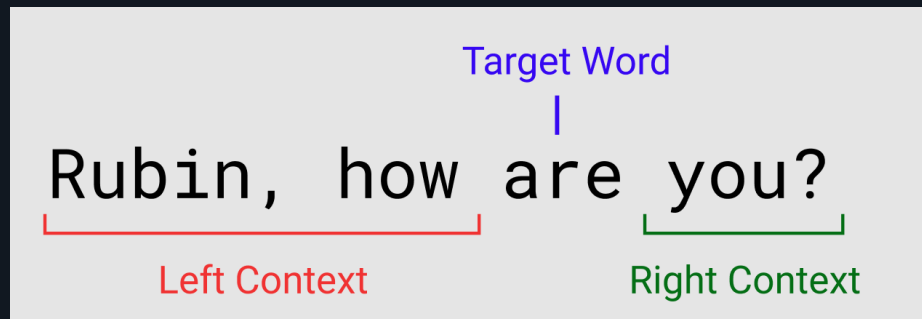| Training Set	Test Set
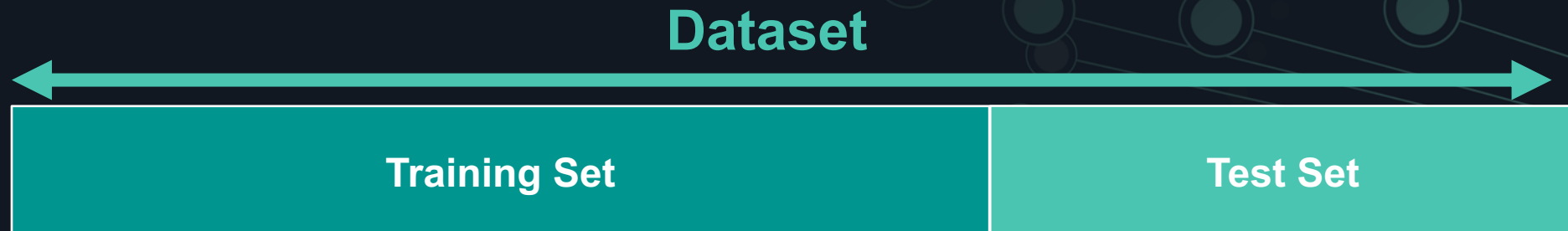
Grover Learning Model

Bidirectional Learning   Unidirectional (Left to Right) Learning

- Grover is modeled after GPT-2
- Grover's three sizes are on par with the BERT and GPT-2, with Grover Mega (it's largest size) at 48 layers
- Trained with computer generated news and human-written news

Grover is capable of detecting, as well as creating fake news

Goals/Hypothesis

Research Question

How would Grover perform as a discriminator for human-written fake news?

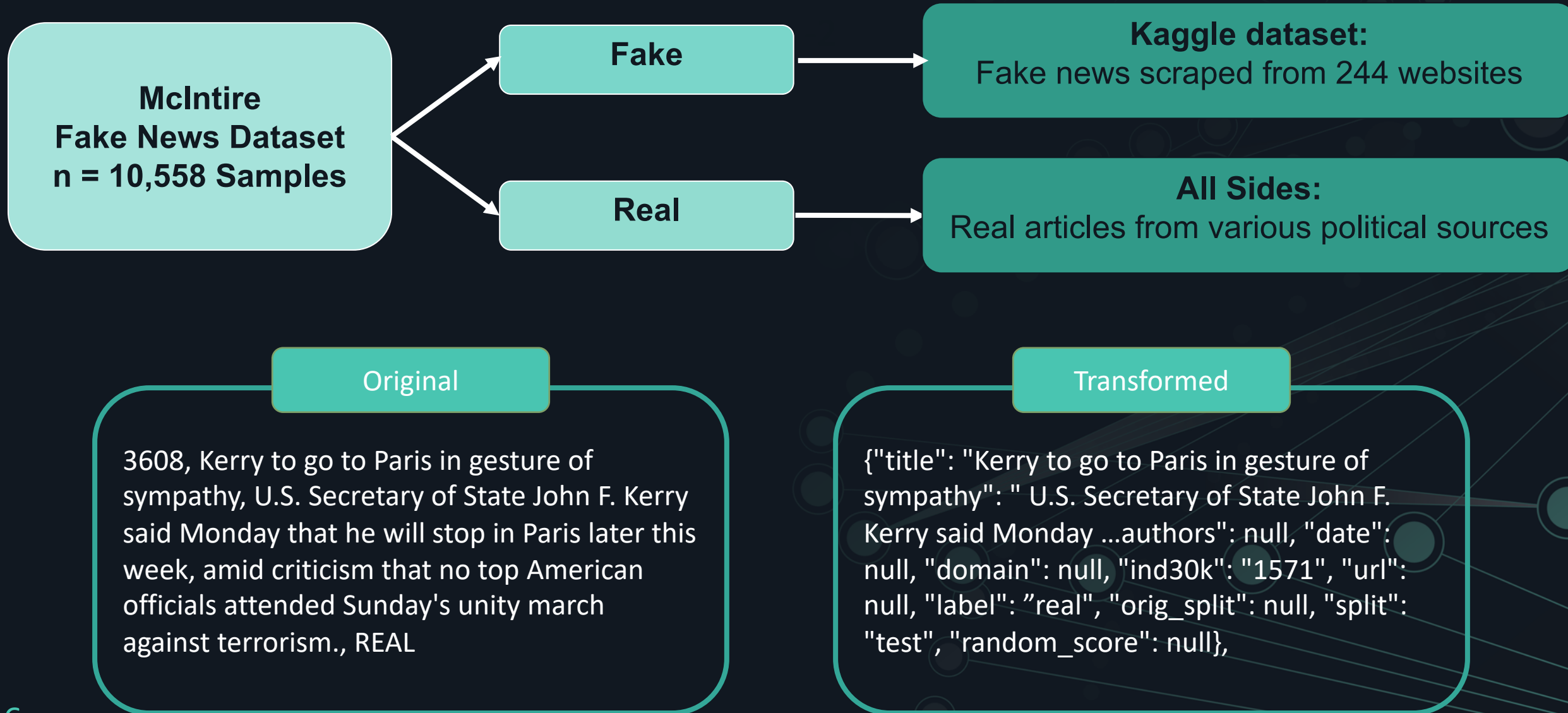| Goal	Hypothesis
Train and test Grover's discrimination model with a dataset of real and fake human news	Grover will perform significantly better when trained with human news than with neural news
Compare with Grover's published results	

Methodology: Datasets and Preprocessing

McIntire
Fake News Dataset
n = 10,558 Samples

Fake → Kaggle dataset:
Fake news scraped from 244 websites

Real → All Sides:
Real articles from various political sources

Original

3608, Kerry to go to Paris in gesture of sympathy, U.S. Secretary of State John F. Kerry said Monday that he will stop in Paris later this week, amid criticism that no top American officials attended Sunday's unity march against terrorism., REAL

Transformed

{"title": "Kerry to go to Paris in gesture of sympathy": " U.S. Secretary of State John F. Kerry said Monday ...authors": null, "date": null, "domain": null, "ind30k": "1571", "url": null, "label": "real", "orig_split": null, "split": "test", "random_score": null},

Methodology: Training and Testing

Train Grover with original input data

Train Grover with McIntire's Fake News Dataset
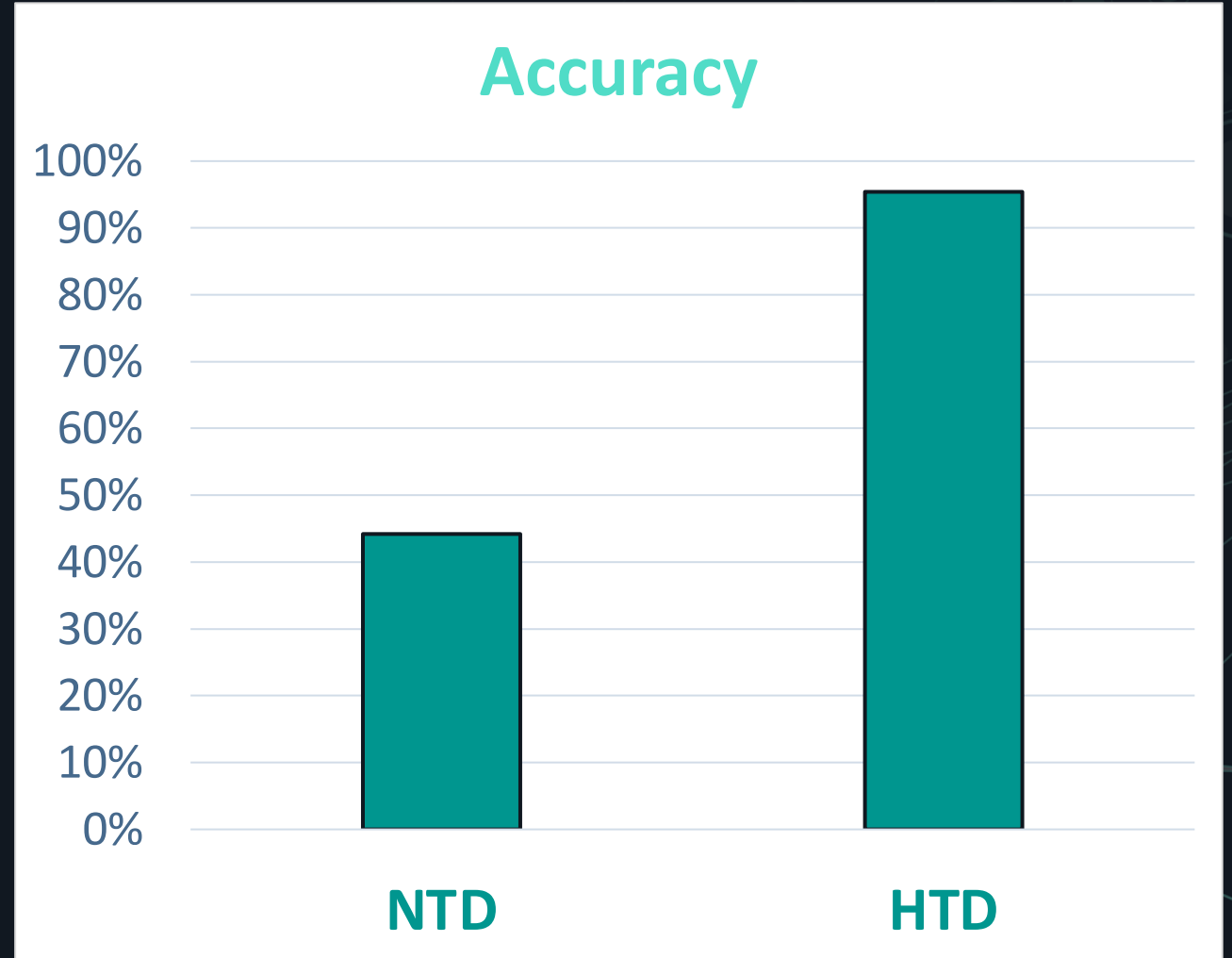
Test Grover with McIntire's Fake News Dataset

Iteration 1	Test	Train	Train	Train	Train
Iteration 2	Train	Test	Train	Train	Train
Iteration 3	Train	Train	Test	Train	Train
Iteration 4	Train	Train	Train	Test	Train
Iteration 5	Train	Train	Train	Train	Test

- Cross Validation is a procedure used if a quantity of data is insufficient for a test/train split
- 5-fold cross-validation used for McIntire's data

Results

	Neural Train Data	Human Train Data
# of Training Instances	10,000	6,336
# of Testing Instances	6,336	6,336
Accuracy	44.2%	95.4%

Experimental results with dataset instances

Accuracy

Results of Grover model tested with the McIntire Fake News Dataset, trained with different datasets

Conclusion

- Today, the biggest problem in Fake News is social media spread
- Grover shows strong capability in detecting fake news based on language patterns

IT'S OVER: Hillary's ISIS Email Just Leaked & It's Worse Than Anyone Could Have Imagined...

POSTED BY FRIENDSOFSYRIA IN WAR CRIMES   ≈ 851 COMMENTS

— Hillary Clinton, Friend of the Syria people? Like the USA

Today Wikileaks released what is, by far, the most entire campaign. This makes Trump's dirty talk v of Barney and Friends.

Even though when Trump called Hillary the 'found truth and 100% accurate, the media has never stopp

Today the media is forced to eat their hats because emails show Hillary, in her own words, admitting to and running ISIS.

John Podesta, Hillary's campaign chair, who was als Obama at the time, was the recipient of the 2014 en today.

Assange promised his latest batch of leaks would le Hillary, and it looks like he was not kidding. The e

Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement

TOPICS: Pope Francis Endorses Donald Trump

WikiLeaks CONFIRMS Hillary Sold Weapons to ISIS... Then Drops Another BOMBSHELL!

Kosar
Breaking News Editor
August 4, 2016

From The Web   Sponsored

Conclusion

Goal
To determine if Grover's discrimination model could be used for detection of neural fake news

Hypothesis
Since the model is trained on the test data, it is assumed that the human data testing round will perform better.

Results
The human data testing round performs significantly better than the neural testing round

Conclusion

Limitations

- Lack of funding, limits capabilities in Deep Learning
- Use of Google Colaboratory
- Limited access to recent datasets

Future

- Include more human news datasets
- Use Grover's Mega size model
- Expand study to GPT-2 Model
- Expand study to BERT Model

Works Cited

[1] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever. "Language Models are Unsupervised Multitask Learners". In: (2019).

[2] Adams, C. (2019, July 17). Expensive, Labour-Intensive, Time-Consuming: How Researchers Overcome Barriers in Machine Learning. Retrieved from https://journal.binarydistrict.com/expensive-labour-intensive-time-consuming-how-researcher-overcome-barriers-in-machine-learning/.

[3] Baly, R., Karadzhov, G., Alexandrov, D., Glass, J., & Nakov, P. (2018). Predicting Factuality of Reporting and Bias of News Media Sources. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing.

[4] Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday*, *21*(11). doi: 10.5210/fm.v21i11.7090

[5] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language under-standing.arXiv preprint arXiv:1810.04805, 2018.

[6] Jiawei Zhang, Limeng Cui, Yanjie Fu, and Fisher B Gouza. 2018. Fake News Detection with Deep Diffusive NetworkModel.arXiv preprint arXiv:1805.08751 (2018)

[7] Kosoff, M. (2018, January 29). The Fake-News Epidemic Is Worse Than We Imagined. Retrieved from https://www.vanityfair.com/news/2018/01/the-fake-news-epidemic-is-worse-than-we-imagined.

[8] Lazer DMJ, Baum MA, Benkler Y et al.: The science of fake news.Science 359:1094–6, 2018x

[9] LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521,** 436–444 (2015) doi:10.1038/nature14539

[10] McIntire, G. (2018, April 18). How to Build a "Fake News" Classification Model. Retrieved from https://opendatascience.com/how-to-build-a-fake-news-classification-model/.

[11] Rowan Zellers, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner, and Yejin Choi. 2019. Defending Against Neural Fake News arXiv:1905.12616.

[12] S. Vosoughi, D. Roy, S. Aral, The spread of true and false news online. Science359, 1146–1151 (2018).

[13] Welch, R. (2017, April 20). PolitiFacts guide to fake news websites and what they ... Retrieved from https://www.politifact.com/punditfact/article/2017/apr/20/politifacts-guide-fake-news-websites-and-what-they/.

Testing Computer Generated Disinformation with Human Written Fake News Using the Neural Model Grover

Aaron Chu-Carroll