

fake-real-news: A Fake News Classifier Proposal

Huang, Sheng 3035534103 & Li, Yik Wai 3035566015

Background

In July 1995, Terry Pratchett said in an interview with Bill Gates, founder of Microsoft, that “Let’s say I call myself the Institute for Something-or-other and I decide to promote a spurious treatise saying the Jews were entirely responsible for the second world war and the Holocaust didn’t happen, and it goes out there on the internet and is available on the same terms as any piece of historical research which has undergone peer review and so on. There’s a kind of parity of esteem of information on the net. It’s all there: there’s no way of finding out whether this stuff has any bottom to it or whether someone has just made it up.” Bill, at the time, did not believe in this version of the future, saying there would be authorities on the internet and that the ways of fact checking would be much more sophisticated than in print. [1]

Nearly 26 years later, Bill’s optimism turns out to be not in accordance with the reality. With the rise of social media, the propagation of fake news is at a dangerous level, threatening societies and democracies. [2] Today, fake news is enjoying better popularity than some serious report and could undermine authentic journalism. [3][4] Fake news has also been weaponized by authoritarian regimes to undermine the integrity of the political systems of other countries. [5][6]

Tackling the problems brought about by fake news is urgent. However, we must find a solution that is not worse than the problem. In the hands of authoritarian governments, fact checking may mean censorship. [7][8][9][10] In 2019, the Singaporean Parliament passed the Protection from Online Falsehoods and Manipulation Act, aiming to tackle the problem of fake news and false information. However, critics say it is a tool for the government to crack down on dissent. [11][12]

To avoid any entity, government or otherwise, being the arbiter of truth and making decisions on the authenticity of information for its own good instead of for the good of the whole society and to mitigate the damage that fake news could bring upon our societies, we need to develop an application that can judge the authenticity of a news piece fairly, efficiently and effectively. To make an application, other mechanisms, such as appeal processes, are also required to ensure the fairness of the system.

Objectives

To build and train a model for the fair and impartial judgement of authenticity of a news article.

To integrate that model with a web framework and make it a web app.

To provide model update given the availability of newer and/or better datasets and user feedback.

Application

The application shall be a web application.

Since some news articles, especially quality content, are hidden behind a paywall and the application may encounter copyright issues if we bypass that wall without publisher consent, the frontend shall have input fields for the user to input article information, including title, author, date, publisher and text, manually. After input, the web app will send the data to the backend for classification. The result will be sent back to the frontend and the user could choose to send a feedback, which will be sent to the backend.

In the backend, the server will receive the news article data, classify it with the trained model, record the data and classification result in the database for future review and send back the result. When the backend receives a user feedback, it will store it for future review. From time to time, the product owner could adjust the model according to the availability of newer and/or better datasets and user feedback.

Dataset

Fake News Corpus by [@several27](https://github.com/several27) on GitHub.
<https://github.com/several27/FakeNewsCorpus>

The dataset covered nearly 10 million articles from hundreds of domains, including some news articles from *The New York Times* and *WebHouse English News Articles* for better balance.

The data mining was primarily done in 2018-2019, with about three quarters complete. Development of the dataset is not active. However, it is the best available dataset we could find in terms of its size, diversity of sources, date of mining, etc.

There are limitations as the dataset creator pointed out, but it shouldn’t affect practical use.

Columns: id, domain, type, url, content, scraped_at, inserted_at, updated_at, title, authors, keywords, meta_keywords, meta_description, tags, summary, source (opensources, nytimes, or webhose)

type include fake for Fake News, satire for Satire, bias for Extreme Bias, conspiracy for Conspiracy Theory, state for State News, junksci for Junk Science, hate for Hate News, clickbait for Clickbait, unreliable for Proceed With Caution, political for Political, reliable for Credible.

Methodology & Model

We will use fastai, a deep learning library based on PyTorch that provides higher level APIs. We will use TextBlock in fastai, which automatically handles tokenization and numericalization, and create dataloaders. Then, we will fine-tune a pretrained language model.

To fine-tune a language model, we use fastai's language_model_learner with AWD-LSTM architecture from Smerity et al. [13] We'll need to use embeddings to convert integer word indices into activations that could be used in a neural network. We'll feed these embeddings into a recurrent neural network using AWD-LSTM. Then we can start training.

When we are satisfied with the performance of the model, we will export the model and build a web app with it.

Tools

For Web:

Linux: A UNIX-like operating system on which we will develop and run our application, which supports a lot of open source software and powerful tools

Python3: A super powerful programming language that has numerous libraries

Flask: A light-weight web framework for Python

Gunicorn: A Python WSGI server for UNIX

SQLite3: A Relational Database Management System that's fit for development

(We don't plan on deploying the app to a host because the size of the project may exceed the free quota of any platform. If we were to deploy, we would switch to PostgreSQL)

For Model:

Anaconda: A distribution of Python and R that is aimed at data science, which we will use to manage environments in HKU CS GPU Farm

Python3: A super powerful programming language that has numerous libraries

Jupyter Notebook: A open source web application that helps us to better interact with code, texts, images, etc., which we will use to train our model

fastai v2: A deep learning library based on PyTorch that provides higher level APIs, which will provide more automation to our development process

Platforms:

GitHub: A platform for hosting software development and version control using Git, which we will use to host and track our development files

Google Colab: A *Jupyter Notebook*-like platform that provides free GPU resources to researchers alike, perfect for light tasks

HKU CS GPU Farm: The CS Departments' own GPU cluster that provides students and faculty with GPU computing power

Timeline

13/03: Submission of Proposal

14/03 – 18/03: Data Engineering - Li
Download the dataset onto GPU Farm and process the data from the chosen dataset to ready-to-use status.

14/03 – 20/03: Creating Initial Prototype – Huang
Create the initial model prototype and make it ready-to-test status

19/03 – 20/03: Experiment Design – Li
Design experiments that could test the effectiveness of the initial model

21/03 – 23/03: Preliminary Experiments – Li
Carry out the designed experiments and report the results

24/03 – 31/03: Modification and Retesting – Li, Huang
Given the test results, modify the model and conduct the testing in turns

01/04 – 03/04: Writing Report – Li, Huang
Write the report for submission

03/04: Submission of Interim Prototype

04/04 – 17/04: Modification and Retesting – Li, Huang
Given the test results, modify the model and conduct the testing in turns

04/04 – 17/04: Building Web App – Huang
Build the web app and integrate the model with it

18/04 – 19/04 – Final Testing – Li
Explore final bugs and limitations and record final testing results

20/04 – 21/04: Making Video – Li
Record and edit the video for presentation

22/04 – 24/04: Writing Report – Li, Huang
Write the report for submission

24/04: Submission of Final Work

Limitations

While there are certainly further limitations to explore later in the project, the limitations that are already looming over are the limitations of our dataset. No matter how smart the deep learning algorithm, it is at most as smart as the data we have. The dataset is not completely processed, which although the creator claims won't affect practical use, it's still not the most ideal situation, and it is slightly outdated if we want to focus on the reports of more recent events, especially after the spread of SARS-CoV-2 and the political crises worldwide since 2019.

References

[1] Flood, Alison (May 30, 2019). ["Terry Pratchett predicted rise of fake news in 1995, says biographer"](#). The Guardian.

[2] Andreas Kaplan (2020). Artificial Intelligence, [Social Media, and Fake News: Is this the end of democracy?](#). Digital Transformation in Media & Society, DOI: 10.26650/B/SS07.2020.013.09.

[3] Merlo, Carlos (2017), ["Millonario negocio FAKE NEWS"](#), Univision Noticia.

[4] Tsang, Stephanie Jean (August 31, 2020). ["Motivated Fake News Perception: The Impact of News Sources and Policy Support on Audiences' Assessment of News Fakeness"](#). Journalism & Mass Communication Quarterly. doi:10.1177/1077699020952129.

[5] Macfarquhar, Neil (August 28, 2016). ["A Powerful Russian Weapon: The Spread of False Stories"](#). The New York Times.

[6] ["Russia is targeting French, Dutch and German elections with fake news, EU task force warns"](#), The Telegraph, January 24, 2017.

[7] Orlowski, Andrew (November 21, 2016), ["China cites Trump to justify 'fake news' media clampdown. Surprised?"](#), The Register.

[8] Pascaline, Mary (November 20, 2016), ["Facebook Fake News Stories: China Calls For More Censorship On Internet Following Social Media's Alleged Role In US Election"](#), International Business Times.

[9] ["After Trump, Americans want Facebook and Google to vet news. So does China."](#), The Washington Post, November 17, 2016.

[10] Dou, Eva (November 18, 2016), ["China Presses Tech Firms to Police the Internet—Third-annual World Internet Conference aimed at proselytizing China's view to global audience"](#), The Wall Street Journal.

[11] Han, Kirsten (2019). ["Big Brother's regional ripple effect: Singapore's recent 'fake news' law which gives ministers the right to ban content they do not like, may encourage other regimes in south-east Asia to follow suit"](#). Index on Censorship. 48 (2): 67–69. doi:10.1177/0306422019858296. S2CID 197704061.

[12] ["Singapore 'fake news' law ensnares government critics"](#). Reuters. 16 December 2019. Retrieved 11 July 2020.

[13] Merity, Stephen, Nitish Shirish Keskar, and Richard Socher. (2017) ["Regularizing and Optimizing LSTM Language Models"](#). arXiv:1708.02182.