

Springboard Capstone Project 3 – Fake news detection – project proposal

Jonny Pearce, 2 July 2023

Introduction/problem statement

The reliability of information is key to businesses, governments and media organisations around the world. The rise of misinformation and fake news has a massive impact on business and government to effect change or develop appropriate responses/behavioural change within society, as well as sell products and provide services. This is an issue that cuts across all content/subject areas, and is ultimately one of trust and credibility. While trust and credibility could be seen as possibly a slightly intangible value, or something that is secondary to commercial or pure business issues, one only has to look at countless examples where lack of trust or credibility has had a catastrophic impact on business interests and concerns. For example, the international banking/financial crises of 2008 were largely about trust and credibility in financial information regarding consolidated mortgage debt and the financial instruments used to package them. More recently in the last decade, the Volkswagen/Audi group has had to deal with the multimillion fallout from the cover-up regarding its falsification of the emissions data for its vehicles.

With the rise of social media and online forums, and with customers becoming content generators, businesses have even more to contend with. On one level, the number and type of companies that publish information has increased massively as technology has grown and the internet has developed and expanded. It's not just traditional publishers or news/media organisations that publish information and therefore have to ensure the accuracy and reliability of their information, and how trusted it is. Any organisation that runs any aspect of social media or an online community now falls into that category (which at one end of the spectrum can be as simple as having a comments feature on your company content).

On another level, businesses need to be aware of and manage their reputations and faith/trust in their products from user/purchaser reviews of their products and services on a whole range of internet and social media sites.

On yet another level, there are a whole range of legal and compliance issues regarding content and content management. Traditional legal issues and liabilities affect anyone responsible for false, abusive, misleading, offensive, etc, information. Regulation and liability will only increase with international authorities around the world all considering new and enhanced responsibilities to manage and control falsification of information, particularly in the light of reported developments in machine learning and artificial intelligence.

Thus, knowing what is true and what is fake in terms of protecting your own business (either as a publishing organisation or as a product producer or service provider) and enhancing your reputation and profile.

So, reliable and effective tools that can help governments, businesses and media/publishing organisations manage fake news and false/misleading information will only become more and more important.

Proposed project/solution

In order to address this problem, it is proposed to build a fake news detector that can distinguish between fake and true content with a high degree of accuracy and reliability.

Datasets

There are a whole range of readily available datasets available on the internet to test the development of such models, but after a review of a number of them, it was decided to focus on ones with a decent amount of text so that a reliable model could be developed. In practice many of the available datasets were too small or limited.

The project will use the following dataset from <https://www.kaggle.com/datasets/jruvika/fake-news-detection?select=data.csv>, which contains headlines, full content and binary ratings on over 4000 news articles.

It is also important that for any dataset used, that any “easy to identify” features are removed, eg, publisher or website url details, which may be a give-away for trusted or non-trusted sources. What we are interested in here is being able to evaluate the credibility or trustworthiness of anonymised or unknown data, rather than looking at say, whether the BBC or New York Times has published.

Methodology

In essence, this is a classification problem (a binary choice between fake and true) involving natural language processing. The general approach will involve the following steps:

- Data wrangling/cleaning/standardisation
- Exploratory data analysis.
- Vectorization of the text data (with consideration given to different text vectorization methods).
- Model training, building and testing, using a number of different machine learning classification algorithms.
- Model selection (including refinement and parameter tuning).

Outputs/deliverables

The project will deliver the following:

- A final, selected model for use in fake news detection.
- Jupyter notebooks containing the code, notes and commentary on the different steps of the project.
- A project report, outlining the project work, analysis, model development and final results.
- A PowerPoint slidedeck, summarising the project.