
GMIT
INSTITIÚID TEICNEOLAÍOCHTA NA GAILLIMHE-MAIGH EO
GALWAY-MAYO INSTITUTE OF TECHNOLOGY

Contents

List of Figures

Chapter 1

Introduction

1.1 Context

Fake news (also known as junk news, pseudo-news, or hoax news) is a form of news consisting of deliberate disinformation or hoaxes spread via traditional news media [1] (print and broadcast) or online social media.Digital news has brought back and increased the usage of fake news, or yellow journalism.The news is then often reverberated as misinformation in social media but occasionally finds its way to the mainstream media as well.

The main aim behind fake news is to mislead and sway public opinion surrounding an agency, entity, or person, and/or gain financially or politically, often using sensationalist, dishonest, or outright fabricated headlines to increase readership. Similarly, clickbait stories and headlines earn advertising revenue from this activity Fake news undermines proper news coverage and makes it more difficult for journalists to cover signifigant news stories. The term 'fake news' became very popular in 2016 when during and after his presidential campaign and election, Donald Trump popularized the term 'fake news' in this sense, regardless of the truthfulness of the news, when he used it to describe the negative press coverage of himself.

"Fake news" has acquired a certain legitimacy after being named word of the year by Collins, following what the dictionary called its "ubiquitous presence" over the last 12 months. At present technology companies are fighting an epedemic of fake news. Facebook has deleted 3.39 billion fake accounts from October 2018 to March 2019. Whatsapp is deleting 2m accounts per month. Fake news spread through Whatsapp in India is responsible for 30 mob lynchings that were said to have been triggered by incendiary rumours spread using the app.

The dynamics and influence of fake news on Twitter during the 2016 US presidential election remains to be clarified, however, a study found that in the five months preceding the election, a dataset of 171 million tweets were gathered and a subset of 30 million tweets, from 2.2 million users, which contain a link

to news outlets was gathered. The study found that of those 30 million tweets, approximately 25% of the news outlets linked containted outright false or extremely biased information[2]. It's important to remember that this is just in relation to American politics.
In Ireland fake news is heavily prevalent. Irish people are among the biggest consumers of Facebook and Twitter in Europe. A survey by Deloitte in 2017 found that Irish adults look at their mobile phone 57 times a day (in comparison to a European average of 41 times). Some 16% admit to looking at their phone more than 100 times a day (against a European average of 8%). A study attempted to examine false memories in the week preceding the 2018 Irish abortion referendum. Participants (N = 3,140) viewed six news stories concerning campaign events—two fabricated and four authentic. Almost half of the sample reported a false memory for at least one fabricated event, with more than one third of participants reporting a specific memory of the event [1].
This problem is getting worse. There is just too much to be gained from spreading fake news that it's becoming more and more prevalent. Much like telephone scams, we often think how can people believe such utter nonsense especially when it comes to badly faked news articles and stories shared on social media and other outlets. However this type of fake news is just the surface. The real danger lies in the stories and other types of information that are expertly crafted to make us believe them. The best lies are often coated in a layer of truth which makes the job of labelling and classifying fake news extremely difficult.

1.2 My project

My project aims to combat this problem. I aim to build and test a number of machine learning models that can classify fake news. I will then select the best model based on the tests performed and deploy that to the cloud so that users can interact with the model to hopefully minimise the spread of fake news. I will give a brief introduction to each part of the project. I will dicuss these parts in much more depth in the Methodology and System design sections of the dissertation.
I have two objectives for this project.

- Create and test a multitude of machine learning models to classify fake news
- Deploy the best model via a chrome extension and have it classify fake news

1.2.1 Models

I designed and implemented five machine learning models during the course of this project. The models are listed below

- Support Vector Machine(SVM).

- Long Short Term Memort(LSTM).
- Naive Bayes
- Neural Net using Keras
- Neural Net using Tensorflow

I will breifly list why I chose these 5 models for this project. I will dicuss at a basic leve their pros and cons. I will dicuss in more depth how they work but for now I will give a high level overview of each model and it's architecture.

Support Vector Machine(SVM)

In machine learning, support-vector machines (SVMs, also support-vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Traditionally SVM's have been seen as having primary usage in linear classification, However SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces. Feature space refers to the n-dimensions where your variables live (not including a target variable, if it is present).For example, consider the data set with:
Target

$$1.\mathbf{Y} \equiv Thicknessofcartiresaftersometestingperiod$$

Variables

$$1.X_1 \equiv Distancetravelledintest$$

$$2.X_2 \equiv Timedurationofthetest$$

$$3.X_3 \equiv Amountofchemical\mathbf{C}intires$$

$$The feature space is \mathcal{R}^{\ni}$$

1.2.2 Data

Data used to train the models was retrived from two websites. Those websites are kaggle and github. I decided to make a number of csv files out of the data from the two sources in order to test the models accuracy. In terms of cleaning the data I wrote two scripts called clean.py and clean2.py. The data wasn't very clean and uniform. Certain fields were missing as well as the data containing an awful lot of non english characters. I had to spend a number of weeks cleaning the data to ensure that I used good data to train the models. I also used this time to validata a random sample of the data to ensure as best I could that the data was accuractely labelled. Due to the size of the data I was unable to verify all the data was accuracte but I was happy with the accuracy of the random sample so I am confident that the data is accurate.

1.2.3 Testing

I used a technique called cross validation to ensure the accuracy of my models. I used an 80/20 test train split. Validation was performed on the testing data after each susccessfult epoch of training and in the end a confusion matrix was calculated. I then compared and contrasted each of the confusion matrices of the various models, calculating the sensitvity and specificity of each model. I will dive deeper into this in the methodology chapter of this dissertation

1.2.4 Deployment

I decided that a useful way to interact with the model would be through the use of a chrome extension. The same principles of deployment I will discuss also apply to Mozilla Firefox. A chrome extension. Extensions are small software programs that customize the browsing experience. The model would be deployed to the cloud and the extension would allow the user to interact with the model via http. The extension would, with permission, send the news article to the model. The model would then classify it and return the result to the user in the form of a status indictor. Green for accuracte and red for fake. Now, the real world is very rarely this black and white and there are often degrees of truth in fake news but as this is more of a proof of concept rather than a full scale, production ready deployment, I am content with having two degrees of truth. Those being real and fake.

Chapter 2

Methodology

In this chapter I will describe the way I went about this project. I will provide details on my approach regarding the research and actualy development work I did for this project. I would like to start with the data.

2.0.1 Data

I began looking for the appopriate data on Kaggle. Kaggle is an online community of data scientists and machine learning practitioners. Kaggle has repositories of publicly avaliabe data for anyone to download and build models for. The data avaliable on kaggle is oftentimes alread pre cleaned so when you download the data it is ready for use straight away. However in my case I had two major issues when looking on Kaggle for the data.

- The data on Kaggle wasn't in the format I wanted.
- There wasn't enough data avaliable on the topic of fake news.

Nevertheless I managed to download a number of datasets from Kaggle. I loaded the data in a jupyter notebook using the pandas library for python and took stock of the data I had retrived. The columns for the datasets were labelled differently. The datasets also had different approaches to the labelling of the correct classifcation of the data with some datasets using numerical values(1 for true, 3 for semi-true and 5 for fake) and others using text values(true for true and fake for false). On github I managed to find a datset called all-the-news by Andrew Thompson that conatined approximately 143,000 news articles.

2.0.2 Models

Now that I had my data I needed to decide on how best to select the best model to solve the problem. I had two selection critera.

- Highest accuracy

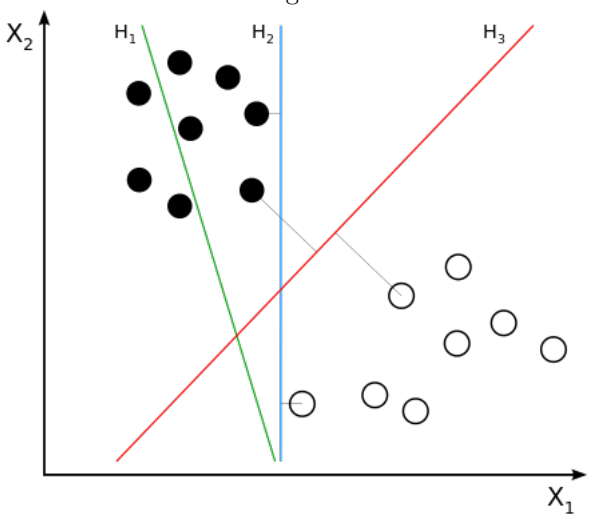
- Scalability

As I was the sole developer for this project I had to make sure that the models I designed were at the level that one person could do it. State of the art models take years and teams full of experts to create and I was the only person doing this project and I had only a short period of time to do it so I had to make my expectations realistic. The first model I decided to create was the support vector machine(SVM)

Support Vector Machine(SVM)

The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space(N — the number of features) that distinctly classifies the data points. An example is shown below.

Figure 2.1: An overview of a SVM

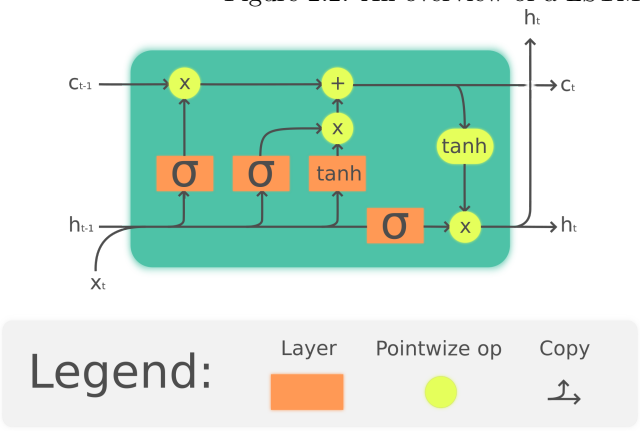


We can see a number of lines drawn in the figure above. A support vector machine aims to find the best line that seperates the two classes. In this case the line H3 does that the best. More formally, a support-vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks like outliers detection.

Long Short Term Memory LSTM

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture. An example is shown below
All RNNs have feedback loops in the recurrent layer. This lets them maintain information in 'memory' over time. But, it can be difficult to train standard RNNs to solve problems that require learning long-term temporal dependencies.

Figure 2.2: An overview of a LSTM cell



This is because the gradient of the loss function decays exponentially with time (called the vanishing gradient problem). LSTM networks are a type of RNN that uses special units in addition to standard units. LSTM units include a 'memory cell' that can maintain information in memory for long periods of time. A set of gates is used to control when information enters the memory, when it's output, and when it's forgotten. This architecture lets them learn longer-term dependencies. GRUs are similar to LSTMs, but use a simplified structure. They also use a set of gates to control the flow of information, but they don't use separate memory cells, and they use fewer gates.

2.0.3 Testing and Validation

As I breifly explained in the introduction section of this thesis I used a technique called cross validation to ensure the accuracy of my models. I will now explain cross validation and why I used it in the way that I did. A technique called regularization is also widely used in the field of machine learning. Another technique called dropout began to change the field of machine learning after it emerged around 2012

Cross Validation

Cross validation is a technique used in the field of machine learning that aims to combat two major problems

- Model Overfitting
- Model Underfitting

Model overfitting is when a statistical model or machine learning algorithm captures the noise of the data. Intuitively, overfitting occurs when the model or

the algorithm fits the data too well. Overfitting a model result in good accuracy for training data set but poor results on new data sets. Such a model is not of any use in the real world as it is not able to predict outcomes for new cases.
Model underfitting is when a statistical model or machine learning algorithm cannot capture the underlying trend of the data. Intuitively, underfitting occurs when the model or the algorithm does not fit the data well enough. Underfitting is often a result of an excessively simple model. By simple we mean that the missing data is not handled properly, no outlier treatment, removing of irrelevant features or features which do not contribute much to the predictor variable.
Cross validation allows us to effectively generalise our model. By keeping, In this case 20%, an amount of data for the purposes of testing, we can see how well our model performs on unseen data. We can plot our results using a confusion matrix and this can be one of many ways of comparing our models. In my case this is exactly what I did. After each successfuly epoch of training, validation was performed on the test data and the results were outputted so it was quite easy to spot when my model would be overfitting. This would be seen by having high levels of accuracy in the training data and then having low levels of accuracy in the test data.

Regularization

Another technique that has been proposed to combat the problem of overfitting is called regularization. Regularization modifies the objective function that we minimize by adding additional terms that penalize large weights. In simple terms, it reduces parameters and shrinks (simplifies) the model.

Dropout

Dropout is a very different kind of method for preventing overfitting that has become one of the most favored methods of preventing overfitting in deep neural networks.While training, dropout is implemented by only keeping a neuron active with some probability p (a hyperparameter), or setting it to zero otherwise. Intuitively, this forces the network to be accurate even in the absence of certain information. It prevents the network from becoming too dependent on any one (or any small combination) of neurons. Expressed more mathematically, it prevents overfitting by providing a way of approximately combining exponentially many different neural network architectures efficiently.

2.0.4 Deployment

Chapter 3

Technology Review

Blah.....

Chapter 4

System Design

Blah.....

Chapter 5

System Evaluation

Blah.....

Chapter 6

Conclusion

Blah.....

Bibliography

[1] Gillian Murphy, Elizabeth F Loftus, Rebecca Hofstein Grady, Linda J Levine, and Ciara M Greene. False memories for fake news during ireland's abortion referendum. *Psychological science*, 30(10):1449–1459, 2019.

[2] Alexandre Bovet and Hernán A Makse. Influence of fake news in twitter during the 2016 us presidential election. *Nature communications*, 10(1):1–14, 2019.