Imperial College London

Department of Electrical and Electronic Engineering

Final Year Project Report 2019

Imperial College London

Project Title:	**Fact or Fake?**
Student:	**Orion M. Mathews**
CID:	**01052855**
Course:	**EE4**
Project Supervisors:	**Dr Mark Witkowski** **Prof. Robert Spence** **Dr James Mardell**
Second Marker:	**Dr David Thomas**

Abstract

News is vital to every major aspect of modern civilisation. It influences the decisions made in politics, economics, health care, and sociocultural interactions. If inaccurate information is spread the consequences can be dire. Manually determining the accuracy of information is tedious, time consuming and, as a result, often disregarded at the individual level. This project report details the development and evaluation of two automated approaches that aim to discriminate between accurate and inaccurate information.

An investigation into the metrics with which news can be analysed was conducted. This investigation led to the first automated approach which analyses the quantity and extremity of opinions present in a news article. Examination of the results revealed no statistically significant difference between the opinion words used in fact and fake news. Further analysis leads to the conclusion that current state of the art sentiment analysis techniques are not advanced enough to differentiate between the complex expression of opinions presented in news pieces.

A second automated approach was constructed in which a consensus score, derived via stance detection, was used to infer the accuracy of a statement. This method utilises a pre-trained artificial neural network which takes a statement and an article, and predicts whether the article agrees, disagrees, discusses or is unrelated to the statement. Based on this neural network a web browser extension was created to make fact checking quicker and easier for news readers. The browser extension was then modified to incorporate the previous sentiment analysis approach. The evaluation showed that enabling the browser extension increased the likelihood a user questioned and checked the accuracy of a social media post. Increased user fact checking is a small but significant step in inhibiting the spread and impact of misinformation.

Acknowledgements

I would like to thank my project supervisors Mark Witkowski, James Mardell and Robert Spence for all the time and guidance they have given me throughout my final year. They have made working on this project engaging and enjoyable and I would not have been able to achieve nearly as much without their help. I would also like to thank Gwen Groocock for her invaluable expertise and advice in a field where I had very little experience. Finally I would like to thank Jon Groocock and in particular my parents, Denis and Julie, for their endless support and inspiration throughout my educational life. The journey would not have been possible without them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Logical decision making is based on information. Inaccurate information gives rise to misguided decisions and the consequent actions can be severe (for instance "Like. Share. Kill." [10]). With the current ubiquity of the internet and social media, inaccurate information has the potential to reach and influence millions of people within an extremely short period of time. Furthermore, several independent studies have shown that, not only is there a serious lack of fact checking performed by the average news reader ([11], [12]), but that incorrect news is actually more likely to be shared than true stories ([13], [14]). Incorrect information has been a problem throughout history and although recent advances in technology have magnified the issue, they may also present a solution.

The aim of this project was to develop an automated method of analysing news in a specific metric so that a news article can be classified as being real, persuasive, false or fake - see table 1.1 for the definitions of each news category. The desired outcome was to aid users in determining the legitimacy of news and therefore hinder the spread of misinformation.

Definitions:	
Information	Information is defined as "facts provided about a situation, person, event, etc." (Cambridge Dictionary [15]). This definition will not assume anything about the bias or truthfulness of the "facts".
News	News is "Newly received or noteworthy information, especially about recent events" (Oxford Living Dictionaries [16]).
Real News	Real news is news that is verifiably correct and contains none, or very little, of the authors opinions.
Persuasive News	Persuasive news is news where the information is correct but the author presents their personal opinion within the news piece.
Fact News	Fact news will be defined as news where the information is correct and no consideration is given to the opinion presence in the article.
False News	False news is news that is incorrect but the misinformation is not intentional. The author was not deliberately producing untruthful news.
Fake News	Fake news is disinformation; the news piece is intentionally incorrect and untruthful (Cambridge Dictionary [17]).
Incorrect News	Incorrect news is news where the facts presented are untruthful and the author's opinion/intent is not considered.
News Metric	A news metric is a feature of news (or news sources) that can be used measured and classify a news piece.

Table 1.1: Summary of word definitions. See appendix A for full definitions with examples and/or discussion.

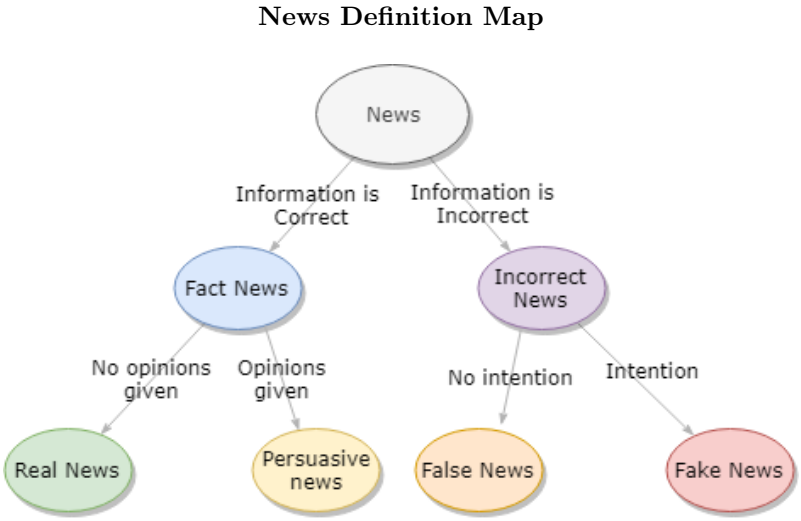

Figure 1.1: Map to summarize each news class and the characteristics that define them.

1.1 Report Structure

Chapter 2 gives a detailed description of the false information problem as well as the current solutions. During the current solutions section, it became apparent that manual approaches are too slow and expensive to effectively address this issue. Though automated methods designed to combat misinformation exist, evidence is given suggesting they fall short of providing a competent solution. This chapter then goes on to cover the necessary background for the proposed solutions, the first of which uses sentiment/opinion analysis while the second relies on a neural network that performs stance detection.

Chapter 3 outlines the requirements needed to be achieved in order for the developed tool to be considered a success. The requirements capture involves detailing the necessary and desired outcomes of the project. Overall, five requirements were established, three of which were necessary while the others were optional. These requirements guided the research and development of the tool.

The requirements capture is followed by an investigation into the metrics that can be used to define and classify news pieces (chapter 4). Of the nine news metrics identified, accuracy and neutrality emerge as the metrics best suited to discriminating between real and fake news. It becomes clear that directly quantifying accuracy and neutrality is exceedingly complex. However, the latter part of this chapter discusses the possibility of inferring the accuracy and neutrality of an article from respective measures in consensus and sentiment.

Chapter 5 describes the design, implementation and testing of an automated approach that seeks to infer the neutrality of a news article via a sentiment score. The proposed tool makes use of VADER - a sentiment lexicon attuned to the analysis of social media posts. A Hypothesis was made claiming that fake news contains more opinions, in both quantity and extremity, than real news. Testing was preformed on two independent fake news data sets. The results were then examined, revealing no statistical difference between the opinions in real and fake news articles. Rigorous evaluation of the method showed the tool to be inadequate at discerning between opinion pieces and standard world news. Therefore, the chapter concludes with the notion that current state of the art sentiment analysis techniques are not sophisticated enough to accurately analyse the opinions presented by news articles.

Chapter 6 begins by documenting a thought experiment designed to identify a method of deducing the trustworthiness of a news source. The experiment confirms consensus as a viable means of determining the accuracy of information. The rest of the chapter details the development, implementation and testing of a new automated approach based on a consensus score. The approach

relies on a pre-trained artificial neural network that performs stance detection. The stance detection system was incorporated into a web browser extension designed to make fact checking more efficient and convenient for news readers. The extension was then evaluated by a qualitative assessment focused on user observation. The results showed that users were more inclined to question the validity of new information and cross-examined assertive statements.

The penultimate chapter presents a conclusion to the project as a whole, accompanied by numerous examples of possible future work. The conclusion leads into the final chapter which acts as a user guide for the developed web browser extension.

Chapter 2

Background

2.1 The Problem

At present time, news can be generated and spread across the world almost instantaneously. Every person who has access to the internet has the ability to write stories and share information. However, this information may not be accurate. The author may have written it with some kind of bias or intent. A good example of this would be: "The godfather of fake news" [18]. Since new information arguably affects all decisions people make, the presence of fake news can have major consequences on politics, economics, health-care, sociocultural interactions and society in general.

Previously, people and news organisations built reputations for being trustworthy by reliably publishing accurate information and giving evidence of the facts they present by referencing sources. The problem nowadays is that, with the rise of the internet and social media, people can become news sources with access to a much larger reader base with no pressure to reference any sources or to be accurate. The lack of pressure comes from the fact that people on the internet can remain anonymous and hence have no reason to care about their reputation or the repercussions of their actions.

There have been times where advances in technology have led to brief periods where facts were verifiable, such as the invention of cameras and camcorders resulting in pictures and videos. The situation is currently at a turning point as further advances in technology have meant that images and even videos (see [19]) can now be modified as well as generated. This has led to a phase of distrust, where politicians (and the population in general) can get away with tactics such as "lying press" - casting doubt upon legitimate news from an opposing standpoint (Wikipedia [20]).

Separating correct and incorrect information is crucial for decision making and therefore progression, though this is no simple task. Determining the accuracy of information presented by a source is difficult because to some degree it will always involve trust. Trust is a human concept; there are no physical laws defining the rules of trust and as a result conflicts, and contradictions are common. Trustworthy reputations are built by experience, but what happens when there is no previous experience? Even experience is not a foolproof method of establishing trust. It is likely that a perfect method to determine trustworthiness does not exist.

In summary, people rely on new information for decision making, and inaccurate information can have dire consequences, for example "Like. Share. Kill." [10]. The rise of the internet and social media has allowed and even incentivised (see [21]) people to efficiently create and spread fake news. Determining whether information is correct or not is a difficult problem. This is because, unless the event was witnessed in person, it is essentially the same issue as finding out if you can trust a source or not (be the source a news organisation, a journalist, a video or even a person who claims to be a witness). It is impossible to stop the generation of false information but efforts can be made to contain its spread and make the public aware of the truth.

2.2 Current Solutions

2.2.1 Human Fact Checking

Currently the best solution to inhibiting the spread of false information is human review. The majority (if not all) of the major social media platforms (such as Facebook, Twitter, Instagram, Google+ etc.) and fact-checking services attempting to combat the spread of false information use human review to flag inaccurate content.

Facebook is a good example of a social media platform that utilises the accuracy of human fact checking. As a result, Facebook is currently doing better at hindering the spread of inaccurate information compared to other social-media platforms such as Twitter (see [22]). Facebook's solution has been to partner with independent third-party fact-checking services such as Full Fact and Factcheck.org. Facebook currently works with 48 of these third party fact-checkers in 23 different countries [23]. Their process for finding and rating false news is as follows: Facebook users can flag content (stories, images and videos) if they suspect it of being false. This content is then reviewed by one or more of the third-party fact-checkers and classified into one of nine categories including true, false or a mixture of accurate and inaccurate [24]. If content is rated as "false", "mixture", or "false headline" then its distribution will be reduced and it will appear lower down the news feed accompanied by related articles from the fact-checkers [23].

This process gives accurate results but suffers from the same problems that face all human fact-checking methods. In a blogpost Full Fact admitted that, "Fact-checking is slow, careful, pretty unglamorous work and realistically [they] know [they] can't possibly review all the potentially false claims that appear on Facebook every day" [25]. Facebook itself has commented on the limits of its human fact-checking process stating that "Fact-checkers don't exist in all countries, and different places have different standards of journalism as well as varying levels of press freedom. Even where fact-checking organizations do exist, there aren't enough to review all potentially false claims online. It can take hours or even days to review a single claim. And most false claims aren't limited to one article — they spread to other sites." [26]. Facebook has acknowledged that they must look to automated techniques (such as machine learning) to truly tackle the problem of fake news (see [26] and [27]).

Overall, human review is accurate but has many critical disadvantages. It is incredibly labour intensive, expensive and too slow to keep up with the creation and spread of false information.

2.2.2 Automated Approaches

Due to the disadvantages of human review and fact-checking, much work has been done to find an automated approach to inhibit the spread of false information. There have been many proposed solutions to this problem, the majority of which exploit natural language processing, probabilistic models or the recent success of machine learning. The aim is to find a method that can be implemented by computers which reliably and efficiently discerns between correct and incorrect information.

FABULA AI - FACT NOT FAKE developed and patented such a method which they termed geometric deep learning. Fabula uses geometric deep learning to predict fake news not from an articles content, but from its spread pattern over social networks [28]. This is based on the idea that people respond to fake news and real news differently. Studies have shown that people are in fact more likely to share fake news articles, causing them to spread faster than real news articles [13]. Using spread patterns to predict fake news means that this method is language independent and is well suited to social networks. The initial model proved its ability to quickly and somewhat accurately spot fake news (see [28]). However, in order to predict fake news via an article's spread pattern the article must spread across a social network, meaning that thousands if not millions of people may have already been influenced by the article's content. If the article is indeed fake, the damage may already have been done.

Google published a paper detailing a method of scoring webpages based on the correctness of the factual information presented. Correctness of the source data and the trustworthiness of the source is estimated using a sophisticated probabilistic model [29]. The paper defines the trustworthiness of a web source as "the probability that it contains the correct value for a fact (such as Barack Obama's nationality), assuming that it mentions any value for that fact" [29]. The system works by extracting many facts from webpages using common information extraction techniques. The extraction of facts is error prone leading to two types of inaccuracy: incorrect fact extraction (the fact extraction results in the error and the fact itself may be correct) and incorrect fact (the fact is correctly extracted by the extraction technique but is itself inaccurate). An inference probabilistic model is then used to jointly estimate the correctness of these facts and the accuracy of the sources by considering facts as true if they are present in the Freebase Knowledge base [30] and extractions as error free if they pass type checking. This method is a scalable and efficient way to estimate the correctness of factual information. The trade-off is that the fact extraction process does not acquire all the facts and there is no perfect way to deal with the errors and incomplete information introduced. The fact checking also uses the Freebase knowledge base which means that unless this knowledge base has been recently updated, new accurate facts may be classed as inaccurate.

There are numerous tools that utilise AI and machine learning to give news articles a validity rating, for example FakeBox [31] and MIT CSAIL's AI [32]. General machine learning models such as logistic regression, naive Bayes classifiers, support vector machines and many others have also been used to classify news as fact or fake (see [33] [34] [35]). These tools/models can achieve very high validation accuracies with some being as high as 95% [36].

The problem with using machine learning to detect fake news is that it requires a large data set to learn from. Fake news data sets exist and are relatively easy to find (see [8] and [37]). The issue is that these data sets are small and the algorithms can only learn features that are present within the training data. Therefore, they learn to classify fake news based on features such as the identity of the source, the style in which the article is written and the vocabulary used in the article's title/text. These features give clues to the validity of the content, but, inferring an articles accuracy base simply on vocabulary and writing style has several drawbacks. Firstly, the models accuracy would decrease with time as new news sources emerge with different writing styles to the ones present in the training data set. Secondly, these models would struggle to differentiate between real and false news. Since by definition of false news, the author, who may usually be trustworthy, has unknowingly published incorrect fact. So, their writing style and vocabulary would not have changed when compared to a real news article they may have written. Finally these models would struggle to classify text that is too short, which is often the case when news is spread on social media.

To summarise, automated methods have made progress when it comes to determining between correct and incorrect information. Many of the current approaches use machine learning which are limited by the data sets they are trained on. An ideal automated technique would be able to reliably discerned between fact and fake/false information while not being subject to changes in the authors writing style and vocabulary.

2.3 Proposed Solutions

The following background on sentiment analysis and stance detection is given, as these were the techniques pursued when attempting to automate the process of discerning between real and fake news. The sentiment analysis section focuses on research into existing lexicons that can be used to quantify the overall sentiment of an article. The section on stance detection is centralised on a multilayer perceptron that predicts the stance an article takes with respect to a given statement. Stance detection, as discussed in section 6, is a method used to quantify consensus.

2.3.1 Sentiment Analysis

Sentiment analysis or opinion mining is the analysis of peoples opinions, sentiments and emotions towards a certain subject or entity [38]. There are many modern techniques to performing sentiment analysis such as classification using supervised and unsupervised learning algorithms and the use of certain rule based methods.

Within a document there are many ways to express positive and negative sentiments. The most common of which is the use of sentiment words or opinion words [38]. Sentiment words are divided into two categories, positive sentiment words and negative sentiment words. For example "good", "great" and "fantastic" are positive sentiment words while "poor", "bad" and "terrible" are negative sentiment words [38]. These sentiment words along with other features, such as certain phrases, idioms, acronyms, initialisms and emoticons make up a sentiment lexicon. A sentiment lexicon is essential for sentiment analysis as the vocabulary used is an important indicator of a documents sentiment [38]. Manually constructing and validating a comprehensive sentiment lexicon is a labour intensive and time consuming process. Fortunately there are existing lexicons and rule based models that have been extensively validated.

The Linguistic Inquiry and Word Count (LIWC) software uses a human curated and validated lexicon with a "master dictionary composed of almost 6,400 words, word stems, and selected emoticons" [39]. Of these 6,400 words there are 620 words that are labelled as positive emotion and 744 words labelled as negative emotion, meaning that the overall lexicon that is relevant to our analysis contains 1,364 features. LWIC has undergone a validation process that has spanned more than a decade [1] and is considered to be one of the most reliable lexicons for extracting sentiment polarity from text.

There are many other sentiment lexicons such as the General Inquirer (GI) [40], Hu-Liu04 [41], the Affective Norms for English Words ANEW [42], The Semantic Orientation CALculator (SO-CAL) [43], SentiWordNet [44] and SenticNet [45]. Each of these lexicons have their relative advantages and disadvantages. Table 2.1 gives a brief overview of all the existing sentiment lexicons that were considered for the sentiment analysis of news.

A lexicon that is of particular interest to the sentiment analysis of news and social media is used in the Valence Aware Dictionary and sEntiment Reasoner (VADER) package [1]. VADER is a lexicon and rule-based sentiment analysis model aimed at the analysis of social media text [1]. VADER provides a gold-standard list of 7500+ lexical features with corresponding sentiment intensity measures [1]. These lexical features were constructed by examining existing human validated sentiment lexicons (LIWC, ANEW, GI), adding additional lexical features commonly used in social media and empirically calculating a sentiment intensity for each feature. The resulting list of features then underwent a rigorous process of cleaning and validation. This sentiment analysis model also incorporates five general rules used to modify the sentiments of each feature present in a sentence. For full details of the construction and validation of the lexicon used in the VADER sentiment analysis model see figure 2.1.

	Lexicon overview	Number of sentiment relevant features
LIWC	Binary (words are positive or negative). There are 620 words labelled as positive and 744 words labelled as negative.	1,364
GI	Binary (words are positive or negative). There are 1,915 words labelled as positive and 2,291 words labelled as negative.	4,206
Hu-Liu04	Binary (words are positive or negative). There are 2,006 words labelled as positive and 4,783 words labelled as negative.	6,800
ANEW	Words are associated with valence scores for sentiment intensity. Sentiment valence ranges on continuous scale from 1 to 9. Neutral valence score is 5. Words with valence scores less than 5 are considered negative. Words with valence scores greater than 5 are considered positive.	1,034
SO-CAL	Words are associated with valence scores for sentiment intensity. Sentiment valence for each word is an integer between -5 (extremely negative) and 5 (extremely positive). There are no objective words in the lexicon.	5,000
SentiWordNet	Words are associated with valence scores for sentiment intensity. Synsets are annotated with three numerical scores relating to positivity, negativity, and objectivity. Difference of each sysnset's positive and negative scores can be used as its sentiment valence [44].	38,000+
SenticNet	Words are associated with valence scores for sentiment intensity. Sentiment valence ranges on continuous scale from –1 (extremely negative) to 1 (extremely positive). Neutral valence score is 0. Words with valence scores less than 0 are considered negative. Words with valence scores greater than 0 are considered positive.	14,244
VADER	Words are associated with valence scores for sentiment intensity. Sentiment valence ranges on continuous scale from –1 (extremely negative) to 1 (extremely positive). Neutral valence score is 0. Words with valence scores less than 0 are considered negative.Words with valence scores greater than 0 are considered positive.	7,500+

Table 2.1: A summary of existing sentiment lexicons.

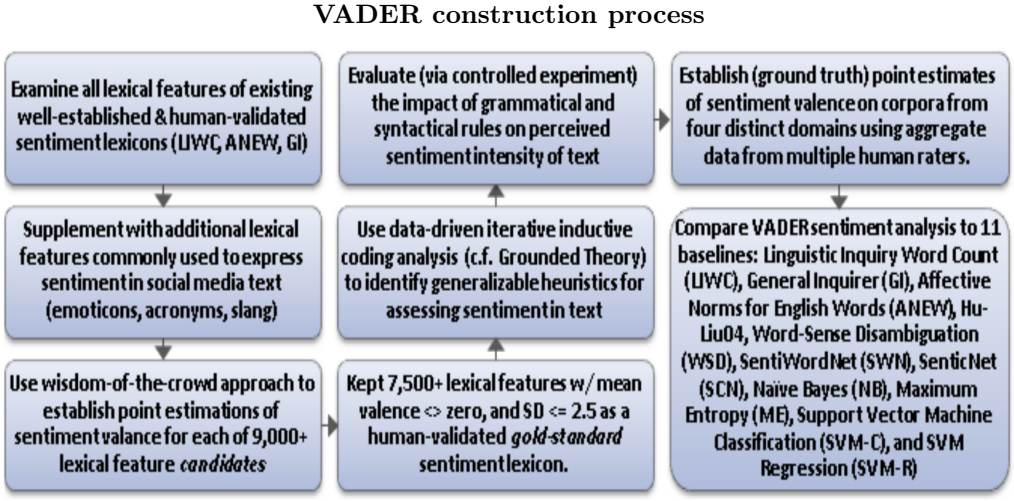

Figure 2.1: 'Methods and process approach overview.' (Hutto and Gilbert, 2014 [1])

Figure 2.2: 'Example of the interface implemented for acquiring valid point estimates of sentiment valence (intensity) for each context-free candidate feature comprising the VADER sentiment lexicon.' (Hutto and Gilbert, 2014 [1])

2.3.2 Stance Detection

Stance detection is the classification of a texts perspective, or *stance*, towards a certain target. For instance, the statement: "Crypto-currency value is rising" may be the target with an article taking a stance that "agrees", "disagrees" or is "neutral" with respect to the statement. Stance detection represents a more complex problem than sentiment analysis but is a necessary step to achieving automated cross checking.

The vast majority of stance detection systems rely on machine learning. As discussed in automated approaches (subsection 2.2.2), directly applying machine learning to the classification of news articles is severely limited by the data sets. However, this is not the case when applying machine learning to stance detection. This is because there are several mainstream data sets that can be used to train stance detection models, see [46] and [47]. These data sets are professionally labelled and relatively large. Furthermore, *The Fake News Challenge* [2] provides a data set specifically tailored to performing stance detection in order to combat fake news.

The Fake News Challenge data set was derived from *Emergent* [48], a data set which "contains

300 rumoured claims and 2,595 associated news articles, collected and labelled by journalists with an estimation of their veracity". The training set used for *The Fake News Challenge* consisted of a headline and article, paired with the underlying stance the article takes with respect to the headline. The possible stances are "agree", "disagree", "discuss" and "unrelated", This data set provides 49972 training examples with an approximate split of 73% unrelated, 18% discuss, 7% agree and 2% disagree.

The Fake News Challenge [2] was put forward as a competition to explore artificially intelligent methods of detecting fake news. As of June 2019 the competition is still ongoing, though the first stage (stance detection) has been completed. Stance detection in the context of the *The Fake News Challenge* requires an AI model to classify whether the body of an article agrees, disagrees, discusses or is unrelated to its headline or the headline of another article. The competing AI models were evaluated based on a weighted two-tier scoring system as illustrated by figure 2.3. The winning team achieved a relative score of 82.02% [2].

The chosen pre-trained stance detection model was developed by UCL Machine Reading [3]. Their system finished third in the competition scoring 81.72% (only 0.3% behind first place). Their single, end-to-end stance detection system consisted of lexical and similarity features passed through a multi-layer perceptron with one hidden layer [3], as shown in figure 2.4. The reasons for choosing this particular system are discussed in section 6.

***The Fake News Challenge* stance detection scoring process**

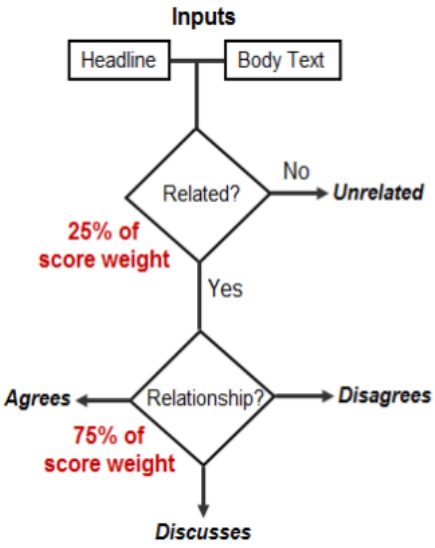

Figure 2.3: "Scoring process schematic" to evaluate competing models in *The Fake News Challenge* [2].

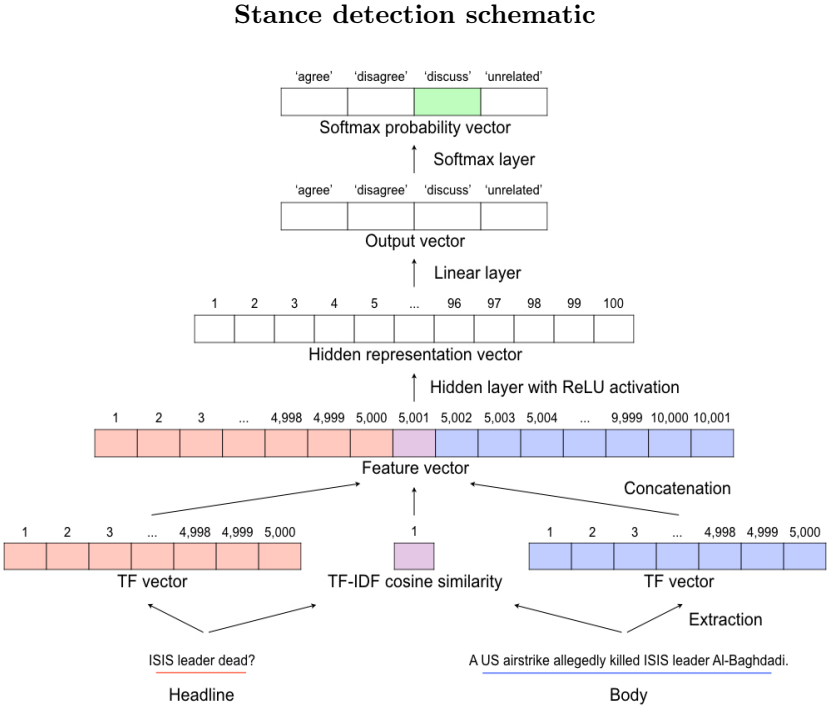

Figure 2.4: "Schematic diagram of UCLMR's system" (B. Riede et al, 2018 [3]).

Chapter 3

Project Requirements

The aim of this project was to develop a tool to aid the average news reader in determining the legitimacy of a news piece. The motivation for developing such a tool comes from the fact that generating and spreading fake news requires very little effort. while its impact on society can be enormous. The following gives a detailed list of requirements used to guide the overall research and development of the tool.

The list of necessary requirements was obtained via interviews with professional journalists, as well as regular news readers. These requirements are essential and must be fulfilled in order for the project to be regarded as a success. The list of desirable requirements was constructed from intuition and research. Achieving these requirements is optional, though each one represents a significant improvement to the value of the project.

Necessary requirements:

1. *Aid in the classification of a news article*

   A news article can be classified as real, persuasive, false or fake. Knowledge of an articles class gives valuable information on the articles accuracy and opinion. Therefore, developing a tool that informs a news reader of the articles class is a direct method of impeding the negative impacts of misinformation. Furthermore, simply drawing attention to a news articles class would lead to an increase in reader awareness. Improved awareness enhances the likelihood a reader recognises and questions potentially incorrect information.

   Consequently, the projects objective can be seen as developing a method of analysing news so that an article or statement can be correctly classified. The tool should generate a score in a specific news metric. This score should give an indication of the class the news article belongs to.

2. *Convenient*

   The reason news readers disregard fact checking is because the process is time consuming and disrupting. When reading a news article or scrolling though a news feed one may come across numerous claims. It is infeasible to manually review every claim, as a rigorous fact checking process can take several hours. The core problem is the inconvenience of fact checking. Therefore, the developed tool must be convenient for users. Reviewing a statement or news article should take no longer than a few minutes, while allowing the reader to continue reading/scrolling. A tool that is convenient and unobstructive is more likely to be employed by users and consequently reduce the spread of incorrect information.

3. *Intuitive*

   Since the developed tool aims to aid the average news reader, it should be intuitive and easy to operate. New users should be able to use the tool with very little help and guidance. Having a small learning curve adds to the user experience by emphasising the convenience and unintrusive aspects of the tool.

Desirable requirements:

1. *Transparent*

   A transparent tool would be simple to understand. That is, users would be able to comprehend the tools inner workings. Knowledge of how the tool works, builds confidence and trust in the derived result. If a user does not trust the result provided, then they will be forced to perform their own investigation into the legitimacy of the information, rendering the tool purposeless and impractical.

2. *Accessible*

   The developed software should be accessible and readily available. Downloading and installing the tool should be a simple process, requiring little effort from the user. This would ensure that the set up process would not act as a barrier to utilising the product.

Chapter 4

News Metrics

4.1 Metric Investigation

The following features of news (or news sources) have been identified as possible metrics in which news can be measured. See the definitions section for exact meaning of a metric in the context of this project and table 4.1 for a brief summary of each metric.

Accuracy: Officially, accuracy is defined as "being exact or correct" [49]. In terms of news, accuracy will be defined as a measure of the information's correctness or factuality, that is, how close the information is to the ground truth.

Neutrality: Neutrality in the context of this project is a measure of opinion. Being neutral is not taking sides or offering judgement on the different points of view. Therefore the information in neutral news must be free of the writers opinions and personal beliefs. For news to be neutral the author should only present the facts and leave the interpretation of these to the audience.

Impartiality: Officially, Impartiality is defined as "Equal treatment of all rivals or disputants" [50]. For the purpose of this project impartiality will be defined as a measure of the treatment of the different points of view. A news piece that is impartial treats all the presented points of view in an equal manner, showing no favouritism. Impartial diverges from the definition of neutral, as opinions can be given, as long as they are equal on both sides. Consequently impartiality is a measure of opinion balance.

Reliability: Reliability will be a measure of how likely it is that a news source will publish a certain type of news. The easier it is to predict what class of news a news source will produce, the more reliable the source is. A reliable news source publishes many news articles that belong to the same class. Unreliable news sources would publish news of many different classes.

For example:

If a news organisation publishes many articles that would be considered real news and very few articles that would be classified as any other type of news, then the organisation can be considered to reliably produce real news. In this case the organisation has a high reliability.

An organisation that reliably produces fake news publishes many fake articles and few articles that can be classed as other news types. This organisation also has a high reliability as it is easy to predict what class of news will be produced.

Consensus: Consensus is "general agreement" [51]. Therefore consensus as a metric will be a measure of how many separate news sources agree with each other. If many news

outlets have spread the same information, then this information has a high level of consensus.

Sentiment: Sentiment is a measure of the general feeling/emotion of a news piece. It can be seen as the overall positivity or negativity surrounding the subject of the article.

Detachment: Detachment is a measure of the journalists emotional approach to news. A detached journalists selects and write stories with a dispassionate and emotionless attitude so that the news piece is presented in a calm and rational manner [52].

Journalistic Objectivity: Journalistic objectivity is a combination of many metrics, the main ones being accuracy, neutrality and detachment [52]. The aim of objectivity in journalism is to encouraged its audience to come up with their own opinions based on neutral and correct information. News that is objective would be classified as real news.

Bias: Bias is favoritism shown to a particular point of view. A bias news article means that the journalist has taken a side; The news piece contains many of the author's personal opinions and judgements, the majority of which support a certain perspective. A non-bias news piece is either neutral (contains no/little opinions) or impartial (contains opinions for/against both sides in equal proportion).

Metric:	**Measures:**
Accuracy	Factuality/truthfulness of a news piece
Neutrality	Quantity and extremity of opinions within a news piece
Impartiality	Overall opinion balance of news
Reliability	Predictability of a news source
Consensus	Agreement with other news sources
Sentiment	Overall positivity or negativity of a news piece
Detachment	Journalists' emotional approach to selecting and writing news stories
Journalistic objectivity	Accuracy and neutrality of news piece
Bias	Quantity and extremity of opinions in favour of a certain point of view in news piece

Table 4.1: Summary of each metric and what it measures.

4.2 Metric selection

The aim of this project, as described in section 3, was to develop a method/tool to analyse news in a specific metric so that a news article can be classified as real, persuasive, false or fake. Since real news contains correct information and few opinions the two metrics most suited to finding real news would be accuracy and neutrality.

Measuring the accuracy of news is a difficult problem and often the only solution is to infer the accuracy of the information by looking at how the news piece performs in another metric. For example, human fact-checkers look to fact-bases and other news organisations to see if the information presented there agrees with the news piece they are investigating. This is not a direct measure of accuracy but is actually a measure of general agreement or consensus. The same is true for automated approaches, they infer accuracy from features such as the way news spreads (Fabula AI, [28]) or the vocabulary used. Overall, it would appear that a direct method to measure

accuracy does not exist and the best way to determine the truthfulness of a news piece is to infer its accuracy from a measure in some other metric, such as consensus.

Unlike accuracy, neutrality can be directly measured by analysing the article's content. This is because neutrality, as with impartiality, sentiment and bias, is fundamentally a measure of opinions. Opinion mining or sentiment analysis is an entire area of research in itself and represents a huge problem space [38]. To illustrate the problems associated with sentiment analysis a sentence is given below with a few examples of the factors that must be taken into consideration.

"Fred thinks the car is not bad but personally I've discovered that in extremely bad weather conditions, putting your foot flat on the brakes won't change your speed!"

- There may be multiple opinion holders, for instance the author and Fred.
- Each of the view holders opinions may be the same, different, completely unrelated or a varying degree of each.
- Within a single document opinions can be expressed about multiple entities (the car and the weather) and multiple features of each entity (the car's brakes).
- There are subjective opinions such as "the car is not bad" and objective statements that imply an opinion, for instance "putting your foot flat on the beaks won't change your speed".
- Opinions can be expressed with varying levels of intensity. Certain words and punctuation, such as "extremely" and "!", alter this intensity.
- Some words can signify a shift in opinion, examples would include the words "but", "however" and "although".
- Certain words can completely flip the opinion polarity as demonstrated by "the car is not bad" being very different from the "The car is bad".

For a more detailed exploration of the problems and solutions in the field of opinion mining see Bing Liu 2012 [38].

All these components make opinion identification and analysis very complex. To give an accurate neutrality measure, an algorithm would have to be able to distinguish between opinions given by multiple people on multiple subjects. It would also have to be able to measure the opinion intensity and polarity towards each of the subjects. This is assuming the opinions have been correctly identified in the first place, which is no trivial task in itself. However when dealing with the classification of news, certain simplifying assumptions can be made.

1. The subject of the opinion is irrelevant. No matter what the subject is, real news would give no, or very little, opinion on it.
2. Opinions have a corresponding sentiment. For example "These political policies would benefit everyone" expresses a positive sentiment about the political policies.

These assumptions are not perfect but they do hold true for the majority of cases and allows sentiment analysis to be preformed on news articles. This situation is similar to case of accuracy. It was previously discussed that accuracy can be determined by measures in other metrics such as consensus. Here, a neutrality measure can be inferred from a sentiment score. Furthermore, although the definition of fake news does not take opinions/bias into consideration (it is simply defined as intentionally incorrect information) a hypothesis relating the opinions in real and fake news can be made and consequently tested.

Hypothesis 1: **Real news contains fewer and less extreme opinions than fake news.**

This hypothesis is derived from the definition of real news. Real news is defined as being neutral; it contains none or very little of the author's opinions. On the other hand, opinions often make a news story more interesting. Since a fake news article's success is determined by how many people read/share it, it can be hypothesised that fake news contains opinions in larger quantity and extremity than real news. The basis of this hypothesis is that news that is abnormally interesting or outrageous can be presumed to attract more readers and spread further.

If the aforementioned assumptions and hypothesis hold, then developing a tool that can correctly identify a news articles sentiment would be an important step towards the classification of a news piece. This is because a sentiment measure can be used to discriminate between real and persuasive news, as well as false and fake news (see figure 4.1). Therefore, sentiment was the metric pursued for the first part of this project.

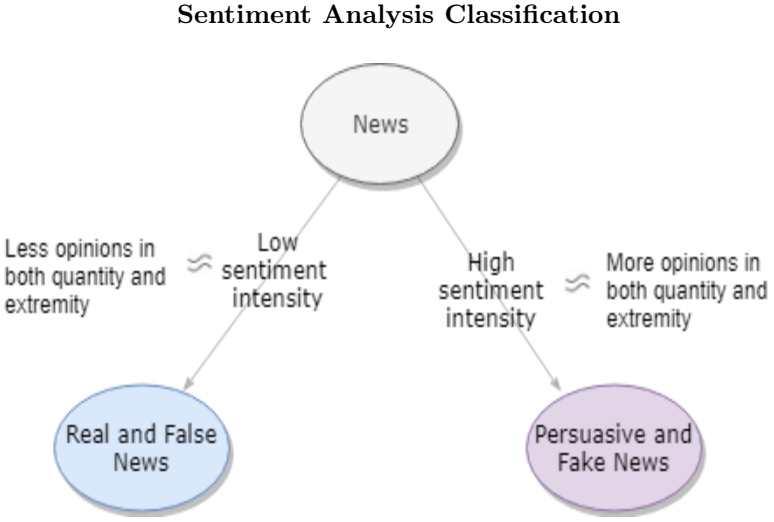

Figure 4.1: Map to show how sentiment analysis would be able discriminate between each news category if hypothesis 1 holds.

Chapter 5

Sentiment Analysis

5.1 Design

The overall design of the sentiment analysis tool is best described by the flowchart presented in figure 5.2. This flowchart illustrates the steps taken to derive a sentiment score for a given text in a way that was convenient and intuitive for users. The following details the engineering decisions made when designing this sentiment analysis tool.

There are many different ways to preform sentiment analysis and each one comes with its relative advantages and disadvantages. One thing all methods have in common is that they all involve a sentiment lexicon. Manually constructing a reliable sentiment lexicon is a labour intensive and time consuming process (see Bing Liu 2012, [38]). The other two options include constructing a sentiment lexicon using machine learning algorithms such as naive Bayes, maximum entropy and support vector machines or finding a pre-existing human curated sentiment lexicon.

Using machine learning to learn the relevant sentiment features of a language comes with certain disadvantages compared to human curated sentiment lexicons and rule based models. The main disadvantage is the need to have a large data set for training and testing. Having a data set that contains many common sentiment-relevant features is fundamental to the quality of the lexicon produced. This is because machine learning algorithms can only learn features if they are represented in the data they are trained on. Other disadvantages include longer computation times and the production of features that are often domain specific and difficult for humans to interpret [1]. For these reasons it was decided to search for an existing sentiment lexicon that was constructed and validated by humans.

Eight sentiment lexicons were considered (see section 2.3). Out of these eight lexicons VADER appeared to be the one most suited to the sentiment analysis of news. Not only is VADER specifically attuned to the analysis of social media text but it also seems to generalise more favourably than most of the competing models [1]. It outperformed all of the eleven state of the art benchmark lexicons it was compared to and even did better than individual humans at rating the sentiment of tweets [1]. For these reasons the sentiment analysis of news articles was carried out using the VADER lexicon and rule-based system.

The VADER sentiment analysis model is readily available as a Python package. Other useful Python packages include Beautiful Soup, Requests and Google-Search. These libraries were essential in automating internet searches and extracting data from websites. Scraping articles from news agency websites was a key component in the testing and evaluation of the sentiment analysis tool. Furthermore, news pieces are natural language documents. Therefore, they can be analysed by the Python Natural Language Toolkit (nltk), which was specifically developed to process human literature. For these reasons it was decided to write the tool using the Python programming language.

It was concluded that users should be able to access the tool while browsing the internet. This was because the tool is most useful when reading an online article or scrolling though a social media news feed. There were several options when it came to providing quick and easy access to the tool. This included installing software or publishing a website where users can copy and paste text to be analysed. However, the most convenient system would simply require the user to highlight the text they wish to check, right click the highlighted text, and select analyse from a dropdown menu. This was achieved by creating a browser extension.

Browser extensions are software applications that add functionality to a web browser such as Chrome, Firefox or Safari. In order to ensure maximum availability of the browser extension, research was conducted to find the web browser with the largest market share. From figure 5.1 it is clear that, as of May 2019, the most popular web browser is Chrome with 57.4% market share. This statistic was confirmed by "StatCounter [5]" and "NetMarketShare [6]" and shows Chrome to have over four times the market share of its closest competitor (Safari). Therefore the extension was developed for the Chrome web browser.

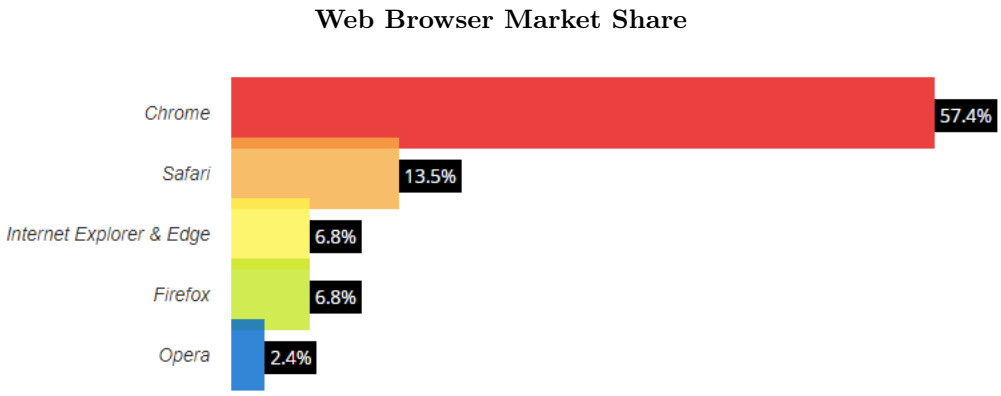

Figure 5.1: Top five web browsers by market share as of May 2019 according to W3Counter: Global stats [4]. This statistic was confirmed by "StatCounter [5]" and "NetMarketShare [6]".

The decision to implement a browser extension led to the following design: A user highlights the text they wish to be analysed. They then right-click the highlighted text causing a context menu to appear. The context menu contains an option to "Analyse Text". If this option is selected, the text is sent from the web page to a local Python script. This script first tokenizes the text into sentences. Then, the VADER sentiment analysis model is applied to each sentence, producing an array of sentiment scores (one for each sentence). The average sentiment, average absolute sentiment, maximum sentiment, minimum sentiment and sentiment variance are then computed. The derived scores, not including the variance, are then multiplied by four to make the results interpretable by the Likert scale shown in figure 2.2. These results are sent back to the web browser where a notification appears to alert the user of the computed scores. A detailed breakdown of the results can then be accessed via a popup that loads when the extensions icon is selected from the web browser toolbar.

When implementing the design, it was concluded that the average absolute sentiment score would give the best indication of a news pieces class. Hence this score should have been used to classify a news piece as real/false or persuasive/fake, as per figure 4.1. However, when applying the tool to news data sets it became apparent that an acceptable decision boundary between these two news groups could not be found. Therefore a direct classification of the news piece was not implemented.

To summarize the design process, six engineering decisions were made. Firstly, it became clear that sentiment lexicons are necessary to perform sentiment analysis, but manually constructing and validating a sentiment lexicon was too time consuming. The remaining options were to compute a sentiment lexicon via machine learning or to find a pre-existing sentiment lexicon. Using machine learning to derive a sentiment lexicon had several critical disadvantages. Therefore research into human curated sentiment lexicons was performed, of which, VADER emerged as the lexicon best suited to the sentiment analysis of news and social media text. VADER was available as a

Python package, which, along with access to popular web scraping libraries, led to the tool being implemented in Python. To ensure maximum availability and convenience for news readers, it was decided to create a Chrome web browser extension. The extensions design is illustrated by figure 5.2. All elements of the design were implemented except for direct text classification. Classification was infeasible as both news groups produced remarkably similar sentiment scores, see section 5.5 for further details.

Sentiment Analysis Flow Diagram

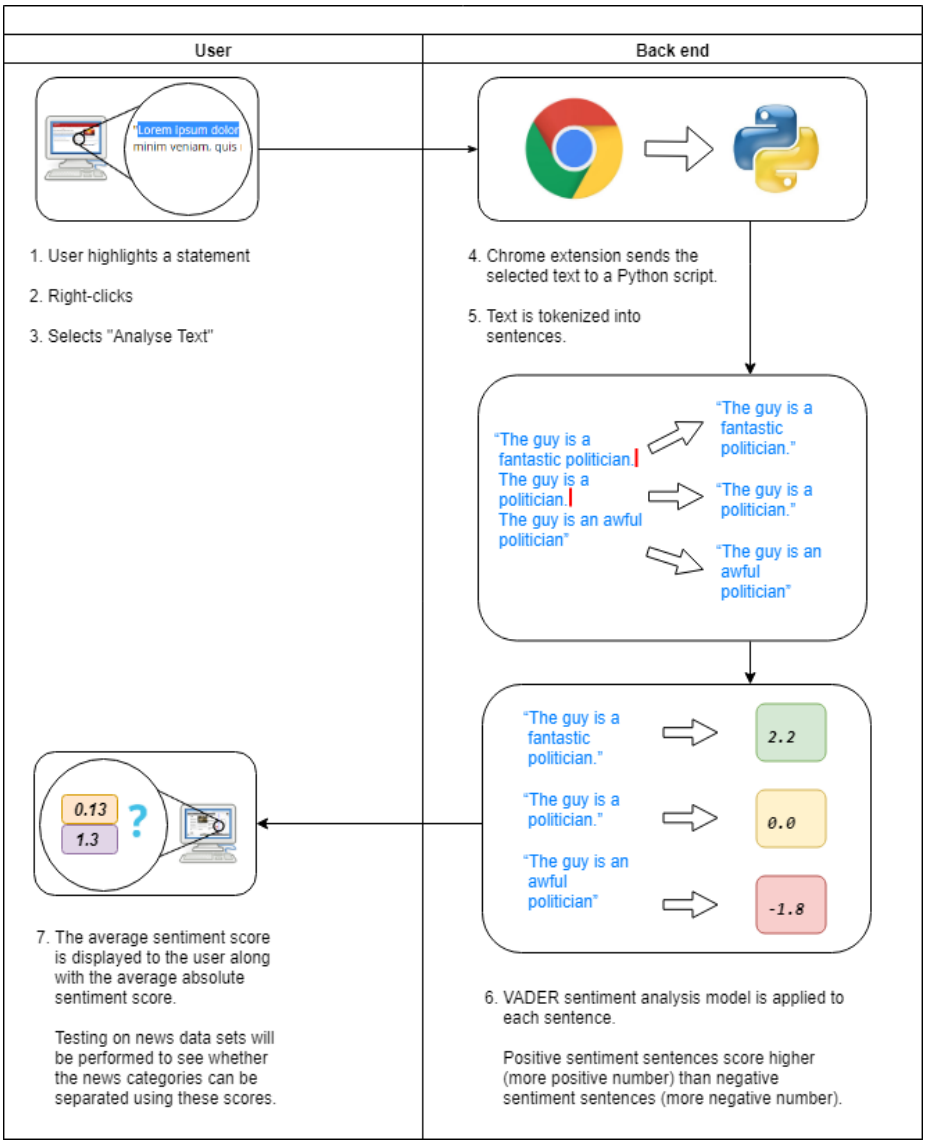

Figure 5.2: Flowchart illustrating the sentiment analysis procedure. Initially the user highlights a piece of text. Upon right-clicking the selected text a context menu appears giving the option to "Analyse Text". If this option is chosen the browser extension sends the selected text to a Python script which uses the Python Natural Language Toolkit (nltk) to tokenize the text into sentences. The VADER sentiment analysis model is then applied to each sentence to obtain a sentiment score. The sentiment score is multiplied by four so that it can be interpreted using the VADER Likert scale, as shown in figure 2.2. The average sentiment score and the average absolute sentiment score are then displayed to the user in a popup along with other sentiment scores that may be of interest (see figure 5.5). Testing on news data sets will determine whether the two news categories, real/false and persuasive/fake, can be separated using these scores (see figure 4.1). If this is possible then the predicted news category will also be displayed to the user.

5.2 Implementation

The sentiment analysis tool can be divided into two parts; a web browser extension and a local host. The browser extension was coded using a mixture of web technologies such as JSON, HTML and JavaScript. The local host was mainly written in Python, though it made use of both JSON and Batch scripts. Overall, eight program scripts were necessary for the tools implementation. These scripts are shown in figure 5.3. All source code can be found at: https://github.com/OrionMat/Sentiment-Analysis [53].

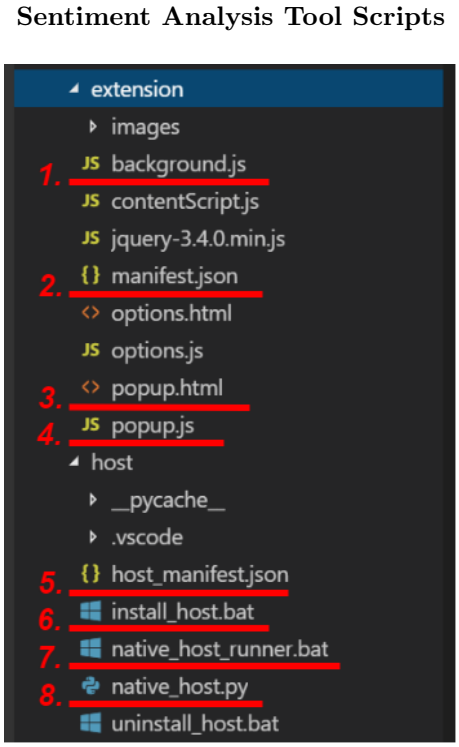

Figure 5.3: The program scripts necessary for the sentiment analysis tool.

The web extension and local host communicate to each other via native messaging. Native messaging allows the exchange of messages via standard input (`stdin`) and standard output (`stdout`) streams. Native messaging requires each message to be serialized with JSON and encoded using the UTF-8 variable width character encoding scheme. A further requirement is that the local host must be registered with Chrome.

Registration of the native messaging host involved installing a manifest file that defined the hosts configuration. This file was named *"host_manifest.json"* (file 5 in figure 5.3). Installing *"host_manifest.json"* required a registry key with value set to the *"host_manifest.json"* path. This was achieved by running the *"install_host.bat"* file which creates this registration key and sets its value in a single command.

Once the *"host_manifest.json"* file had been registered, the web extension was able to establish a native messaging port with the local Python script *"native_host.py"*. This occurs because the *"host_manifest.json"* file contains an element called "path" with value set to the path of *"native_host_runner.bat"*. *"native_host_runner.bat"* acts as the native hosts binary, hence starting *"native_host.py"* in a separate process. *"native_host.py"* was then able to share information with the extension by sending and receiving messages. This process is illustrated in figure 5.4.

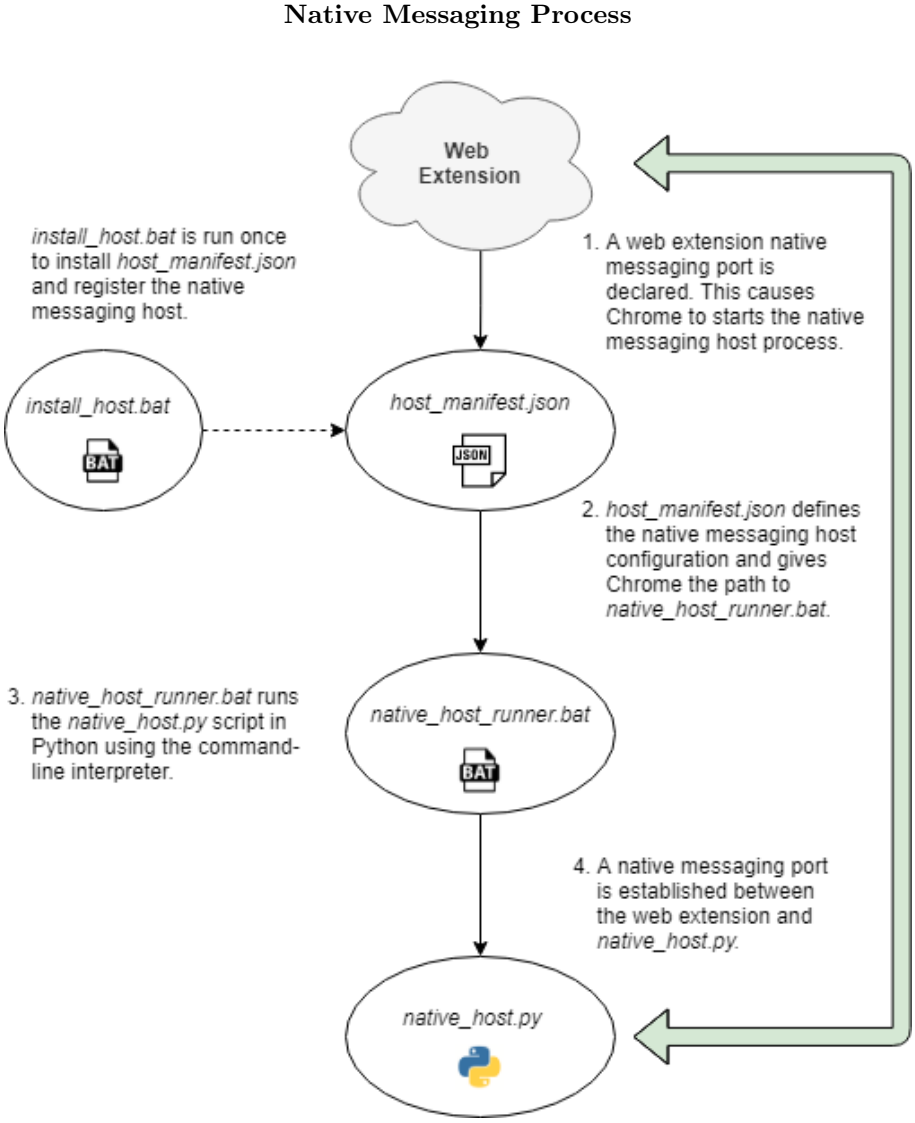

Figure 5.4: Flowchart illustrating the native messaging process to set up a port between the web extension and local Python script. The green arrow represents this native messaging port which allows messages to be transferred. *"install_host.bat"* is executed once and only once while all other files are involved every time the web extension declares a native messaging port.

5.2.1 Web Browser Extension

The web browser extension is responsible for sending user selected text to the local host and displaying the received sentiment scores. There are four scripts involved in this process. These are *"background.js"*, *"manifest.json"*, *"popup.html"* and *"popup.js"*. These scripts are labelled 1, 2, 3 and 4 in figure 5.3.

The *"background.js"* script is the extensions event handler. It remains idle until the context menu is opened and the user selects "Analyse Text" from the list of options. If this occurs the script acquires the highlighted text and saves it to Chrome storage. Then, the script opens a native messaging port to send the text to *"native_host.py"*. After sending the selected text to the native host the script awaits the sentiment analysis results. Upon receiving the results, the background script immediately saves them to Chrome storage. The script then creates a notification based on the average sentiment score which is displayed to the user.

"popup.html" and *"popup.js"* create and control the Chrome popup that appears when the user selects the extension's icon from the web browser's tool bar. The popup presents the selected text followed by the text's sentiment scores, as shown in figure 5.5. The selected text acts as the popups header while the sentiment scores are displayed as part of a table.

The header and table were defined in *"popup.html"*. The header and each element of the table were assigned a unique identification (ID) attribute. This unique ID allowed *"popup.js"* to update the header value and table elements using the jQuery ID selector. *"popup.js"* updates the popup by retrieving the previously saved data from Chrome storage and injecting it into *"popup.html"*.

The web extension made use of four Chrome APIs. These APIs were *contextMenus*, *storage*, *runtime* and *notifications*. *contextMenus* was used to add the option "Analyse Text" to the dropdown menu that appears when right clicking on highlighted text. The Chrome *storage* API was used to save and retrieve the selected text and sentiment analysis results. The *runtime* API contained the *connectNative* method which set up the native messaging port to *"native_host.py"*. Finally, the Chrome *notifications* API was employed to create a notification that alerts users of the average sentiment score as well as the completion of the analysis process.

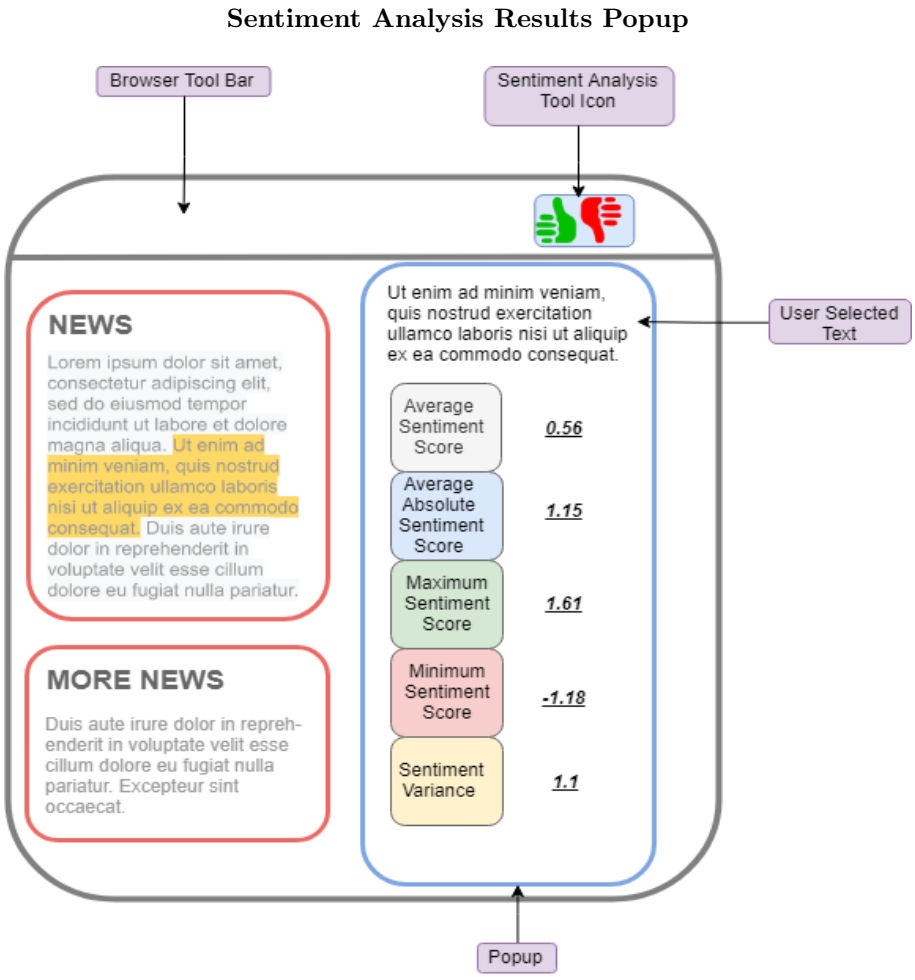

Figure 5.5: A diagram illustrating the results computed and displayed by the sentiment analysis tool. The results are presented in a popup that appears when the tool icon is selected. See appendix B.1 for a screen shot of the implemented tool.

All Chrome APIs, apart from *runtime*, acted on JavaScript objects. For example, manipulating the Chrome context menu to include the option "Analyse Text" involved declaring a JavaScript object with three fields, *id*, *title* and *contexts*. This object was then passed as a parameter to the Chrome context menu API. The *runtime* API simply took a sting parameter, which was the name of the native host file, that is *"native_host"*.

5.2.2 Local Host

The local host, *"native_host.py"*, performs sentiment analysis on the user selected text. This text is sent from the web extension as a native message. Analysing the text's sentiment involves pre-processing the text, applying the VADER sentiment analysis model to the manipulated data, then calculating the required statistics. The required sentiment analysis statistics include: average sentiment, average absolute sentiment, maximum sentiment, minimum sentiment and sentiment variance. These values are then sent back to the web extension.

As VADER rates the sentiment of individual words and sentences, the model can not be directly applied to the text received from the web extension. Therefore, the Python Natural Language Toolkit (nltk) was used to tokenize the raw text into sentences. After sentence tokenization VADER is applied to each sentence in turn, resulting in a list of sentiment intensity scores (one for each sentence). The sentiment scores range between -1 and 1. These scores are difficult to interpret. However, after multiplying the scores by four the results can be compared to Likert scale shown in figure figure 5.6.

To compute all the relevant sentiment scores the Numpy scientific computing package was utilised. Hence, the list of sentiment scores was converted to a Numpy array, which, allowed Numpy built in methods to be applied to the data. All sentiment scores were calculated using Numpy statistic methods. These included: `mean`, `absolute`, `var`, `max`, `min` and `round`.

Sentiment Analysis Local Host Flowchart

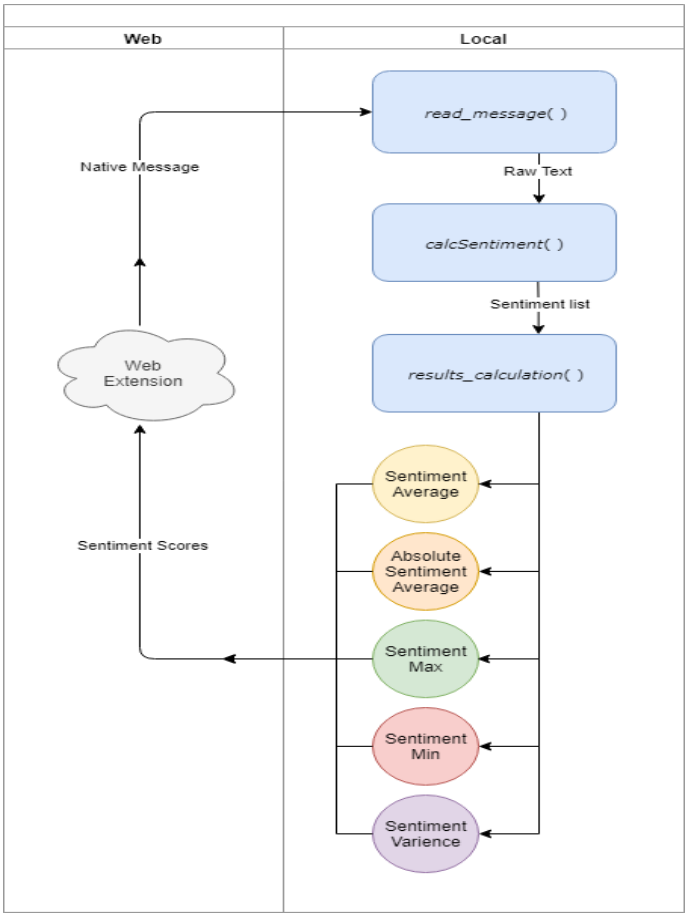

Figure 5.6: Flowchart showing the local host procedure when calculate sentiment scores.

It is not possible to describe every detail of the implementation. Hence, certain processes and functions have been out left. For instance the `send_message()` and `receive_message()` functions defined in *"native_ host.py"* which process the data into a form that can be transmitted by Chrome native messaging. A large amount code was also written for the testing and evaluation of the tool which has not been discussed in detail. As previously mentioned, all code can be found at: https://github.com/OrionMat/Sentiment-Analysis [53].

5.3 Testing

Testing forms a key component of the sentiment analysis tool. This is because not only is testing used to evaluate the tools success, but it also plays a critical role in determining the classification boundary between the different news groups. Obtaining an adequate classification boundary depends on two factors. First, the validity of the hypothesis that led to the tools development. If fake news does not contain opinions in significantly different quantity or extremity than real news, then the tools ability to classify news will be exceedingly limited. Secondly, the tool relies on the VADER sentiment analysis model. Hence, if the VADER model cannot accurately derive the sentiment of a news piece, then determining the classification boundary is not possible. Consequently, this section details two types of testing. One to see if the sentiment scores between real and fake news are distinguishable in a predictable manner. The other to see if the derived sentiment values match reality.

A preliminary test was performed to gain an initial impression of VADER's ability to accurately derive the sentiment of a news piece. This test made use of a BBC news data set [54], consisting of $2,225$ articles divided into five news categories (business, entertainment, politics, sport and technology). A python script was written which loaded the articles from the downloaded CSV file, pre-processed the raw data and then ran the VADER sentiment analysis model on the resulting text. It is logical to assume that the news category (business, entertainment, ect.) plays an important role in the articles overall sentiment. For instance, one would assume the technology category would contain less sentiment than the entertainment or sport category. The aim of this test was to see if the tool would behave as expected.

Figure 5.7 shows the scatter plot obtained by plotting the average sentiment for each article. This scatter plot shows that the majority of data points are clustered around a horizontal line at around zero sentiment intensity (or slightly above zero). Therefore, a large proportion of articles had a sentiment score that was very close to zero. This was likely due to the fact that many of the news articles sentences are objective and involve no sentiment. A further reason for this, was that in taking the articles average sentiment, sentences with a positive sentiment score will cancel out sentences with a negative sentiment score. If the quantity and extremity for both positive and negative sentence sentiments are similar then the overall score will be close to zero. Therefore this was actually a measure of impartiality or sentiment/opinion cancellation. This was taken into consideration when analysing further data sets and can be avoided by simply plotting the sentence sentiments or by taking the absolute value of the sentiment before averaging.

Another observation is that a large proportion of technology articles (purple) are tightly Clustered around the horizontal at zero, whereas the sport (yellow) and entertainment (green) articles are more spread out. This difference is quantified by the variance as shown in table 5.1. This makes sense on a human level as sport and entertainment articles would be expected to contain more sentiment than technology articles.

This initial test revealed an aspect of the project that had not been previously considered. This aspect is data representation. Though the above observations were made from the scatter plot of BBC article sentiments, better insights and more obvious patters emerge when changing the data representation. Figure 5.8 is a histogram of the BBC article sentiments for the five news categories. A kernel density estimate (KDE) was preformed using this histogram resulting in figure 5.9. Presenting the data in this form reveals a lot more about the different news categories as well as emphasising the previous observations.

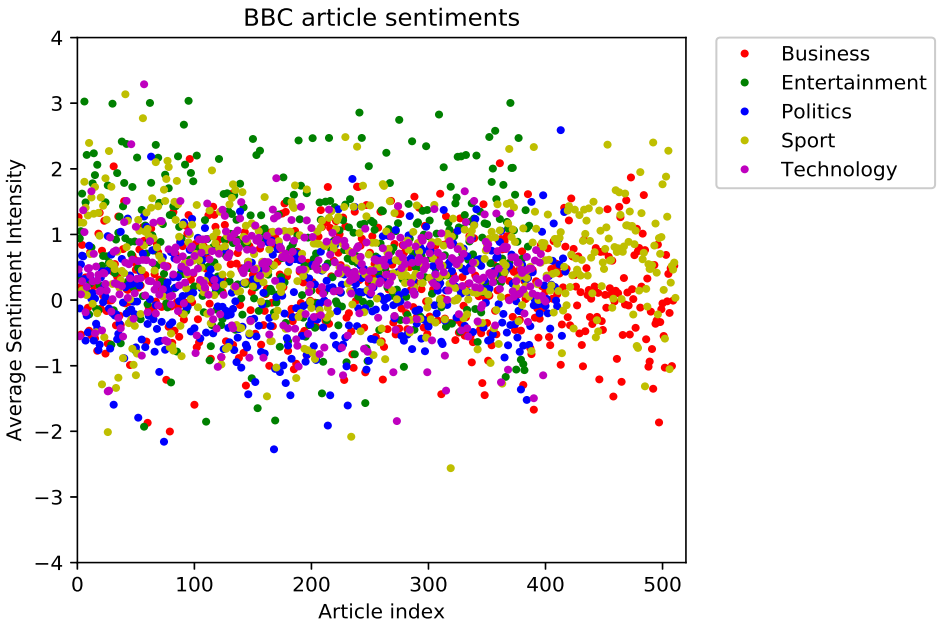

Figure 5.7: Scatter plot of the average sentiment intensity for BBC articles belonging to five different news categories.

Article Average	Sentiment Intensity Mean	Sentiment Intensity Variance
BBC Business	0.217	0.509
BBC Entertainment	0.744	0.863
BBC Politics	0.068	0.435
BBC Sport	0.555	0.604
BBC Technology	0.382	0.387

Table 5.1: Table showing the mean and variance of the sentiment intensity scores for BBC articles.

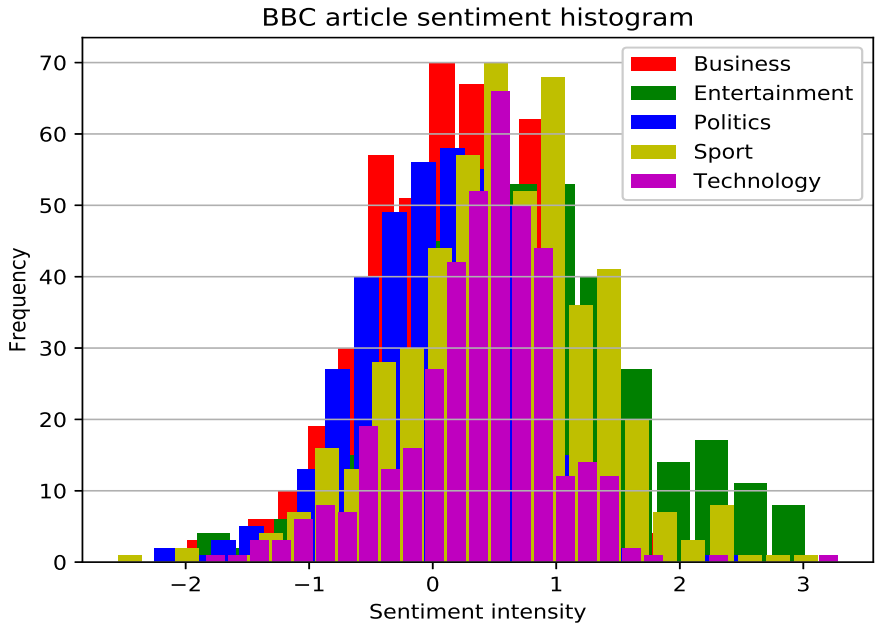

Figure 5.8: Histogram of the average sentiment intensity for BBC articles belonging to five different news categories.

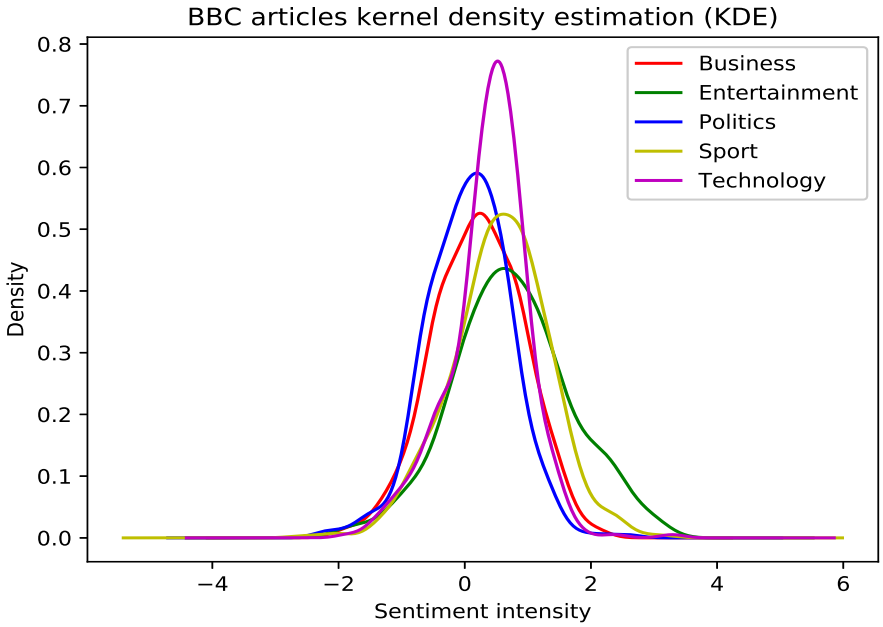

Figure 5.9: Kernel density estimation of the average sentiment intensity for BBC articles belonging to five different news categories.

To test the tools ability to distinguish between fact and fake news, three data sets were obtained with pre-labelled fact and fake news articles. The first data set consisted of 182 articles with an even split between the fact and fake classes. In this data set the ground truths were collected from BuzzFeed. The second data set contained 240 articles, 120 of which were labelled as fact and the other 120 were labelled as fake. The ground truths in this data set were collected from PolitiFact. These two data sets were acquired from the data repository FakeNewsNet [7]. The third data set was acquired from Kaggle [8] (a data science website owned by Google) which has 20,800 articles labelled as "reliable" and "unreliable".

Sentiment analysis was preformed on the FakeNewsNet data set, as well as a subset of the articles provided by Kaggle. The subset consisted of 200 randomly selected articles (100 fact, 100 fake) from the available 20,800 articles. A subset was used, as tokenizing 20,800 articles into sentences and calculating a sentiment score for each sentence was computationally expensive and unnecessary as the other data sets only had approximately 200 articles. Along with computing the average sentiment of an entire article, the sentiments of the sentences themselves were taken and plotted. This avoids positive and negative sentence sentiment scores cancelling out. Hence, retaining all information contained within the data. The results of this analysis is presented in the following section.

5.4 Results

As previously mentioned in section 5.3, it is difficult to discern any patterns from the average sentiment scatter plot. As shown in figure 5.10, the vast majority of articles are clustered around the neutral horizontal and the distribution of extreme sentiment articles seems random. Though overall it appears as if more fake news articles lie outside the neutral region than fact news articles. This plot shows that there is clearly no linear decision boundary that can separate the two news categories.

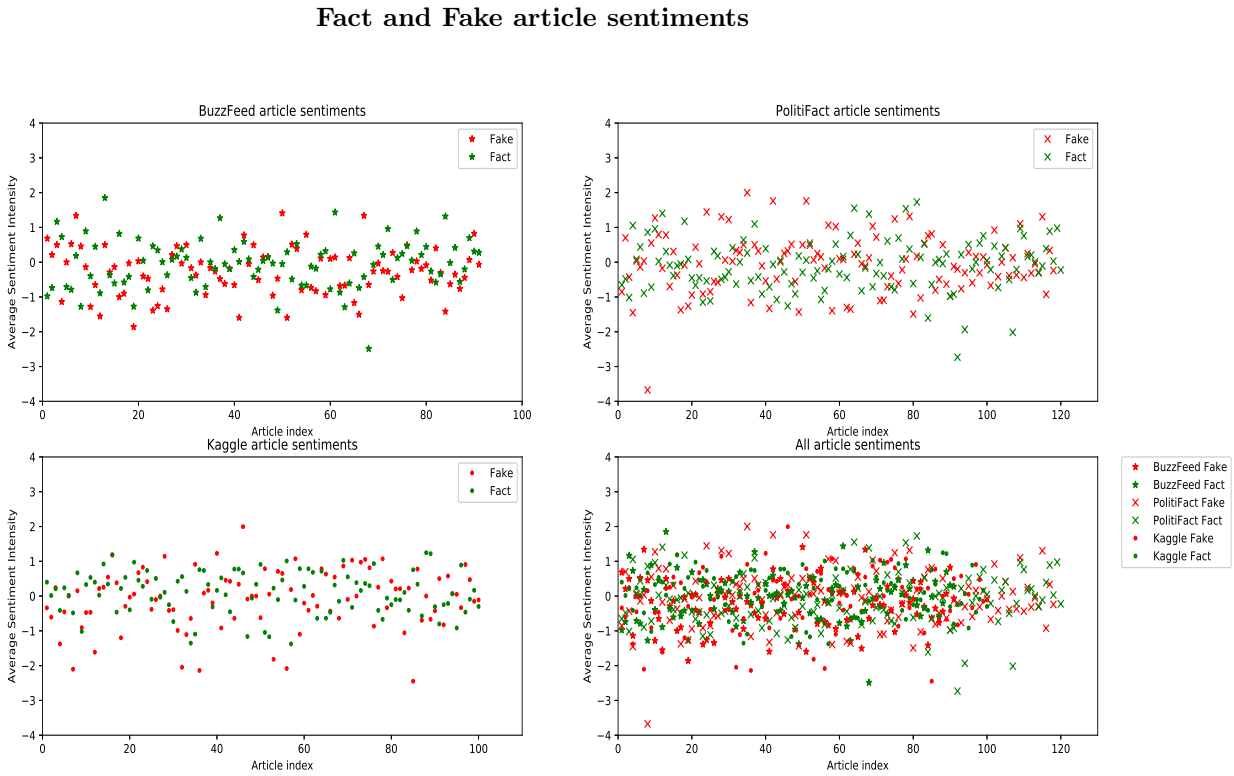

Figure 5.10: Scatter plot of article average sentiment intensity for three different fact/fake news data sets. Green represents fact news and red represents fake or false news. The three data sets are provided by BuzzFeed [7] (top left), PolitiFact [7] (top right) and Kaggle [8] (bottom left). The bottom right graph is a plot of the article average sentiment for articles in all data sets.

A second method of representing the data is to plot the average articles sentiment in order. This resulted in figure 5.11. This figure confirms the previous observation; fake news articles tend to be less neutral than fact news articles. This is indicated by the bottom left plot in figure 5.11. This plot combines all the data sets and as shown, both lines, fact and fake, lie below the neutral axis for the majority of the graph. This, in addition to the line of fake article sentiments lying below the line fact news article sentiments means fake articles are less neutral than fact articles. Though the difference is very little. A further observation is that fact news articles tend to be more positive that fake news articles. Though, again, the difference may be insignificant.

Ordered Fact and Fake article sentiments

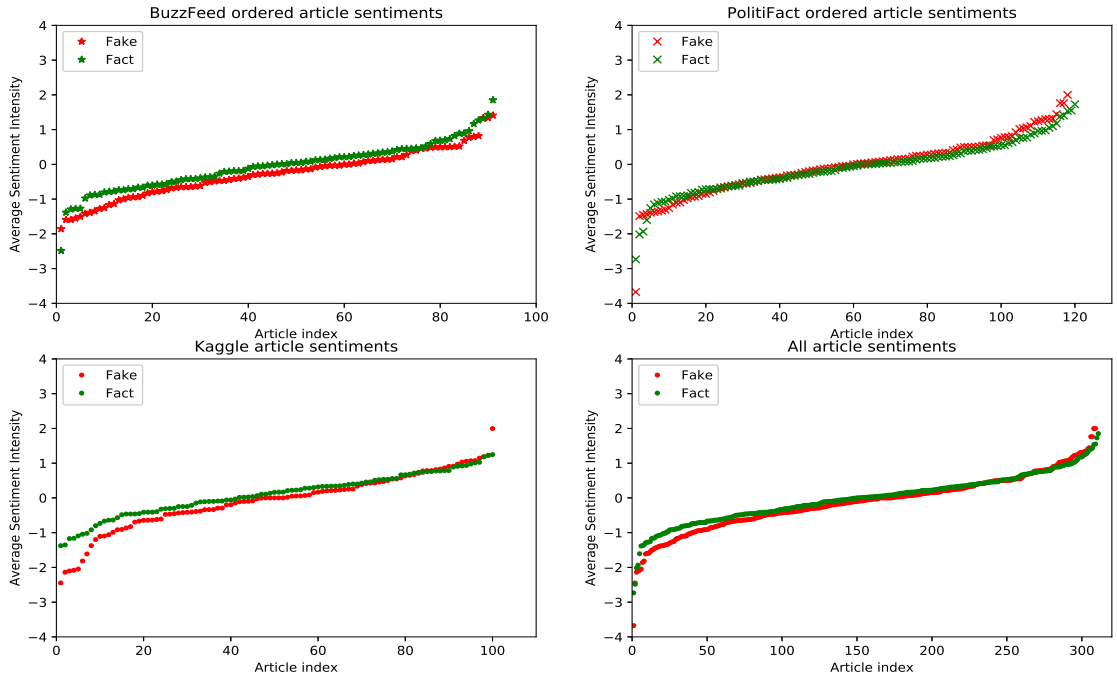

Figure 5.11: Ordered plot of article average sentiment intensity for three different fact/fake news data sets. Green represents fact news and red represents fake or false news. The three data sets are provided by BuzzFeed [7] (top left), PolitiFact [7] (top right) and Kaggle [8] (bottom left). The bottom right graph is a plot of the ordered article average sentiment for articles in all data sets.

Figure 5.12 is an ordered plot of the average absolute article sentiments. Taking the absolute value before averaging avoids sentiment cancellation. As the values can only be positive, the plot never falls below the neutral axis and the minimum sentiment score is zero. It is now obvious that fake news articles are less neutral. This is because, in the plot consisting of all articles from all data sets (bottom right in figure 5.12), the red line predominantly lies above the green. Furthermore, in three of the four plots the maximum sentiment article are fake, as shown by the red curve pealing away form the green line as the article index increases.

Ordered absolute Fact and Fake article sentiments

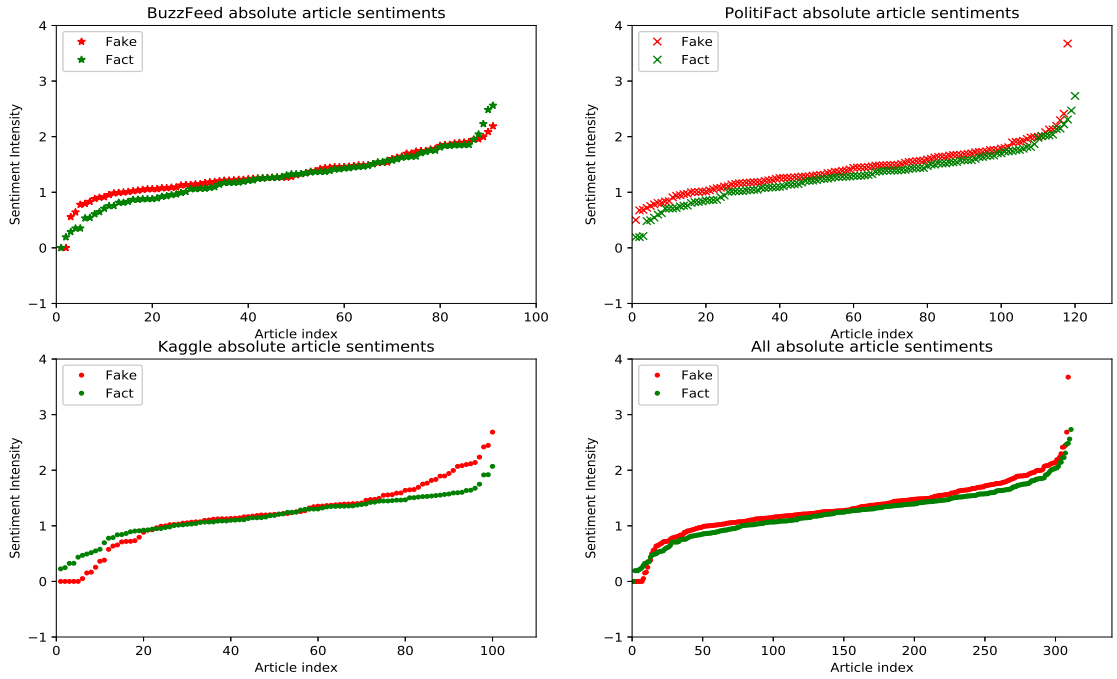

Figure 5.12: Ordered plot of article average absolute sentiment intensity for three different fact/fake news data sets. Green represents fact news and red represents fake or false news. The three data sets are provided by BuzzFeed [7] (top left), PolitiFact [7] (top right) and Kaggle [8] (bottom left). The bottom right graph is a plot of the ordered article average absolute sentiment for articles in all data sets.

Fact and Fake article sentiment histograms

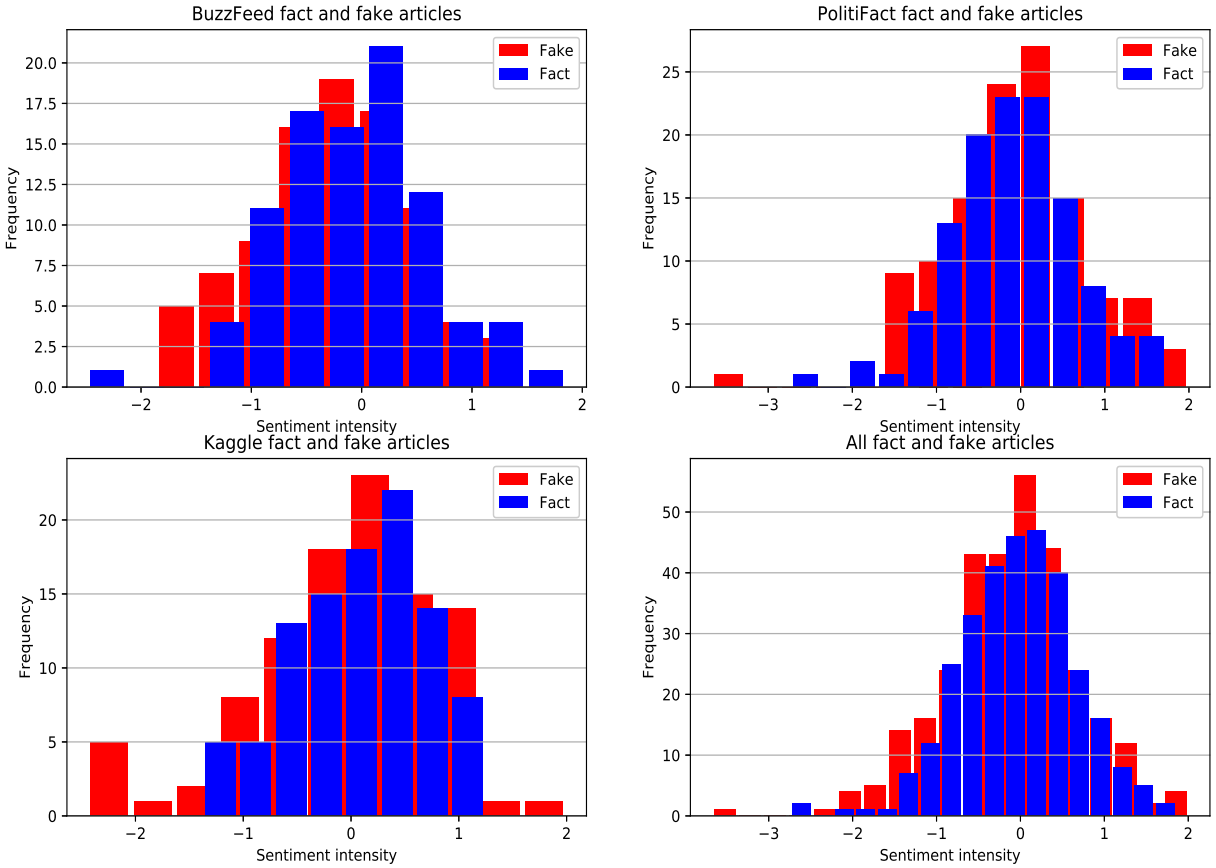

Figure 5.13: Histograms of the average sentiment intensity for news articles belonging to three different fact/fake news data sets. Blue represents fact news while red represents fake/false news. The data sets used are: BuzzFeed [7] (top left), PolitiFact [7] (top right) and Kaggle [8] (bottom left). The bottom right figure is a histogram of the average sentiment for articles in all data sets.

The kernel density estimate of the histogram provided in figure 5.13 shows just how similar the two news groups are. The average news article sentiments are almost completely overlapping in this density representation. Both fact and fake news appear to follow a relatively normal distribution centred around zero sentiment intensity and almost identical variances. All four plots show fact new to have a slightly sharper peak around the neutral sentiment score. The same is not true for the kernel density estimate of the average absolute sentiment score given in figure 5.15. Though overall the same conclusion of near identically is reached.

Kernel Density Estimate (KDE) of fact and fake news article sentiments

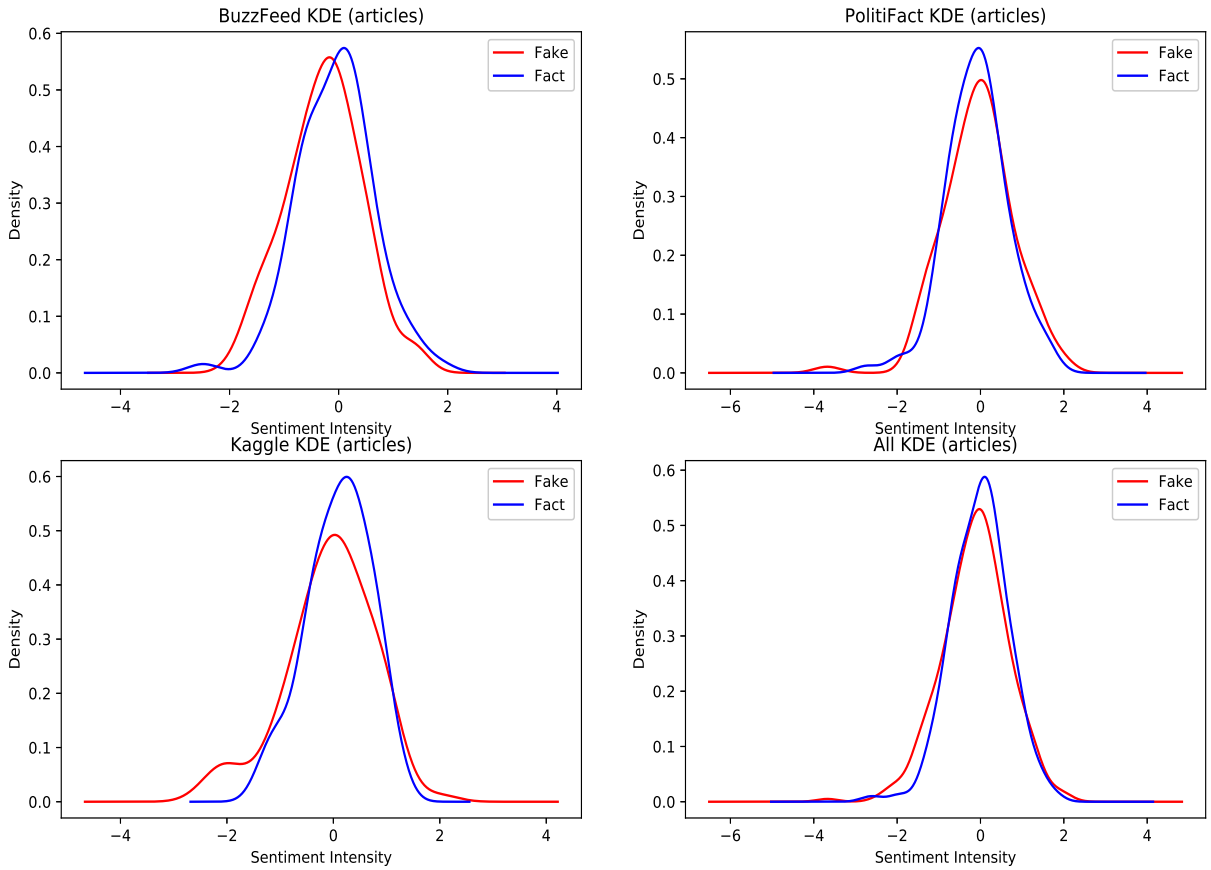

Figure 5.14: Kernel density estimates (KDE) of the average sentiment intensity for news articles belonging to three different fact/fake news data sets. The KDE of fact news is plotted in blue while the KDE of fake news is plotted in red. The data sets used are: BuzzFeed [7] (top left), PolitiFact [7] (top right) and Kaggle [8] (bottom left). The bottom right figure is the KDE of the average sentiment for articles in all data sets.

Kernel Density Estimate (KDE) of absolute fact and fake news article sentiments

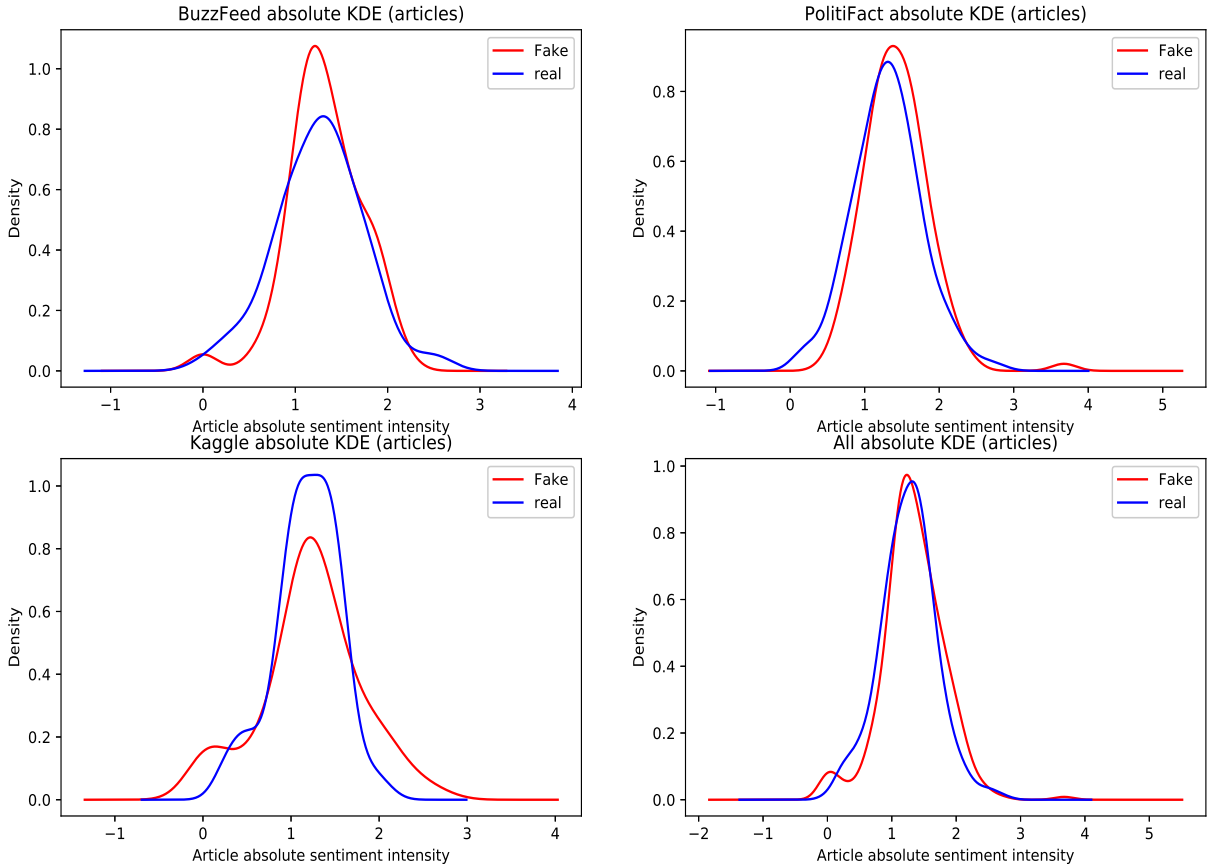

Figure 5.15: Kernel density estimates (KDE) of the average absolute sentiment intensity for news articles belonging to three different fact/fake news data sets. The KDE of fact news is plotted in blue while the KDE of fake news is plotted in red. The data sets used are: BuzzFeed [7] (top left), PolitiFact [7] (top right) and Kaggle [8] (bottom left). The bottom right figure is the KDE of the average absolute sentiment for articles in all data sets. The mod operation was performed before taking the average so no sentiment cancellation occurs.

Sentiment scores are derived for individual sentences. Therefore, the list of sentence sentiment scores contains the maximum amount of information. Plotting each sentence score in fact and fake news articles gives the following scatter plot. It is interesting to note that in all data sets, fact news articles contained more sentences than fake news articles. This is shown by the region to the right of each plot in figure 5.16 where only green markers are present.

Fact and Fake sentence sentiments

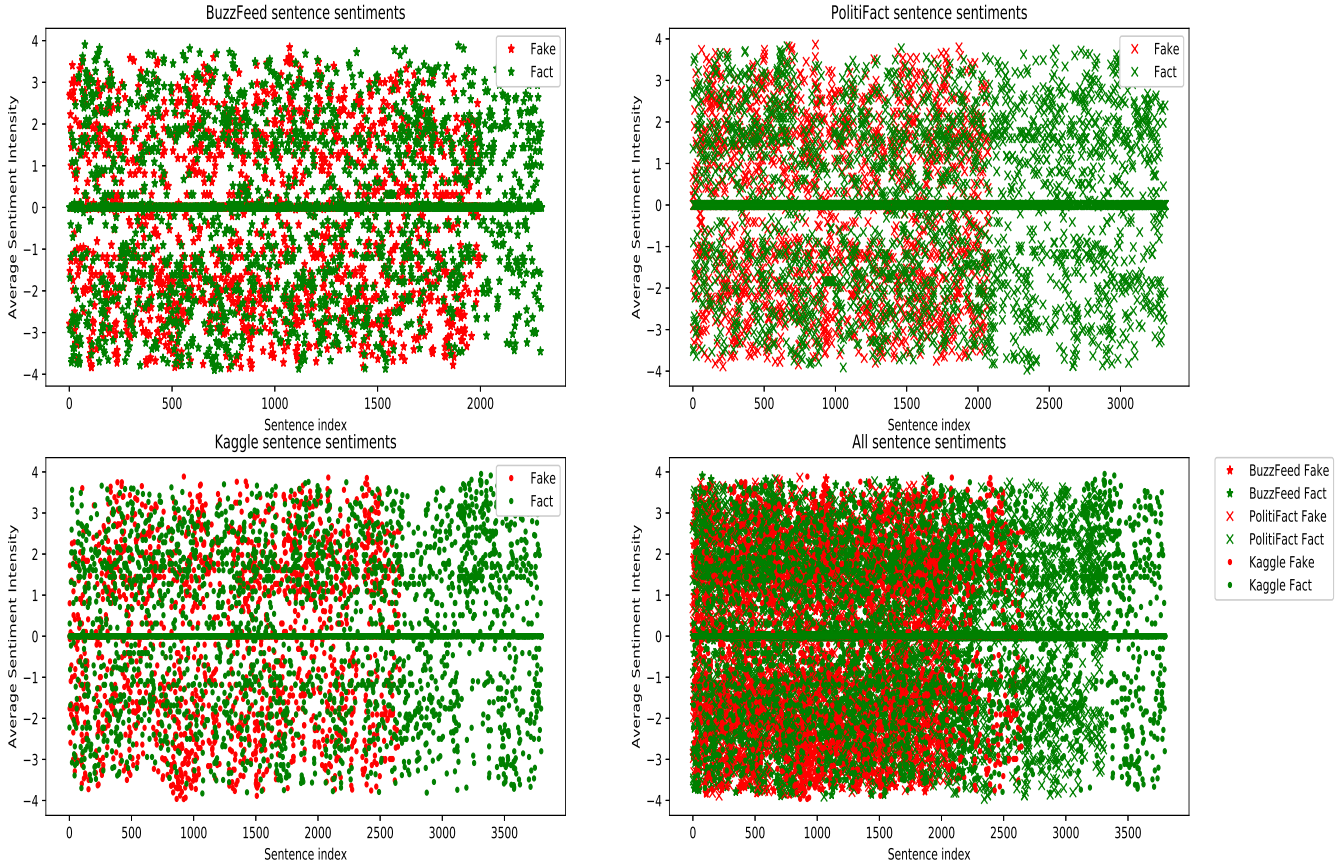

Figure 5.16: Scatter plot of sentence sentiment intensity for articles in three different fact/fake news data sets. Green represents sentences in an article labelled fact news and red represents sentences in an article labelled fake or false news. The three data sets are provided by BuzzFeed [7] (top left), PolitiFact [7] (top right) and Kaggle [8] (bottom left). The bottom right graph is a plot of all sentence sentiment intensities in all articles in all data sets.

The histograms of sentence sentiment intensity reveal a surprising distribution in the frequency of sentences for each sentiment intensity. The distribution is no longer normal, as was the case when taking an average sentiment over the articles sentences. The sharp middle peak is expected as it indicates that most sentences are neutral and involve no sentiment. However, there are two smaller peaks, one either side of the main neutral peak. Overall, figure 5.17 shows that a large majority of sentences are neutral and not many sentences have extreme sentiment scores or close to neutral sentiment scores, the rest tend to cluster around ± 2. For comparison the word "good" has a sentiment score that is approximately 2 while the word "careless" has a sentiment score close to -2.

Fact and Fake sentence sentiment histograms

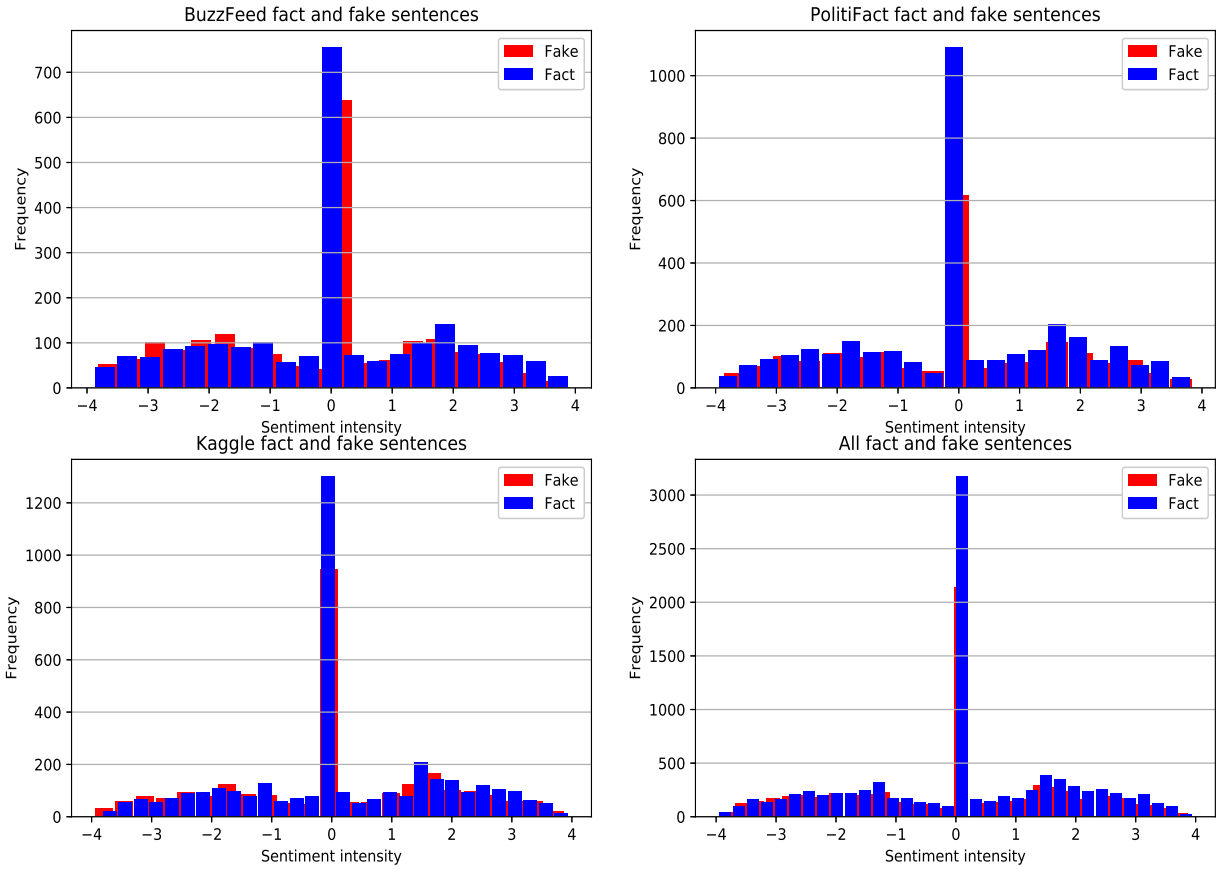

Figure 5.17: Histograms of sentence sentiment intensity for news articles belonging to three different fact/fake news data sets. Blue represents sentences in articles labelled as fact news while red represents sentences labelled as fake/false news. The data sets used are: BuzzFeed [7] (top left), PolitiFact [7] (top right) and Kaggle [8] (bottom left). The bottom right figure is a histogram of the sentence sentiments for all articles in all data sets.

Kernel Density Estimate (KDE) for fact and fake news sentence sentiments

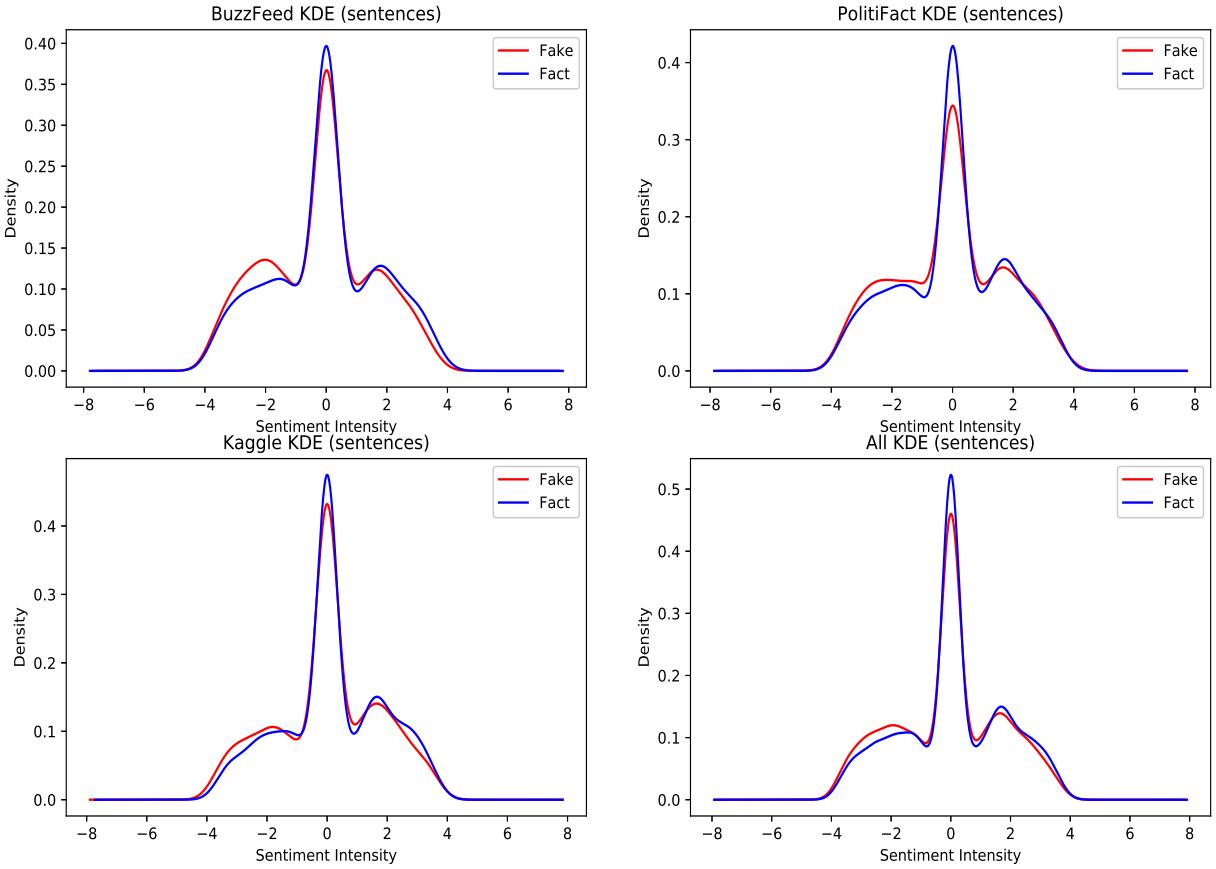

Figure 5.18: Kernel density estimates (KDE) of the sentence sentiment intensities for news articles belonging to three different fact/fake news data sets. The KDE of fact news is plotted in blue while the KDE of fake news is plotted in red. The data sets used are: BuzzFeed [7] (top left), PolitiFact [7] (top right) and Kaggle [8] (bottom left). The bottom right figure is the KDE of the sentence sentiments for all articles in all data sets.

Plotting the sentiment intensity for each word and symbol in the VADER lexicon provides a possible explanation for the observed side peaks. As shown in figure 5.19 the VADER lexicon itself exhibits this double peak distribution. Therefore, it is reasonable to assume that since the majority of lexical features in VADER have an associated sentiment of approximately ±2; any sentiment analysis on large amounts of text will also display this distribution.

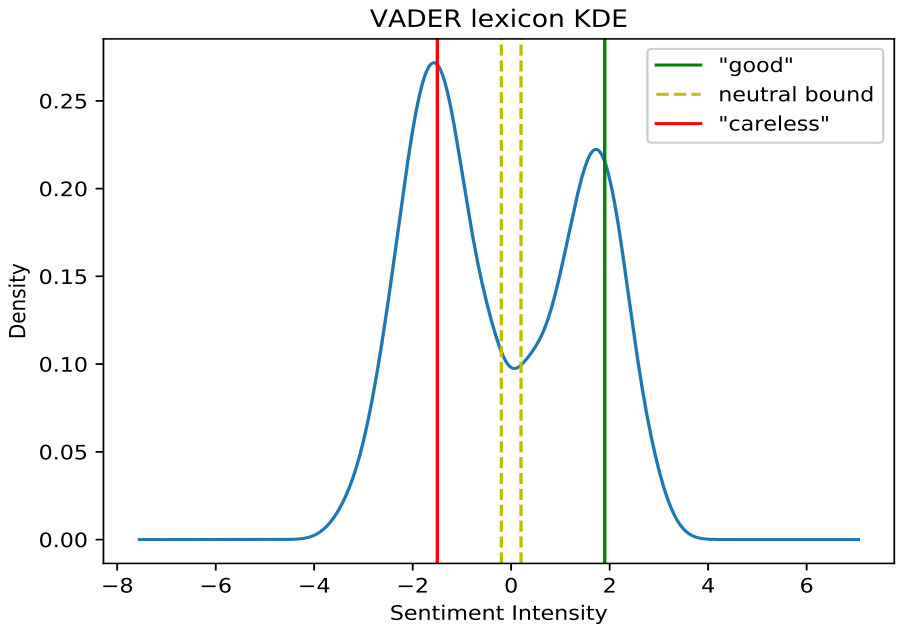

Figure 5.19: Kernel density estimate of the VADER lexicon. Red and green lines indicate the sentiment value of the words "careless" and "good" respectively. All officially neutral words are between the neutral bound.

5.5 Evaluation

Close inspection of the results showed minor differences between the sentiment scores of fact and fake news articles. However these differences are extremely small. Therefore, no acceptable classification boundary can be found to separate the two news groups based solely on the computed sentiment scores. All observations suggest that, in practice, current sentiment analysis techniques cannot be used to differentiate between fact and fake news. This evaluation is based on qualitative interpretation of the results. To obtain a more rigorous conclusion, a statistical hypothesis test was preformed.

The two-sample Kolmogorov–Smirnov statistics test (K-S test) can be used to test the equality of probability distributions. Therefore, it is used here to compare the sentiments of the fact and fake news articles to see whether or not there exists a significant difference between the two sentiment distributions. The K-S test does this by quantifying the distance between the two empirical distribution function of the two samples.

The chosen null hypothesis is that sentiment cannot be used to separate fact news from fake news. Hence, the alternate hypothesis is that sentiment can be used to distinguish between the two from of news. The only statistical assumption made is that the two news groups are independent. A Non statistical assumption is that all news article sentiments are correctly computed, that is, humans would agree with the sentiment score obtained via sentiment analysis. Under the null hypothesis there will not be a significant difference in the distributions of fact and fake news. The significance threshold below which the null hypothesis will be rejected was chosen to be 5%.

Using the python SciPy package allowed the two-sample K-S test to be imported as a function which took both the fact and fake sample distributions as its arguments. The obtained p-value for the average sentiment of all articles was 0.21 while the p-value calculated for the average absolute sentiment of all articles was 0.19. Both these p-values are larger than the chosen significance level. Therefore the null hypothesis cannot be rejected, meaning the two sample distributions are practically identical and sentiment cannot be used to separate fact news from fake news.

Evaluating the tools ability to distinguish between opinion pieces and real news involved writing a python script to scrape news articles from The New York Times website. Twenty random articles were scraped from two New York Times news categories. These categories were opinion and world news. Sentiment analysis was then performed on these articles producing the plots in figure 5.20. These plots show that the tool struggles to determine the difference between opinion pieces and world news articles. It would appear that current state of the art sentiment analysis techniques don't seem to be able to capture the significant differences between the two news group.

The New York Times opinion pieces and world news

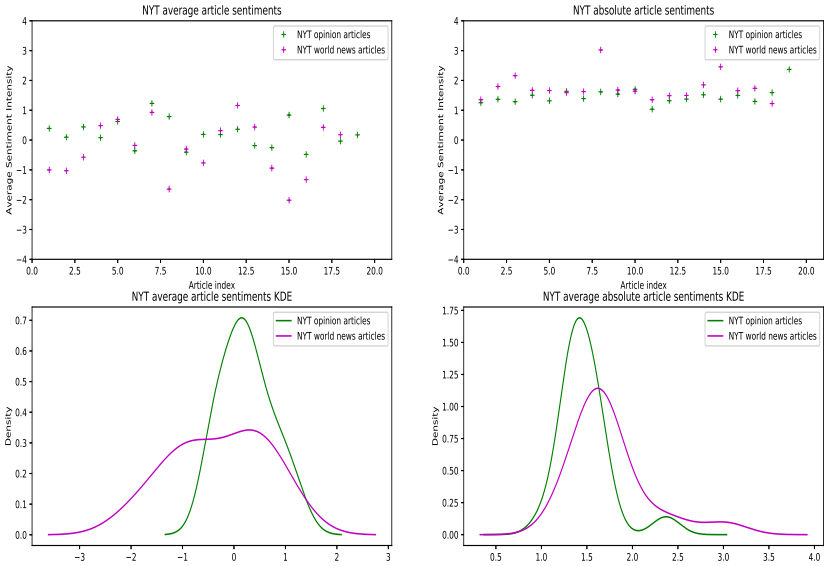

Figure 5.20: A random comparison of New York Times opinion pieces and New York Times world news articles. This shows that the VADER sentiment analysis model struggles to distinguish between opinion pieces and world news articles as the curves cannot be easily separated.

Overall, the tool failed to find a reasonable classification boundary between fact and fake news. Hence it does not meet the first necessary requirement defined in the requirements capture (see section 3). However, the tool does meet all other requirements as it is objectively convenient, intuitive, transparent and accessible. Though being ineffective at classifying a news piece is reason enough to peruse another approach.

Chapter 6

Stance Detection

The previously designed web browser extension can be manipulated to perform stance detection. Stance detection is vital for automated cross checking and can be exploited to derive a consensus score. As described in section 2.3.2, the *The Fake News Challenge* held an artificial intelligence stance detection competition. Therefore, a pre-trained stance detection neural network can be used to predict an articles stance relative to a headline or statement. The pre-trained model developed by UCL Machine Reading [3] was chosen to perform stance detection. This was mainly due to its simplicity and accuracy. The model finished third in the competition scoring 81.72%, which was only 0.3% behind first place. the models simplicity allowed prediction to be made much faster and system integration was easier when compared to the competing systems.

The overall stance detection system design is illustrated in by the flow chart given in figure 6.1. All source code implementing the stance detection tool can be found at:

https://github.com/OrionMat/Consensus-Analysis.

Consensus Analysis Flow Diagram

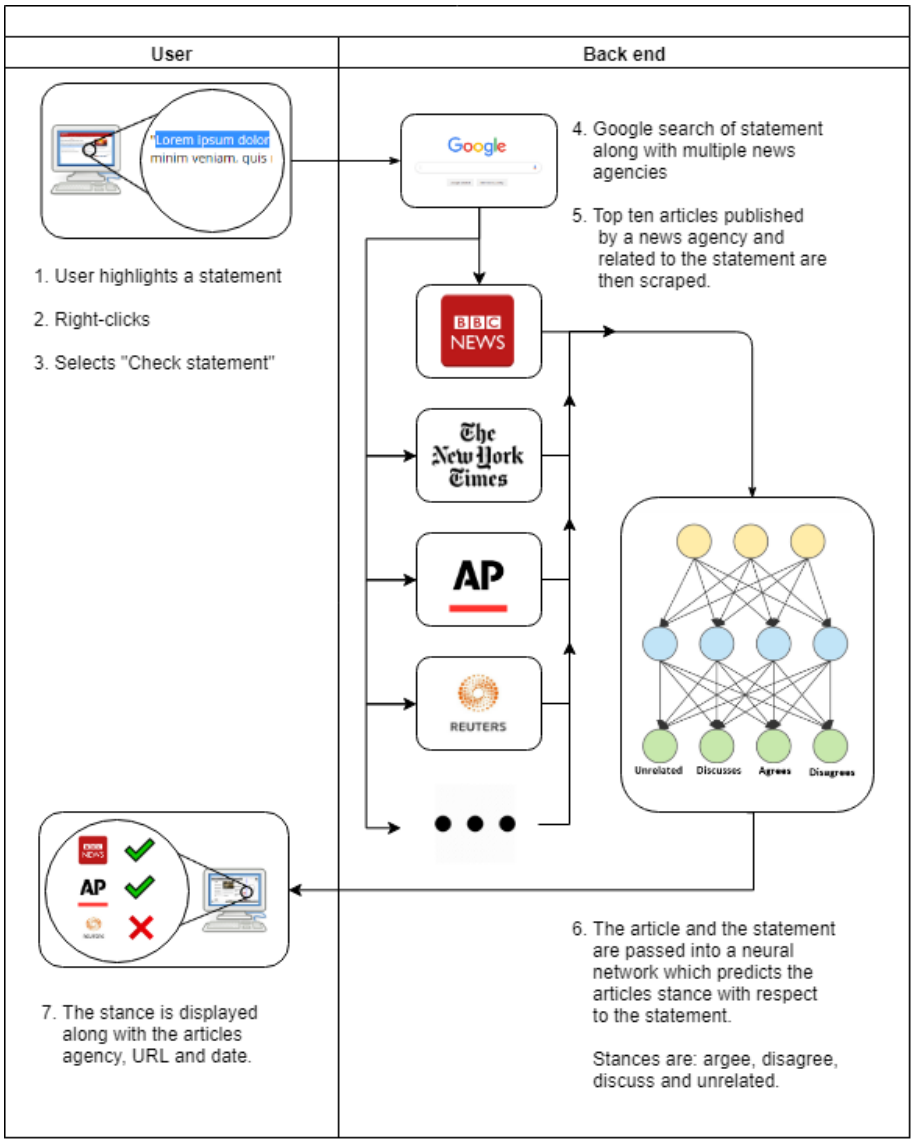

Figure 6.1: Flowchart illustrating the consensus analysis procedure. Initially the user highlights a statement or claim. Upon right-clicking the selected text a context menu appears giving the option to "check statement". If this option is chosen the browser extension sends the selected text to a python script which executes a Google search of the statement appended to a specific news agency. The top ten articles related to the selected statement are then scraped from the news agency website. The article and the statement are then passed though a neural network which predicts the article's stance relative to the statement. That is, the network predicts whether the articles agree, disagree, discuss or are unrelated to the statement. The result, that is the stance, is then sent back to the browser extension to be displayed in a popup alongside the articles agency, URL and date.

Consensus Analysis Results Popup

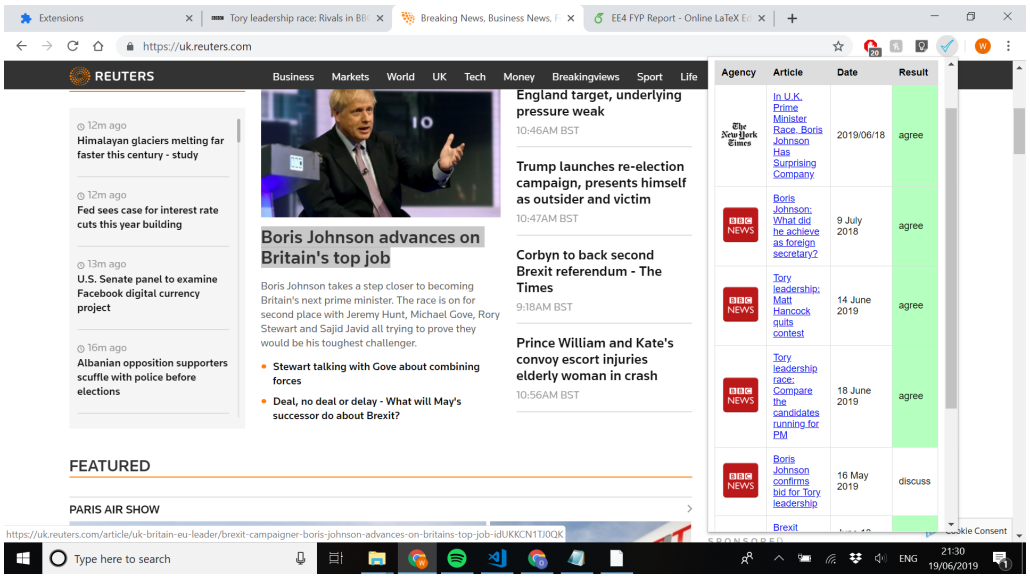

Figure 6.2: Screenshot of Consensus analysis tool results. The results are presented in a popup that appears when the tool icon is selected.

The Media Bias Chart

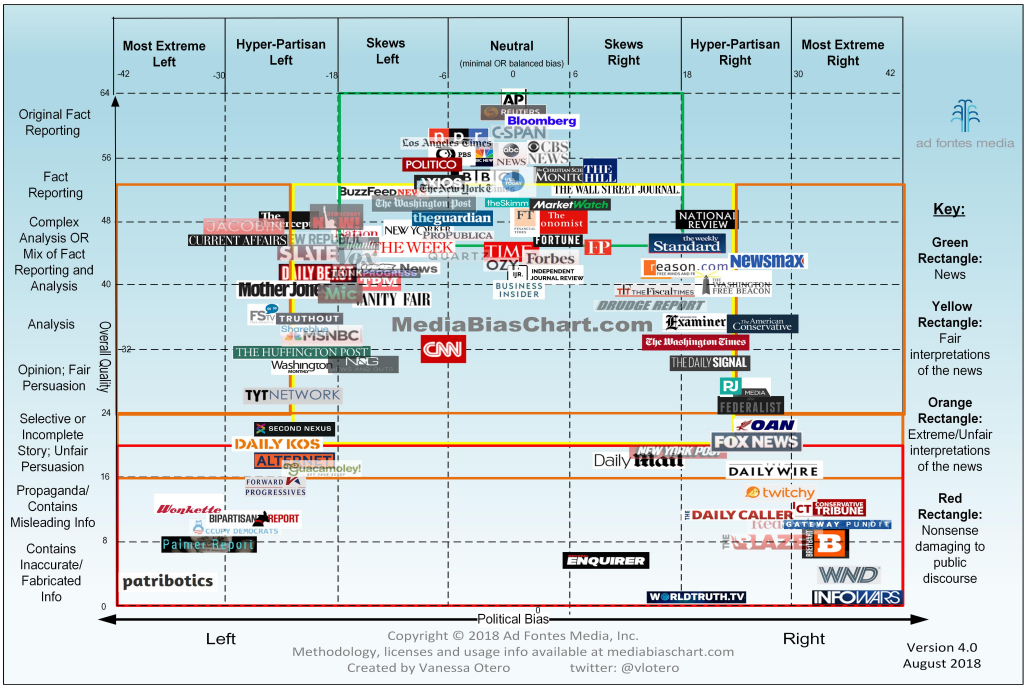

Figure 6.3: "Media Bias Chart: Version 4.0". (Vanessa Otero, 2018 [9]). This chart illustrates the variation in bias (x-axis) and quality (y-axis) observed in mainstream news sources. This chart was used to choose reliable news sources as part of the stance detection system.

User	Positives	Negatives
1	Encouraged me to question things I read rather than just believing them	Quite slow
2	Gave me confidence in an article when matches were found	Some more niche articles didn't work well
3	Easy to use	Struggled to find matches for smaller stories/statements
4	Works well for significant articles/headlines	"Couldn't't match very new articles even from major agencies"
5	Made me think about news more	"Would be useful to have an indicator that it is working e.g. loading bar"
6	The article date tool is very useful	Would be good to choose ones own agencies to check against
7	Colour coding the results is good	Results should pop up automatically rather than being clicked on
8	Results are easy to glance over	"Information about the author not just the agency would be helpful"
9	User friendly	Some metric like an overall agree/disagree would be useful
10	Matches articles from big news stations well	"Not so good for checking facts e.g. malaria kills 800 000 people every year"
11	Quite easy to use	Takes a long time
12	Easier than doing a Google search manually	Couldn't't tell if it was working or not until it finished
13	Liked being given the date of each result	I'd like to be able to specify the news sites it checks against
14	Nice that it works in the background so does not get in the way	Takes quite a long time
15	Result colours make reading results easy	It would be nice to be able to define whether or not I think a source is reliable and then have that come up in the results
16	I thought about the thing I was reading a lot more	It would be good to have results pop up automatically
17	The results were easy and quick to read	The results are quite slow
18	Tool is unobtrusive	Could be good to chose different agencies to check against
19	Matches give confidence when reading articles on social media	Some loading indicator would be useful
20	Made reading news more interesting to see what other agencies agreed with it	Only worked well with major articles

Table 6.1: Qualitative results taken by observing users over a 20 minute period. Users tended to be browsing social media and news websites.

Chapter 7

Conclusion

Two tool were created to combat the spread of fake news. The first relied on a sentiment score while the second made use of stance detection. Though the sentiment analysis tool proved to be inaccurate at analysing the sentiment of news articles, this was not the case for social media posts. Both tools were integrated into a sing web browser extension which increased the chances users that a statement would be questioned and fact checked.

Bibliography

[1] C. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," 2014. [Online]. Available: http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf

[2] D. Pomerleau and D. Rao, *Fake News Challenge: Exploring how artificial intelligence technologies could be leveraged to combat fake news.* [Online]. Available: http://www.fakenewschallenge.org/

[3] B. Riedel, I. Augenstein, G. P. Spithourakis, and S. Riedel, "A simple but tough-to-beat baseline for the Fake News Challenge stance detection task," *CoRR*, vol. abs/1707.03264, 2017. [Online]. Available: http://arxiv.org/abs/1707.03264

[4] W3Counter, "Web browser market share," 2019. [Online]. Available: https://www.w3counter.com/globalstats.php

[5] StatCounter, "Web browser market share," 2019. [Online]. Available: http://gs.statcounter.com/browser-market-share

[6] NetMarketShare, "Web browser market share," 2019. [Online]. Available: https://netmarketshare.com/browser-market-share.aspx?options=%7B%22filter%22%3A%7B%22%24and%22%3A%5B%7B%22deviceType%22%3A%7B%22%24in%22%3A%5B%22Desktop%2Flaptop%22%2C%22Mobile%22%2C%22Handheld%22%5D%7D%7D%5D%7D%2C%22dateLabel%22%3A%22Trend%22%2C%22attributes%22%3A%22share%22%2C%22group%22%3A%22browser%22%2C%22sort%22%3A%7B%22share%22%3A-1%7D%2C%22id%22%3A%22browsersDesktop%22%2C%22dateInterval%22%3A%22Monthly%22%2C%22dateStart%22%3A%222018-06%22%2C%22dateEnd%22%3A%222019-05%22%2C%22segments%22%3A%22-1000%22%7D

[7] S. Kai, M. Deepak, W. Suhang, L. Dongwon, and L. Huan, "FakeNewsNet: A data repository with news content, social context and dynamic information for studying fake news on social media," *arXiv preprint arXiv:1809.01286*, 2018.

[8] Kaggle. (22/01/18) Fake news: Build a system to identify unreliable news articles. [Online]. Available: https://www.kaggle.com/c/fake-news/data

[9] V. Otero, *Media Bias Chart.* [Online]. Available: https://www.adfontesmedia.com/

[10] Y. Adegoke, "Like. share. kill." 13/11/2018. [Online]. Available: https://www.bbc.co.uk/news/resources/idt-sh/nigeria_fake_news

[11] A. Kanski, *Study: 86% of people don't fact check news spotted social media*, 2017. [Online]. Available: https://www.prweek.com/article/1431578/study-86-people-dont-fact-check-news-spotted-social-media

[12] E. Brown, *9 out of 10 Americans don't fact-check information they read on social media*, 2017. [Online]. Available: https://www.zdnet.com/article/nine-out-of-ten-americans-dont-fact-check-information-they-read-on-social-media/

[13] P. Dizikes, "Study: On twitter, false news travels faster than true stories," 08/03/2018. [Online]. Available: http://news.mit.edu/2018/study-twitter-false-news-travels-faster-true-stories-0308

[14] Z. Kleinman, "Fake news 'travels faster', study finds," 09/03/2018. [Online]. Available: https://www.bbc.co.uk/news/technology-43344256

[15] *Information Definition*. [Online]. Available: https://dictionary.cambridge.org/dictionary/english/information

[16] *News Definition*. [Online]. Available: https://en.oxforddictionaries.com/definition/news

[17] *Fake News Definition*. [Online]. Available: https://dictionary.cambridge.org/dictionary/english/fake-news

[18] A. Subedar, "The godfather of fake news," 27/11/2018. [Online]. Available: https://www.bbc.co.uk/news/resources/idt-sh/the_godfather_of_fake_news

[19] J. Vincent, "Watch jordan peele use ai to make barack obama deliver a psa about fake news," 17/04/18. [Online]. Available: https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peele-buzzfeed

[20] Wikipedia, "Fake news," 18/01/19. [Online]. Available: https://en.wikipedia.org/wiki/Fake_news

[21] E. J. Kirby, "The city getting rich from fake news," 05/12/16. [Online]. Available: https://www.bbc.co.uk/news/magazine-38168281

[22] H. Allcott, "Trends in the diffusion of misinformation on social media," 16/09/2018. [Online]. Available: https://arxiv.org/abs/1809.05901

[23] Facebook business, "Third-party fact-checking on facebook." [Online]. Available: https://en-gb.facebook.com/help/publisher/182222309230722?helpref=faq_content

[24] J. Wakefield, "Facebook employs uk fact-checkers to combat fake news," 11/01/19. [Online]. Available: https://www.bbc.co.uk/news/technology-46836897

[25] TEAM FULL FACT, "Full fact to start checking facebook content as third-party factchecking initiative reaches the uk," 11/01/19. [Online]. Available: https://fullfact.org/blog/2019/jan/full-fact-start-checking-facebook-content-third-party-factchecking-initiative-reaches-uk/

[26] Facebook Newsroom, "Hard questions: How is facebook's fact-checking program working?" 14/06/2018. [Online]. Available: https://newsroom.fb.com/news/2018/06/hard-questions-fact-checking/

[27] M. Hughes, "Facebook just bought an ai startup to help it fight fake news," 2018. [Online]. Available: https://thenextweb.com/artificial-intelligence/2018/07/02/facebook-just-bought-an-ai-startup-to-help-it-fight-fake-news/

[28] E. Schmitt, D. Mannion, M. Bronstein, and F. Monti, "FABULA AI." [Online]. Available: https://fabula.ai/

[29] X. L. Dong, E. Gabrilovich, K. Murphy, V. Dang, W. Horn, C. Lugaresi, S. Sun, and W. Zhang, "Knowledge-based trust: Estimating the trustworthiness of web sources," 12/02/15. [Online]. Available: https://arxiv.org/pdf/1502.03519.pdf

[30] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor., "Freebase: a collaboratively created graph database for structuring." [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.538.7139&rep=rep1&type=pdf

[31] A. Edell, "Fakebox." [Online]. Available: https://machinebox.io/docs/fakebox

[32] R. Baly, G. Karadzhov, D. Alexandrov, J. Glass, and P. Nakov, "Predicting factuality of reporting and bias of news media sources," 02/10/2018. [Online]. Available: https://arxiv.org/pdf/1810.01765.pdf

[33] C. Pease, "Machine learning tackles the fake news problem," 21/08/18. [Online]. Available: https://towardsdatascience.com/machine-learning-tackles-the-fake-news-problem-c3fa75549e52

[34] S. Gilda, "Evaluating machine learning algorithms for fake news detection," 14/12/17. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8305411

[35] K. Foy, "Using machine learning to detect fake news," 03/10/2017. [Online]. Available: https://www.ll.mit.edu/news/using-machine-learning-detect-fake-news

[36] A. Edell, "I trained fake news detection ai with >95% accuracy, and almost went crazy," 11/01/18. [Online]. Available: https://towardsdatascience.com/i-trained-fake-news-detection-ai-with-95-accuracy-and-almost-went-crazy-d10589aa57c

[37] W. Y. Wang. (01/05/17) Liar, Liar, Pants on Fire: A new benchmark dataset for fake news detection. [Online]. Available: https://arxiv.org/abs/1705.00648

[38] B. Liu, "Sentiment analysis and opinion mining," 22/04/2012. [Online]. Available: https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf

[39] Pennebaker, J.W, Booth, R.J, Boyd, R.L, Francis, and M.E, "Linguistic inquiry and word count: Liwc2015," 2015. [Online]. Available: http://liwc.wpengine.com/

[40] Stone, P.J, Dunphy, D.C, Smith, and M.S, "The general inquirer: A computer approach to content analysis," 1966. [Online]. Available: https://psycnet.apa.org/record/1967-04539-000

[41] B. Liu and M. Hu, "Opinion mining, sentiment analysis, and opinion spam detection," 2004. [Online]. Available: https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#lexicon

[42] M. M. Bradley and P. J. Lang, "Affective norms for english words (anew): Instruction manual and affective ratings," 1999. [Online]. Available: https://www.uvm.edu/pdodds/teaching/courses/2009-08UVM-300/docs/others/everything/bradley1999a.pdf

[43] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," 2011. [Online]. Available: http://www.sfu.ca/~mtaboada/docs/research/Taboada_etal_SO-CAL.pdf

[44] S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," 2010. [Online]. Available: http://nmis.isti.cnr.it/sebastiani/Publications/LREC10.pdf

[45] E. Cambria, D. Olsher, and D. Rajagopal, "Senticnet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis," 2014. [Online]. Available: https://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/viewPaper/8479

[46] S. M. Mohammad, P. Sobhani, and S. Kiritchenko, "Stance and sentiment in tweets," *Special Section of the ACM Transactions on Internet Technology on Argumentation in Social Media*, vol. 17, no. 3, 2017.

[47] P. Sobhani, D. Inkpen, and X. Zhu, "A dataset for multi-target stance detection," in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Valencia, Spain: Association for Computational Linguistics, Apr. 2017, pp. 551–557. [Online]. Available: https://www.aclweb.org/anthology/E17-2088

[48] W. Ferreira and A. Vlachos, "Emergent: a novel data-set for stance classification," 2016.

[49] Accuracy definition. [Online]. Available: https://dictionary.cambridge.org/dictionary/english/accuracy

[50] *Impartial definition*. [Online]. Available: https://en.oxforddictionaries.com/definition/impartiality

[51] *Consensus definition*. [Online]. Available: https://en.oxforddictionaries.com/definition/consensus

[52] *Journalistic Objectivity*. [Online]. Available: https://en.wikipedia.org/wiki/Journalistic_objectivity

[53] Orion, *Fact or Fake: EE4 final year project*. [Online]. Available: https://github.com/OrionMat/Sentiment-Analysis

[54] BBC, *BBC news article data set*. [Online]. Available: http://mlg.ucd.ie/datasets/bbc.html

[55] *Propaganda Definition*. [Online]. Available: https://dictionary.cambridge.org/dictionary/english/propaganda

[56] *Satirical News Definition*. [Online]. Available: https://en.wikipedia.org/wiki/Satire

[57] B. Marr, "Fake news: How big data and ai can help," 01/03/17. [Online]. Available: https://www.forbes.com/sites/bernardmarr/2017/03/01/fake-news-how-big-data-and-ai-can-help/#331bdec70d56

[58] A. Kittur, H. Chi, and B. Suh, "Crowdsourcing user studies with mechanical turk," in *Proc. CHI 2008, ACM Pres*, 2008, pp. 453–456.

[59] I. C. London, *Data Protection Policy*. [Online]. Available: http://www.imperial.ac.uk/admin-services/secretariat/information-governance/data-protection/data-aware/

[60] D. Ghulati and D. S. Riedel, *FACTMATA: Preserving the quality of the Internet*. [Online]. Available: https://factmata.com/

[61] Chrome, "Chrome runtime.port." [Online]. Available: https://developer.chrome.com/apps/runtime#type-Port

Appendix A

Definitions

News: The official definition of news is "Newly received or noteworthy information, especially about recent events" (Oxford Living Dictionaries [16]). This intentionally vague description gives no indication of the information's accuracy, reliability or bias. It does not specify what media the news has been spread by; if it is opinion and/or has been spread with some kind of intent. This umbrella term encompasses many different types of news, six of which will be considered in this project (as listed below). This official definition will be retained throughout this report, though, to standardise the analysis it will be assumed that the news is in written form.

Real News: Real news is news that is entirely accurate and non-bias. The information given is correct and consistent with real world events. The author gives no opinion and has the sole intention of informing readers of the facts. This ideal is impossible to achieve, as an author cannot guarantee that they had no bias or sentiment that may have influenced their writing or choice of story. It is also extremely difficult to be sure that all sources are completely accurate with their version of events. Therefore, the definition of real news will be relaxed slightly and defined as news that is verifiably correct and contains few opinions.

Persuasive News: Persuasive news is news where the information is correct but the piece aims to convince readers to share a certain point of view, that is, to influence their opinion on a specific topic. The most noteworthy example is propaganda, which is news where the information conveys only one side of an argument/story (Cambridge Dictionary [55]).

Fact News: Fact news will be defined as news where the information is correct and no consideration is given to potential bias or intent behind the article. Fact news can be seen as the combination of both real and persuasive news.

False News: False news is news that is incorrect but the misinformation is not intentional. The author was not deliberately producing untruthful news.

Fake News: Fake news is disinformation; the news piece is intentionally incorrect and untruthful (Cambridge Dictionary [17]).

Incorrect News: Incorrect news is the combination of fake and false news. The facts presented in the article are untruthful while the author's bias and/or intent is not considered.

News Metric: A feature of news (or news sources) that can be measured and used to classify news. Example metrics include accuracy, reliability and bias. These metrics could be used to classify news as real, false, fake, persuasive etc.

Information: Officially, information is defined as "facts provided about a situation, person, event, etc." (Cambridge Dictionary [15]). For the purpose of this report, the definition of information will not assume the "facts" are correct. Therefore, a fake news piece still contains information, though, the information may not be correct. For example, the statement "Veles is the fake news capital of the world" would be considered information regardless of whether or not Veles is indeed the fake news capital.

Appendix B

Sentiment Analysis Tool Popup

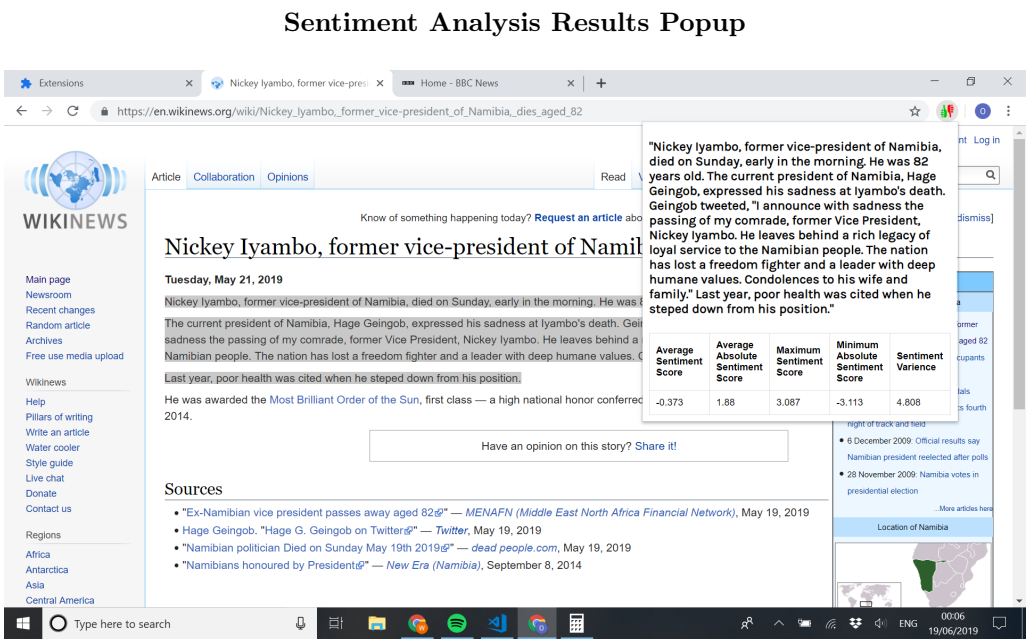

Figure B.1: Screenshot of sentiment analysis tool results. The results are presented in a popup that appears when the tool icon is selected.