Fool's Errand: Looking at April Fools Hoaxes as Disinformation through the Lens of Deception and Humour

Edward Dearden and Alistair Baron

Lancaster University, UK
initial.surname@lancaster.ac.uk

Abstract. Every year on April 1st, people play practical jokes on one another and news websites fabricate false stories with the goal of making fools of their audience. In an age of disinformation, with Facebook under fire for allowing "Fake News" to spread on their platform, every day can feel like April Fools' day. We create a dataset of April Fools' hoax news articles and build a set of features based on past research examining deception, humour, and satire. Analysis of our dataset and features suggests that looking at the structural complexity and levels of detail in a text are the most important types of feature in characterising April Fools'. We propose that these features are also very useful for understanding Fake News, and disinformation more widely.

Keywords: disinformation · deception · April Fools · fake news

1 Introduction

People celebrate April Fools' day each year on April 1st by playing pranks on each other for hilarity's sake. This tradition has transferred over to the traditional media, the most famous example of which is the BBC's 1957 'Swiss Spaghetti Harvest' film[1], which tricked many UK television viewers into believing that a farm in Switzerland grew spaghetti as a crop. With the rise of the web, news sites and companies started releasing annual hoaxes.

In today's world of disinformation and 'Fake News', understanding different forms of deception in news and online media is an important venture. In the 2016 US Presidential Election, dissemination of 'Fake News' was pointed to as one of the crucial factors leading up to Donald Trump's victory and subsequent ongoing tenure as 45th President of the United States of America. If it is true that deceptive news articles swayed a major democratic election, it is certainly important for research towards better understanding and solving of the problem.

One of the main differences between April Fools' articles and typical deceptive texts is the author's intent. The author of an April Fool is not trying to deceive so much as amuse. In this way April Fools' hoaxes are similar to Irony and Satire,

[1] http://news.bbc.co.uk/onthisday/hi/dates/stories/april/1/newsid_2819000/2819261.stm

which expect the reader to understand based on context that what is literally being said is not true. By looking at April Fools' news hoaxes, we investigate whether the change of intent affects the linguistic features of deception in April Fools' compared to in 'Fake News'.

By using April Fools' news hoaxes, we can look at a dataset of verifiable false bodies of text spanning back 14 years. Similar work with satirical news articles has yielded interesting results [26]. While it is true April Fools' hoaxes are not completely similar to 'Fake News', mainly in terms of motivation, our hypothesis is that they will provide insight into the linguistic features put on display when an author is writing something fictitious as if it is factual.

The main contributions of this work are:

- Introducing a new dataset of hoax April Fools' articles.
- Investigating the linguistic features of April Fools' hoaxes, particularly how they relate to features of deception and humour.
- Discussing how these features may be useful in the detection of Fake News.

2 Background

As April Fools' reside in a space somewhere between deception and humour, we will provide a brief background in the areas of deception detection and humour recognition. We will also discuss current NLP approaches to Satire and 'Fake News' detection.

2.1 Deception Detection

Deception research often focusses on 'non-verbal' cues to deception, e.g. eye movement. However, we are interested in the verbal cues to deception, i.e. the features hidden within the text. Without non-verbal cues, humans identify deception with very low degrees of success [12]. Much of the research on verbal cues of deception has been completed in the context of Computer Mediated Communications (CMC). This type of communication can be either spontaneous (synchronous) or preplanned structured prose (asynchronous), such as news, which is of interest for the present research.

Works in synchronous deception detection have involved looking at text from spoken and written answers to questions [17], email [10], and chat-based communication [6, 8]. Carlson et al. [4] provide a good overview of how different factors can affect the deception model, such as medium, the liar's social ability, and the author's motivation. There are certain groups of features that these works suggest are present in deception. One of these groups is 'Cognition Features'. Lying requires a higher level of cognition than telling the truth so often lies seem to be less complicated and more vague. There is also a tendency towards negative emotional language because liars feel guilty about lying. Certain features suggest that liars have more distance from the story they are telling, e.g. reduced pronouns and details.

These works are useful for looking at the linguistic behaviour of liars, but they do not carry over too well to asynchronous communication, where a deceiver can edit and revise what they have written. Toma and Hancock [27], looking at fake online dating profiles, found that certain features of synchronous deception were not present in asynchronous deception. Liars also more frequently exhibited exaggeration of their characteristics. Other works have looked at fake hotel reviews [2, 18]. Features relating to understandability, level of details, writing style, and cognition indicators provided useful clues for identifying fake reviews, though some features may have been genre-dependent. Markowitz and Hancock [11] looked at fraudulent academic writing, an area similar to the news domain where a formalised writing style may mask certain stylistic features of deception. Fake works exhibited overuse of scientific genre words, as well as less certainty and more exaggeration. Stylometric approaches to looking at deception have included Afroz et al. [1] who found that certain style features seemed to leak out even when an author was trying to hide them or imitate someone else. In some ways April Fools' articles are an example of imitation, in which an author is writing a fictional article, mimicking the style of real news.

2.2 Fake News

Conroy et al. [5] provide an overview of computational methods that can be used to tackle the problem of Fake News, including linguistic approaches. Current linguistic research into detecting fake news includes Pérez-Rosas et al. [20] who used features from LIWC [19] for the detection fake news. They found that fake news contained more function words and negations as well as more words associated with insight, differentiation and relativity. Fake News also expressed more certainty and positive language. These results are interesting, but it must be considered that the dataset used was crowdsourced using Amazon Mechanical Turk, meaning the authors of this news were unlikely to be accustomed to writing news articles. Horne and Adali [9] found fake news to be a lot more similar to satire than normal news and also that the title structure and use of proper nouns were very useful for detecting it. Rashkin et al. [21] found that features relating to uncertainty and vagueness are also useful for determining a text's veracity.

2.3 Humour Recognition

Unlike most deceptive texts, April Fools' articles have a motivation of humour. Bringing ideas in from the area of humour recognition therefore may help us characterise hoax articles. Much of the work in humour recognition has focused on detecting humour in shorter texts such as one-liner jokes.

Mihalcea and Strapparava [14] showed that classification techniques can be used to distinguish between humourous and non-humourous texts. They used features such as alliteration, antonymy, and adult slang in conjunction with content features (bag-of-words). Mihalcea and Pulman [13] discussed the significance of 'human-centeredness' and negative polarity in humourous texts. Reyes et al.

[24] looked at a corpus of one-liners and discussed their features. Reyes et al. [25] investigated the features of humour and contrasted to those of irony.

2.4 Irony

Irony is a particular type of figurative language in which the meaning is often the opposite of what is literally said and is not always evident without context or existing knowledge. Wallace [30] suggest that to create a good system for irony detection, one cannot rely on lexical features such as Bag of Words, and one must consider also semantic features of the text. Reyes et al. [25] created a dataset generated by searching for user-created tags and attempted to identify humour and irony. The features used to detect irony were polarity, unexpectedness, and emotional scenarios. More recently, Van Hee et al. [28] investigated annotated ironic tweet corpora and suggested that looking at contrasting evaluations within tweets could be useful for detecting irony. Van Hee et al. [29] also created a system to detect ironic tweets, looking beyond text-based features, using a feature set made up of lexical, syntactic, sentiment, and semantic features.

2.5 Satire

Satire is a form of humour which pokes fun at society and current affairs, often trying to bring something to account or criticise it. This is often achieved using irony and non-sequitur. Satire is similar to April Fools' in that the articles are both deceptive and humourous. The only difference is that satire does tend to be political, whereas April Fools' are often more whimsical.

Burfoot and Baldwin [3] created a system to identify newswire articles as true or satirical. They looked at bag-of-words features combined with lexical features and 'semantic validity'. Rubin et al. [26] used linguistic features of satire to build an automatic classifier for satirical news stories. Their model performed well with an F1-Score of 87% using a feature set combining absurdity, grammar, and punctuation.

3 Hoax Feature Set

The purpose of this work is to identify the features of April Fools' articles, and to see if what we learn is also true of fake news, and possibly disinformation more generally. We want to avoid highly data-driven methods such as bag-of-words because these will learn content and topic-based features of our specific dataset meaning we would not necessarily learn anything about April Fools' or deception more generally. We specifically look at the use of features from the areas of deception detection and humour recognition.

Some previous works have used LIWC [19] to capture Neurolinguistic features of deceptive texts. While we did not use LIWC directly, we did consider important LIWC features from previous work when devising our own features.

For many of our features, we utilise tokenisation and annotations from the CLAWS Part-of-Speech (PoS) tagger [7] and the UCREL Semantic Annotation System (USAS) [22]. The code we used for extracting features, including the output from CLAWS and USAS, are available for reproducibility purposes with the rest of our code[2].

The features we used have been split into seven categories so as to logically group them together to aid analysis and understanding of the results. These categories are: Vagueness, Detail, Imaginative Writing, Deception, Humour, Complexity, and Formality. All features were normalised between 0 and 1.

Vagueness features aim to capture the idea that hoax articles may be less detailed and more ambiguous because the stories are fabricated. Ambiguity was captured by calculating the proportion of words in a text for which there were multiple candidates for annotation. Three types of ambiguity were used: Part-of-Speech Ambiguity, Semantic Ambiguity, and WordNet Synset Ambiguity. Vague descriptions might use more comparative and superlative words as opposed to hard, factual statements [18]. Groups of PoS and Semantic tags were gathered to represent exaggeration, degree, comparative, and superlative words.

Detail features are almost the opposite of vagueness. Genuine news article should contain more details because the events described actually happened. Increased cognition is needed to invent names and places in a text. For this reason we look at the number of proper nouns in a text. Similarly, a fake article may avoid establishing minute details such as dates. We therefore look at Dates, numbers, and Time-related words. Motion words, spatial words, and sense words also establish details that may be less present in deceptive texts.

Imagination features have been used in deception research by Ott et al. [18], based on the work of Rayson et al. [23], which involved comparing informative to imaginative texts. It is worth noting that we are comparing informative texts to pseudo-informative texts, rather than informative to openly imaginative texts. However, they were previously useful in detecting deceptive opinion spam [18], so we evaluate their use here. Rayson et al. [23] identify different PoS tags that are more present in imaginative and informative writing. We used tags that were highlighted from the following PoS groups: conjunctions, verbs, prepositions, articles, determiners, and adjectives.

Deception features are the features of synchronous verbal deception. We include them to investigate if any of the features of spontaneous deception are preserved in spite of a change in medium. Features of asynchronous deception are more relevant to this task and have been distributed between more specific categories, such as Complexity and Details. These synchronous deception features are: First-person pronouns, Negative Emotional Language, and Negations.

Humour features are those from the area of humour recognition. As with deception, some humour features (notably ambiguity) fit better into other categories. The humour features used were: Positive emotion, Relationships, Contextual Imbalance, Alliteration, and Profanity. Contextual Imbalance is characterised as being the average similarity of all adjacent content words in the text.

[2] `https://github.com/dearden/april-fools`

Similarity was calculated by comparing the vectors of words using the in-built similarity function of spaCy[3]. Positive Emotions and Relationships were both gathered using USAS semantic categories. Profanity was gathered from a list of profanities banned by Google [4]. Alliteration was measured by calculating the proportion of bigrams in the text that began with the same letter.

Formality features aim to capture elements of style in news documents that may show how formal they are. April Fools' may be generally less formal or have less editorial oversight. We used three features based on aspects of the Associated Press (AP) style book: AP Number, AP Date, and AP Title Features. These features checked if the text obeyed AP guidelines in their writing of numbers, dates, and titles. An example of an AP guideline is that all numbers under 10 must be spelled out (e.g. 'four' as opposed to '4'). Spelling mistakes were also counted and used as a feature, using the enchant spell checker[5].

Complexity features represent the structure and complexity of an article. They comprise: punctuation, reading difficulty, lexical diversity, lexical density, average sentence length, and proportion of function words. Punctuation was the number of punctuation marks in the text, found using a regular expression. To calculate the reading difficulty feature, we used the Flesch Reading Ease index. We used a list of function words from Narayanan et al. [16].

4 Data Collection

4.1 April Fools Corpus

When building our dataset, the first challenge we faced was finding news articles that were definitely April Fools' hoaxes. One cannot simply collect all news articles from April 1st as the majority of news from this date is still genuine. It is also infeasible to manually go through all news published on this day every year. So instead we utilised a website that archives April Fools' each year[6]. The collection of links published on this site is crowd-sourced so there are some issues arising from the fact that only the popular/amusing hoaxes are uploaded. However, this problem is fairly minor; in fact crowd sourcing may serve to diversify the kinds of website from which hoaxes are sampled. The site archives April Fools' articles from 2004 onwards, providing 14 years of hoaxes.

We used Beautiful Soup [15] to scrape all of the hoax links. We performed some preprocessing to remove hoaxes that one could tell did not constitute a news story from the URL. Next we processed all of the linked webpages, extracting the headline and body of each hoax separately. The wide range of sites in the corpus made automatic scraping too error-prone, so the final approach was largely manual. Efforts were made to ensure no boilerplate or artefacts from

[3] `https://spacy.io/`

[4] `https://github.com/RobertJGabriel/Google-profanity-words/blob/master/list.txt`

[5] `https://github.com/rfk/pyenchant`

[6] `aprilfoolsdayontheweb.com`

the website were included as these could have caused the classifier to pick up features such as the date as being features of April Fools. For the same reason, we also removed any edits to the article disclosing its April Fools' nature.

There were various categories of April Fools' articles found, the most common of which were news stories and press releases. News stories are distinct from press releases which we classed as texts that are self referential; usually taking the form of announcements or product reveals. For example, a press release might be a website announcing that they have been bought out by Google, whereas a news story might be an article by the BBC saying that Google has bought out said company. Press releases were manually filtered out for the present study in order to keep the focus on news, and to avoid the features of press releases obscuring those of April Fools' articles. This resulted in a final April Fools' (AF) corpus comprising of 519 unique texts, spread across 371 websites.

4.2 News Corpus

To create a comparable corpus of genuine news articles, Google News was utilised to automatically scrape news articles from the 4th–5th April of the same years (2004–2018) This time range was chosen so the kinds of topics in the news would be of a similar nature. We will refer to these articles as "NAF" articles. The stories were found using 6 search terms that aimed to catch similar topics to those represented in the AF articles. We did this to avoid learning about the differences in topics of articles rather than whether or not an article is a hoax. These search terms were: "news", "US", "sport", "technology", "entertainment", and "politics". Despite our efforts, it was difficult to match the topic distribution exactly: not all the websites in the AF corpus have archived articles going back to 2004. We acknowledge this is a problem but do not consider it too critical as the features we are looking for are not data-driven and so should not be influenced by topic. We then took all of these URLs and automatically scraped the text using the newspaper python package[7]. Using this method we scraped 2,715 news articles.

For each year (2004–2018), we selected the same number of articles as there were in the AF corpus. The 519 AF articles were spread over 371 websites, the most common of which occurred 19 times. To try and match this distribution, we capped the number of articles that could be taken from any given site at 20. Once we had selected our genuine (NAF) articles, we manually checked the text of each article to ensure that the full text was scraped correctly and that the text only contained the news article itself, without boilerplate noise. We went through the same process as for the AF articles of removing any texts that did not fit in the category of News, such as personal blogs. When an article was removed, we replaced it by choosing a news article from a later page of the Google search that found it. Once this process was finished, we had an NAF corpus of 519 articles spread over 240 websites. Table 1 shows a summary of the corpus, which is made available for further research. April Fools' articles contain

[7] `http://newspaper.readthedocs.io/en/latest/`

Table 1: Summary of April Fools (AF) and Non-April Fools (NAF) corpora.

	Articles	**Websites**	**Avg Words**	**Std Words**
AF	519	371	411.9	326.9
NAF	519	240	664.6	633.2

fewer words on average. Both AF and NAF articles vary significantly from the mean in their lengths.

4.3 Limitations

This is a small dataset and has various notable limitations. The genuine articles tend to be from a smaller pool of more established websites as it is these websites that are more prominent when searching for news online. Only news articles are contained in the dataset. Further work may extend to blogs and press releases. Sometimes the distinction between blogs and news is arbitrary but we tried to be consistent. Multiple genuine news articles occasionally cover the same story, but this was rare and no one story was ever repeated more than twice. A minimum length of 100 characters was enforced to remove anomalous texts such as video descriptions, however this may have removed some genuine articles. While we bear them in mind, we do not see these limitations as major barriers to the research. We will analyse the data using both quantitative and qualitative techniques that allow us to take a deep dive into the data and understand the language being in April Fools' articles for the first time. We do not believe a significantly larger corpus could be built in a reasonable time period.

5 Analysis

5.1 Classifying April Fools'

To evaluate the comparative strength of our feature groups for predicting hoaxes, we used a Logistic Regression classifier with 10 fold cross-validation. We used default parameters of Logistic Regression (from scikit-learn), with standardization to zero-mean and scaling to unit variance ($x' = x - \bar{x}/\sigma$). A basic Logistic Regression classifier serves our needs as we are primarily concerned with investigating the behaviour of features with an interpretable model, and not maximising classification accuracy through tuning or more elaborate classifiers. The results of these classifications can be seen in Figure 1.

From the classification results, we can see that our features provide some information to differentiate between April Fools' hoaxes and genuine news articles. The results are not as high as the F_1-Score of 0.87 found by Rubin et al. [26] for the related task of satire detection, though they are similar to results from fake

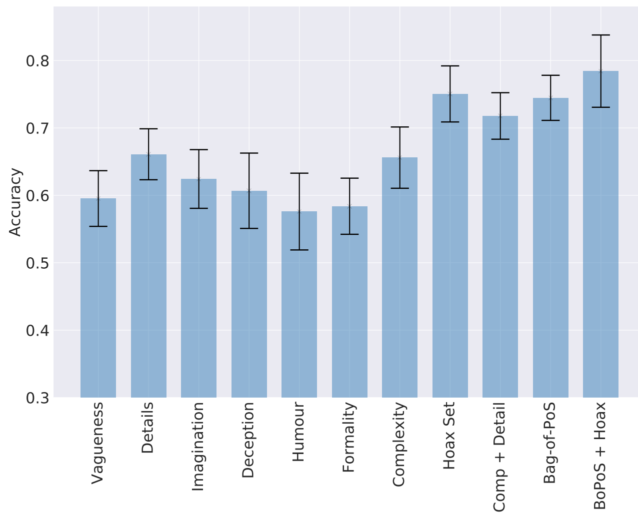


Fig. 1: Mean accuracies of Logistic Regression classifiers across 10 Fold Cross-Validation. Error bars show standard deviation of accuracies across the 10 folds.

news detection, such as those of Horne and Adali [9] who achieved an accuracy of 71% using Bodies of text and 78% using headlines.

Looking at the individual feature groups, Complexity and Detail Features perform best, though not as well as the full Hoax Set. Deception literature suggests that deceptive accounts contain fewer specific details and are generally less complex [4]. Humour performing badly is not surprising as understanding the joke of an AF hoax requires a lot of context and pre-existing knowledge. The features of humour we used in the Humour feature-set were relatively simplistic, more complex, context-aware features may be needed to identify the humour in April Fools' hoaxes. The poor performance of Formality features could suggest that AF Hoaxes are still written to the same journalistic guidelines and standards as their genuine counterparts.

Given the success of the Complexity and Detail features, we classified articles using only these features, achieving an accuracy of 0.718, not far from that of the entire Hoax Set (0.750). This further suggests that looking at details and complexities within a text are crucial when trying to determine if an article is a hoax.

We looked at a non-tailored Bag-of-Part-of-Speech (BoPoS) approach, to compare our curated features to a more data-driven approach. Each PoS tag in CLAWS is used as an individual feature, the occurrences of which are counted for each document. BoPoS was chosen over the more standard Bag-of-Words (BoW) approach because BoW is prone to identifying differences in content and topic, rather than style. BoPoS achieved an accuracy of 0.745, similar to the hoax set. This is not overly surprising as many of the hoax features were part-of-speech counts. These sets do not completely overlap, however. When the hoax set was added to the BoPoS features, the classifier improved its accuracy. This suggests that the non-part-of-speech features in the hoax set provide useful additional information. BoPoS performing similarly to the hoax set therefore suggests that there must be additional PoS frequencies which characterise AF hoaxes.

5.2 Classifying "Fake News"

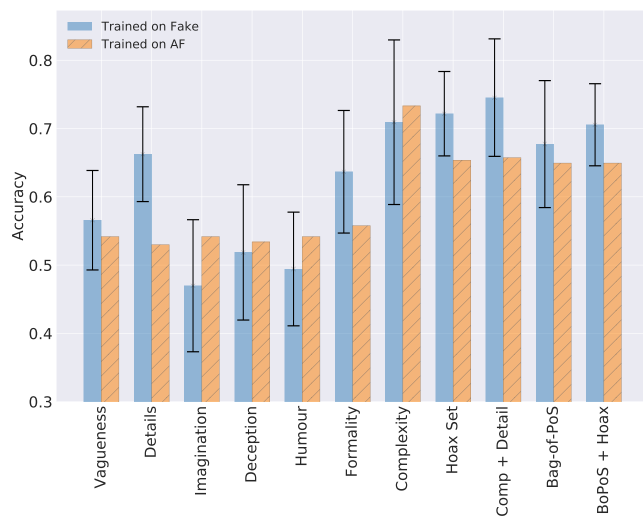


Fig. 2: Accuracies of Logistic Regression classifiers for detecting fake news, trained on Fake News using 10 fold cross-validation and April Fools. Error bars show standard deviation of accuracies across the 10 folds.

Next, we aim to see if we can use the same feature set to effectively identify Fake News. For this we used the fake news dataset introduced by Horne and Adali [9]. This dataset consists of a mixture of articles gathered from well-known fake news outlets and legitimate sites as well as articles gathered by Buzzfeed for an article about fake news in the 2017 election[8]. This is a small dataset (250 articles) split evenly between real and fake. The classification results, again using logistic regression and cross-validation, for fake news can be seen in Figure 2. For each feature set, one classifier was trained on fake news and evaluated using 10 fold cross-validation, and another trained on April Fools' and tested on fake news.

The classifier trained on fake news using the hoax features achieved an accuracy of 0.722, similar to that achieved by the classifier trained on the hoax features for April Fools' (0.750). This suggests that at least some of the features

[8] `https://tinyurl.com/jlnd3yb`

useful for detecting April Fools' hoaxes are also useful in the identification of deceptive news. Complexity features performed well on the fake news dataset (0.709), performing almost as well as the full Hoax Set. Details were useful as before but vagueness features performed significantly less well.

When trained on April Fools and predicting fake news with the Hoax Set, an accuracy of 0.653 was achieved. It is possible that some of the same features are useful but their behaviour is different for fake news. Still, the accuracy is not far off the Hoax Set, so there may be some features that manifest themselves similarly for both AF hoaxes and fake news. Finding these features could provide insight into deception and disinformation more generally.

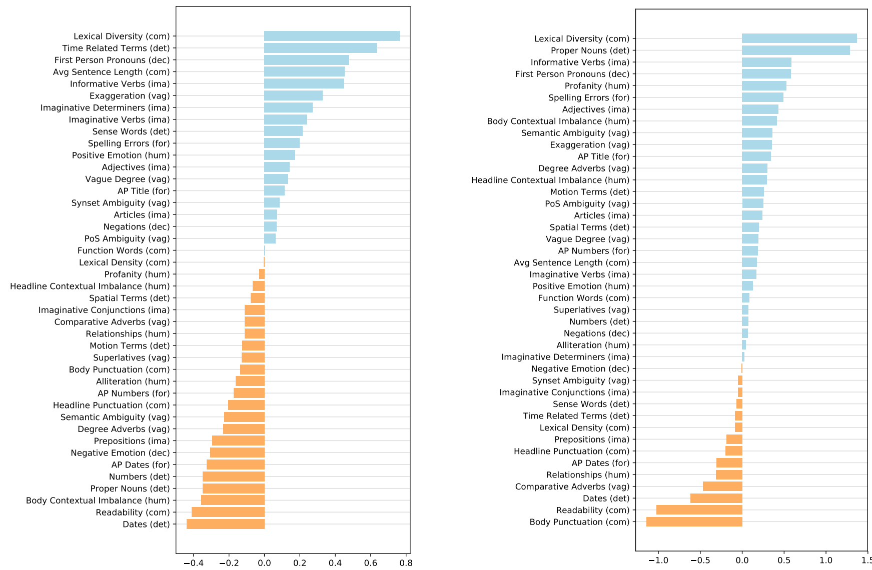
BoPoS performed less well on Fake News, with an accuracy of 0.677, suggesting that PoS tags are not as important when looking at fake news. This, combined with the fact that the hoax features maintained a similar accuracy and complexity features did almost as well as the entire feature set, suggests that the structural features are more important when identifying fake news. BoPoS also did worse when trained on AF and tested on fake, and its drop in accuracy was similar to that of the hoax set. This suggests that there are some PoS tags that are distributed similarly for April Fools and fake news.

5.3 Individual Feature Performances

To see how important individual features were to the classifier, we looked at Logistic Regression weights as shown in Figure 3. For some features it is interesting to see how they are distributed. To this end, frequency density plots for some of our features are provided in Figure 4.

There are differences in structural complexity between AF and genuine articles. Lexical Diversity is the most highly weighted feature. As we can see in Figure 4a, the feature separates hoaxes from genuine articles quite significantly. This could mean hoax texts use more unique words, but it could also be down to the difference in length. High values of lexical diversity correlate to shorter texts and, as we can see in Table 1, the AF articles are shorter, on average. This does still show, however, a difference in complexity. Average sentence length and readability being important features also suggests a difference in complexity. Genuine articles slightly tend towards a shorter average sentence length. NAF articles also tend towards being slightly more difficult to read, though again the difference is not huge.

The story is similar with Fake News – with structural complexity providing key features. Lexical Diversity behaves the same as in April Fools (Fig 4a). This could again be something to do with average document length, but also could suggest a higher proportion of unique words. Reading difficulty also remains important, though the difference in distribution between fake and real is far more prominent, with genuine articles generally more difficult to read. This means that they generally contain longer sentences and words with more syllables. This difference could suggest that fake news articles are more simplistic than genuine texts. Body punctuation, a feature not weighted as highly for AF, appears to

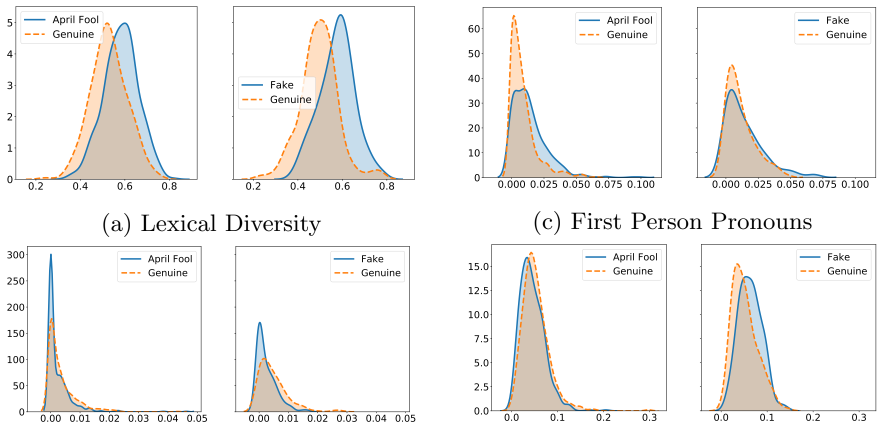


(a) April Fools (b) Fake News

Fig. 3: Logistic Regression weights for the Hoax Set. A large positive weight suggests an important feature of April Fools / Fake News and a large negative weight suggests an important feature of genuine news.

be very important for identifying fake news. More punctuation implies complex structures such as clauses and quotes.

There are also differences in the level of detail between AF and NAF. Genuine articles tend to contain fewer time-related terms. This seems to go against the idea that genuine articles contain more detail. However, if you look at the occurrences of this feature in the text, the most frequent time-based term is 'will'. This combined with the fact that April Fools tended towards fewer dates (Fig 4b) and numbers suggests that AF hoaxes refer to events that will happen, but do so in vague terms. This backs up the idea that April Fools are less detailed and more vague. There are also more references to the present. AF hoaxes seem to be more interested in the present and future than the past. AF hoaxes containing fewer dates is interesting, as one might expect that an AF article would mention the date more than a regular article. This is true as far as references to April are concerned, April Fools had more of those. However, the number of references to the month was roughly the same (April Fools actually had slightly fewer overall), though for genuine news it was spread across more months. This may be because real news stories are the culmination of multiple past events that need to be referenced in the story. More significant than references to months, were references to days of the week. Genuine articles contained many more, which backs up the idea of real texts building more detailed stories.



(a) Lexical Diversity

(c) First Person Pronouns

(b) Dates

(d) Proper Nouns

Fig. 4: Density plots of notable features.

The distribution of proper nouns between AF and NAF is fairly similar (Fig 4d), possibly skewing towards fewer in April Fools. This could suggest fewer details, i.e. names and places, being established in the fake documents. Similarly to complexity, the differences in details are not huge, but do seem to be present.

The detail features do not quite behave the same in Fake News articles as in AF. Proper nouns are one of the most important features for characterising fake articles (Fig 4d). However, unlike AF hoaxes, fake articles tend towards containing more proper nouns than genuine articles. This does not suggest less detail. When looking at the corpus, Fake News articles seem to use a lot of names, often the same ones, such as “Trump” and “Hillary”. Interestingly they massively over use the name “Hillary”, both suggesting that they are less formal (using the lone forename of a politician), and also that they may have an obsession. Dates are the only other detail feature to be weighted highly for fake news. Fig 4b shows that this feature behaves similarly as it did for AF hoaxes, though not as dramatically so. Fake articles are more likely to contain very few, or no, dates. These findings suggest that there are not the same types of difference in detail between AF and fake news, though detail does still hold some significance: it was the second best performing feature group.

Not all the important features link to detail and complexity. First person pronouns were an important feature for both AF hoaxes and fake news. The word ‘we’ was overused in particular by April Fools and to a lesser extent by fake news. This goes against the ideas from traditional deception detection [4] that suggest liars use fewer first person pronouns. In our data, the fake texts use more self-references. This could point towards false articles being more personal and less formal, rather than a feature of deception.

Some of the highly weighted features of fake news are not in common with April Fools. For example, profanity and spelling errors. Both could point towards a reduced level of formality. This would make sense as not being a feature of April Fools. AF writers are usually writing for outlets that publish genuine news, and so likely conform to many of the same standards as genuine news. Fake news, however, comes from less journalistically formal websites.

One of the most obvious differences between April Fools and Fake News in Figure 3 is that Fake News has a smaller group of features that are very important. Lexical diversity, proper nouns, body punctuation, and readability are significantly higher weighted than anything else. Three of these four features relate to structural complexity and the other to detail. This could suggest that, in the case of fake news, the 'fakeness' lies in the structure of the words rather than the words themselves.

Our results suggest that April Fools and Fake News articles share some similar features, mostly involving structural complexity. The level of detail of a document is also important for both AF hoaxes and fake news, though these features do not behave exactly the same way. Some of the features of deception are present in April Fools, notably those relating to complexity and detail but also first person pronouns, though their behaviour is reversed. The basic features of humour we gathered seem to be less important. A more advanced study of the humour would be required to try and identify it within the AF hoaxes. A successful approach would likely require substantial context and world knowledge.

To compare them to the findings from our feature set, and demonstrate how we can gain new insight by looking at features prominent in the data, as well as those from past literature, we looked at some of the PoS tags that were highly weighted by the BoPoS classifier. Some familiar features show up. Certain time-related tags such as 'quasi nominal adverbs of time' (e.g. "now", "tomorrow") and singular weekday nouns (e.g. "Monday") are highly weighted. Proper Nouns are also highly weighted for fake news in particular. Coordinating conjunctions (e.g. "and", "or") are a prominent feature of NAF articles. More coordinating conjunctions implies more detail and complexity. It is good to see that some of the most highly weighted parts of speech back up our finding that detail and complexity are important in defining April Fools' articles and Fake News.

6 Conclusion

In this paper we have introduced a new corpus of April Fools' hoax news articles. We also created a feature set based on past work in deception detection, humour recognition, and satire detection. Using this feature set, we built a system to classify news articles as either April-Fools' hoaxes or genuine articles. The resulting accuracy of 0.750 suggests that the features we identified are useful in identifying April Fools' hoaxes, though not without room for improvement. We then tested our system on a small dataset of fake news to see if April Fools' hoaxes are similar enough to fake news that similar features can be used to detect both. An accuracy of 0.722 was achieved on the Fake News dataset, suggesting that these features are useful for both tasks.

We analysed our features using a combination of qualitative and quantitative techniques to observe the differences between April Fools' hoaxes and genuine articles. This analysis suggests that the structural complexity and level of detail of a text are important in characterising April Fools. This was also the case for Fake News, though structural complexity seemed more important and the changes in details differed slightly from those in April Fools. Our findings suggest that there are certain features in common between different forms of disinformation and that by looking at multiple varieties, we can hope to learn more about the language of disinformation in general. We also showed that by using a mixture of analysis techniques, we can gain far more insight than we can purely from classification. The corpus we have introduced will also be useful in wider fake news research by providing a dataset of news articles which are completely untrue, similar to how satirical news articles are already being used.

Despite similar features being effective at classifying both April Fools' hoaxes and Fake News, we showed that not all these features behave the same way between the two text types. It is possible that some of these differences in feature behaviour come down to the deceptive intent of the texts. April Fools' are an interesting form of disinformation because the author does not believe what they are writing and is not trying to deceive anybody. By looking at a wider variety of false texts, we can further understand the way that the author's motivation and belief affect the way false information is written.

This early work has provided a new dataset for use in the area of Fake News detection and has highlighted directions for future work, describing features useful for detecting April Fools' articles and showing that they may also be present in fake news. Our findings may provide important insight into deceptive news going into the future.

Bibliography

[1] Afroz, S., Brennan, M., Greenstadt, R.: Detecting hoaxes, frauds, and deception in writing style online. In: Proceedings of the 2012 IEEE Symposium on Security and Privacy, SP '12 (2012)

[2] Banerjee, S., Chua, A.Y.K., Kim, J.J.: Using supervised learning to classify authentic and fake online reviews. In: Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication, IMCOM '15 (2015)

[3] Burfoot, C., Baldwin, T.: Automatic satire detection: Are you having a laugh? In: Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, ACLShort '09 (2009)

[4] Carlson, J.R., George, J.F., Burgoon, J.K., Adkins, M., White, C.H.: Deception in computer-mediated communication. Group Decision and Negotiation **13** (2004)

[5] Conroy, N.J., Rubin, V.L., Chen, Y.: Automatic deception detection: Methods for finding fake news. In: Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community (2015)

[6] Derrick, D.C., Meservy, T.O., Jenkins, J.L., Burgoon, J.K., Nunamaker, Jr., J.F.: Detecting deceptive chat-based communication using typing behavior and message cues. ACM Trans. Manage. Inf. Syst. **4**(2) (2013)
[7] Garside, R.: The claws word-tagging system (1987)
[8] Hancock, J., E. Curry, L., Goorha, S., Woodworth, M.: On lying and being lied to: A linguistic analysis of deception in computer-mediated communication **45**, 1–23 (01 2008)
[9] Horne, B.D., Adali, S.: This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. CoRR (2017)
[10] Keila, P.S., Skillicorn, D.B.: Detecting unusual email communication. In: Proceedings of the 2005 Conference of the Centre for Advanced Studies on Collaborative Research, pp. 117–125, CASCON '05 (2005)
[11] Markowitz, D.M., Hancock, J.T.: Linguistic traces of a scientific fraud: The case of diederik stapel. PLOS ONE (2014)
[12] Masip, J., Bethencourt, M., Lucas, G., SEGUNDO, M.S.S., Herrero, C.: Deception detection from written accounts. Scandinavian Journal of Psychology **53**(2), 103–111 (2012)
[13] Mihalcea, R., Pulman, S.: Characterizing humour: An exploration of features in humorous texts. In: Computational Linguistics and Intelligent Text Processing (2007)
[14] Mihalcea, R., Strapparava, C.: Making computers laugh: Investigations in automatic humor recognition. In: Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, HLT '05 (2005)
[15] Nair, V.G.: Getting Started with Beautiful Soup (2014)
[16] Narayanan, A., Paskov, H., Gong, N.Z., Bethencourt, J., Stefanov, E., Shin, E.C.R., Song, D.: On the feasibility of internet-scale author identification. In: 2012 IEEE Symposium on Security and Privacy (2012)
[17] Newman, M.L., Pennebaker, J.W., Berry, D.S., Richards, J.M.: Lying words: Predicting deception from linguistic styles. Personality and social psychology bulletin **29**(5) (2003)
[18] Ott, M., Choi, Y., Cardie, C., Hancock, J.T.: Finding deceptive opinion spam by any stretch of the imagination. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, HLT '11 (2011)
[19] Pennebaker, J.W., Francis, M.E., Booth, R.J.: Linguistic inquiry and word count: Liwc 2001. Mahway: Lawrence Erlbaum Associates **71** (2001)
[20] Pérez-Rosas, V., Kleinberg, B., Lefevre, A., Mihalcea, R.: Automatic detection of fake news. CoRR (2017)
[21] Rashkin, H., Choi, E., Jang, J.Y., Volkova, S., Choi, Y.: Truth of varying shades: Analyzing language in fake news and political fact-checking. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (2017)
[22] Rayson, P., Archer, D., Piao, S., McEnery, A.M.: The ucrel semantic analysis system. (2004)

[23] Rayson, P., Wilson, A., Leech, G.: Grammatical word class variation within the british national corpus sampler. Language and Computers **36** (2002)
[24] Reyes, A., Rosso, P., Buscaldi, D.: Humor in the blogosphere: First clues for a verbal humor taxonomy. Journal of Intelligent Systems **18**(4), 311–332 (2009)
[25] Reyes, A., Rosso, P., Buscaldi, D.: From humor recognition to irony detection: The figurative language of social media. Data Knowl. Eng. **74** (2012)
[26] Rubin, V., Conroy, N., Chen, Y., Cornwell, S.: Fake news or truth? using satirical cues to detect potentially misleading news. In: Proceedings of the Second Workshop on Computational Approaches to Deception Detection (2016)
[27] Toma, C.L., Hancock, J.T.: Reading between the lines: Linguistic cues to deception in online dating profiles. In: Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work, CSCW '10 (2010)
[28] Van Hee, C., Lefever, E., Hoste, V.: Exploring the realization of irony in twitter data. In: LREC (2016)
[29] Van Hee, C., Lefever, E., Hoste, V.: Monday mornings are my fave : #not exploring the automatic recognition of irony in english tweets. In: Proceedings of COLING 2016, 26th International Conference on Computational Linguistics, pp. 2730–2739, ACL (2016)
[30] Wallace, B.C.: Computational irony: A survey and new perspectives. Artif. Intell. Rev. **43**(4) (2015)