

Efficiencies of Feature Engineering in the Machine Learning approach for Fake News Classification

Katrin Donetski

Ward Melville High School, 380 Old Town Rd, East Setauket, NY 11733

Abstract

The rapid infiltration of fake news is a flaw to the otherwise valuable internet, a virtually global network that allows for the simultaneous exchange of information. While a common, and normally effective, approach to such classification tasks is designing a deep learning-based model, the subjectivity behind the writing and production of misleading news invalidates this technique. Deep learning models are unexplainable in nature, making the contextualization of results impossible because it lacks explicit features used in traditional machine learning. This paper emphasizes the need for feature engineering to effectively address this problem: containing the spread of fake news at the source, not after it has become globally prevalent. Insights from extracted features were used to manipulate the text, which was then tested on deep learning models. The original unknown yet substantial impact that the original features had on deep learning models was successfully depicted in this study.

Keywords: Fake news detection, Deep learning, Feature Engineering

1. Introduction

The daunting speed and global scale at which information can be digitally exchanged is a crucial trademark of the 21st century. While seemingly helpful, as this allows individuals to raise awareness to otherwise isolated issues, it, unfortunately, floods the internet with misleading, fake information spread at an equally fast rate. The rapidly growing political divide in our nation, one of the largest in its history, has become a motive for producing fake news. Evident by the results of the 2020 presidential election, a narrow 4 point margin in the popular vote between the candidates of the two major political parties demonstrates the extent of partisan polarization [1]. Most prominent in politics, fake news is used to exaggerate or sometimes entirely fabricate stories to spread a harmful narrative about the opposing party, often targeting specific individuals. By the same token, the spread of fake news drives the two parties further and further apart, when each side is almost solely relying on biased news to make political decisions.

Aside from politics, there are monetary incentives that motivate smaller-scale, and some larger, news outlets to fabricate titles and content to draw a larger, more credulous audience in the long-term. Readers are often inclined to follow news sources that align with their beliefs [2]. Thus, sources will appeal to the biased perspectives of certain audiences to gain the most attention and, ultimately, profit. There is large subjectivity regarding the writing and expansion of fake news, therefore making its classification similarly irregular. Classification algorithms in machine learning are a common approach to this issue: there is a set of data (news) that need to be categorized into classes (true or false). Both traditional machine learning and deep neural networks are used, with the latter yielding higher accuracies. Due to the overabundance of fake news, there is a heavy reliance on such data-driven models using Deep Learning; these models solely find relationships between the system's variables (input and output) - undoubtedly important for classification on a global scale. However, the anonymity of the physical behavior of the system prohibits contextualization of the model because no explicit features are identified. In certain applications, understanding the causal relationship between variables is far more valuable than the mere classification, which is sufficient with medical data. The feature engineering of traditional machine learning methods provide insight into the weighting or quantitative impact of certain features. **Identifying the**

different underlying motives of fake news are the essence of this research, such that its spread can be contained at the source, not after it has impacted elections or reached the opposite side of the world.

2. Related Work

Fake news has been around since the origins of language. In a similar format to its modern implications, individuals have spread false or merely exaggerated rumors, formally known as propaganda, to give themselves a higher ranking. Efforts to spread propaganda have only evolved into more resilient methods. Now with the global network that is the internet, virtually any person can state their opinion on a digital platform for millions of people to see, regardless of credibility or underlying motives.

Due to the longevity of this crisis, its evolution can be observed in full: from its role in ancient conflicts to its current daily presence in social media. This provides the opportunity to identify distinctive characteristics of fake news throughout history and test their validity on modern fake news. Some indicators include longer titles, more all-capitalized words, more proper nouns yet fewer total nouns, fewer stop words [3]. Another supervised machine learning approach extracts features such as simple bag-of-words, n-grams, - the permutations of word combinations, term-frequency - specifically tf-idf that accounts for the significance of the term, not solely just its frequency.

In summary [4], there are currently five main approaches used to classify fake news: language, topic-agnostic, machine learning, knowledge based, and hybrid. The language approach considers the content by analyzing structure of language, grammar, syntax, among others. Topic-Agnostic looks at other features such as number of advertisements, longer headlines, whether author name is included. Machine learning is the traditional statistic based method of training data to improve and fine-tune algorithms. Examples include crowdsourcing, rumor identification, and the Twitter crawler [5]. Knowledge based uses external sources to verify validity of news; for instance, expert oriented fact checking, computational oriented fact checking, crowdsourcing oriented. Lastly, hybrid which combines human and machine learning based methods.

Natural Language Processing systems, a subset of the language approach, are used to understand, analyze, and quantify elements of the subjective human language. It has been observed that Deep Learning (DL), an unsupervised learning subset of Machine Learning (ML), is an approach for various NLP related tasks [6], including fake news detection. Specifically, common Deep Learning approaches are NN, CNN, and RNN, with CNN having an apparent fast performance and “functional evidence for representational learning and feature extraction.” Overall, DL eliminates the need for background knowledge as well as demands of human engineering. Thus, this has been a popular method for fake news. One study [7] reinforces the supremacy of DL classification; it compares three DL models, Long Short-Term Memory (LSTM), Neural Network with Keras (NN-KERAS), and Neural Network with TensorFlow (NN-TF), to two standards supervised ML models, Naive Bayes and Support Vector Machine. As expected, the DL models, specifically LSTM, significantly outperforms with an average accuracy rate of 94.21%, with others following closely, such as KNN with 92.99%. Another study [8] applied a combination of DL approaches, a joint CNN-LSTM (Convolution Neural Net and LSTM) model. It uses the CNN as a “trainable feature detector” for the input text resulting in powerful convolutional features. These are then inputted to the LSTM which generates a description of the content, allowing it to classify or map content onto its correct label. The highest accuracy was lower at 72.25%, which could be credited to the small dataset, however, it was higher than the individual SVM and CNN models, validating the efficacy of combining DL methods. A similar approach, using a [9] CNN-RNN on a much larger dataset yields a 99% accuracy rate. DL methods have undoubtedly brought higher accuracy rates and although the above are very impressive results, they lack a crucial component, the insight and contextualization of features, that is only attainable through traditional feature engineering. There are potential biases in fake news that DL accounts for but because there are no officially determined features, the observer/researcher is unable to understand the distinctive trends and loses the opportunity to find motives of fake news. These hidden motives could shed-light-on how to eliminate fake news at its first appearance, instead of simply filtering it out after it eventually surfaces.

3. Methods

3.1 Dataset and Tools

For the classification purpose, a simple pre-labeled dataset where each text or article has a true/false label is sufficient. The dataset chosen, the ISOT dataset [10, 11], is from the University of Victoria, Canada. It includes real sources from 2016 - 2017 : articles labeled “true” were collected from Reuters.com and “fake” texts were from a variety of sources flagged by PolitiFact - an American organization for fact-checking articles - and Wikipedia [12]. Below is the description of the data from which our features are extracted.

Table 1 – Distribution of ISOT dataset.

News	Total Number of Articles	Subject	Number of Articles
Real	21417	World	10145
		Politics	11272
Fake	23481	Government	1570
		Middle East	778
		US	783
		Left	4459
		Politics	6841
		News	9050

Among standard python libraries (NumPy, pandas, Matplotlib), various additional packages were imported to access certain features: readability, emotion, sklearn, nltk, text2emotion, spacy. Other packages to implement machine learning algorithms include XGBRegressor, Linear Regression, Logistic Regression, Random Forest Classifier, and Sklearn.metric (used to evaluate the accuracy of models).

As seen in Tables 2 and 3, the preliminary details were initially extracted from the texts: article title, text, type and date of publishment.

Table 2 – Real News excerpt.

Real	
Title	As U.S. budget fight looms, Republicans flip their fiscal script
Text	WASHINGTON (Reuters) - The head of a conservative Republican faction in ...
Subject	politicsNews
Date	December 31, 2017

Table 3 – Fake News excerpt.

Fake	
Title	Donald Trump Sends Out Embarrassing New Year's Eve Message; This is Disturbing
Text	Donald Trump just couldn't wish all Americans a Happy New Year and leave it ...
Subject	News
Date	December 31, 2017

The original preprocessing didn't remove the incorrect punctuation and preexisting mistakes in fake news articles. As a result, deep learning methods could extract these details as biased features without considering that the proper punctuation criteria is a minor factor. Many misleading articles are generated by the same

well-respected news outlets that produce truthful articles. Automated grammar checks are applied to both types of articles making this characteristic largely inconsistent, therefore unreliable.

While there was this minimal organization of the data, we applied the following additional steps to prepare the data:

- To start, only each “true” text had its source as the first word(s). This was extracted to create a separate “source” column as seen below. Since “fake” articles didn’t have a source, no feature could be extracted with this information, instead it would only harm other features: readability scores, tf-idf (term frequency), etc.
- Additionally, a small but significant number of rows, mostly among the “fake” articles, were often completely empty in the text column whilst still having an assigned title. Approximately 500 rows of this nature were dropped to prevent any unnecessary errors in the future.
- Approximately 20% of the concatenated dataset was reserved to properly evaluate and fine-tune the algorithms previously modeled on the other 80%, the training dataset. After applying `train_test_split()`, the final training and testing dataset amount to 33817 and 10450, respectively.

Other characteristics, somewhat distinct to either true or false articles, include the subject (World-News, Political-News, Government-News, Middle east, US News, left-news, politics, News), length of the article, and presence of a source. If any of these elements remain, there is potential for biased features and, ultimately, a biased, inaccurate model only viable for this specific dataset.

Stop words, often defined as the most common words of a language, are filtered out as they provide low-level [13] information that could hinder the focus on more important information, such as in the calculation of readability scores. Horne and Adalt [3] introduce the idea that fake news has significantly less stop words; a potential explanation could claim that fake news writers “are attempting to squeeze as much substance into the titles”, necessary to create the desired reaction. A reader, for instance, is overwhelmed by this “substance”, possibly negative details, regarding a politician and is promptly convinced the only “correct” reaction is to oppose said politician. While the stop words provide little meaning, their abundance could reveal a distinction. Thus, there is value in both ridding the text of stop words and leaving them included; two types of datasets, both with and without stop words, were finalized.

3.2 Feature Extraction

The following features were selected for analysis

- **Emotion.** Being that fake news is often written to incite an emotional [14] reaction in the reader, it can be reasoned that quantifying the emotions of truthful and misleading text will yield different results. Ideally, truthful news should be as fact-based and literal as possible if it is to accurately report the newsworthy event. The following emotions - Happy, Angry, Surprise, Sad, Fear - were extracted from the datasets via the `text2emotion` python package. These quantified values were then inputted into Machine Learning methods: Multiple Linear Regression, Logistic Regression, and XgBoost.
- **Readability scores** were calculated for each text to estimate the reading difficulty and other elements of complexity: Gunning Fog score and grade level; Flesch Kincaid score and grade level; Flesch score, ease, and grade level; Coleman Liau score and grade level; Dale Chall score and grade level; Ari score, grade level, and ages; Linsear Write score and grade level; Spache score and grade level; Smog score and grade level. Only numeric values were imputed into the standard models (same as emotion, above).

- **General Features** describe those without package dependencies, including average sentence length and stop word count ratio. The former refers to the average number of words in a sentence of a particular text; the latter is the frequency of stop words in a text, weighted by the total word count of a text.
- **Term Frequency.** A list of all unique words through the entirety of the text was compiled. Then, each word's frequency is calculated to create a list of the 200 most frequent words.

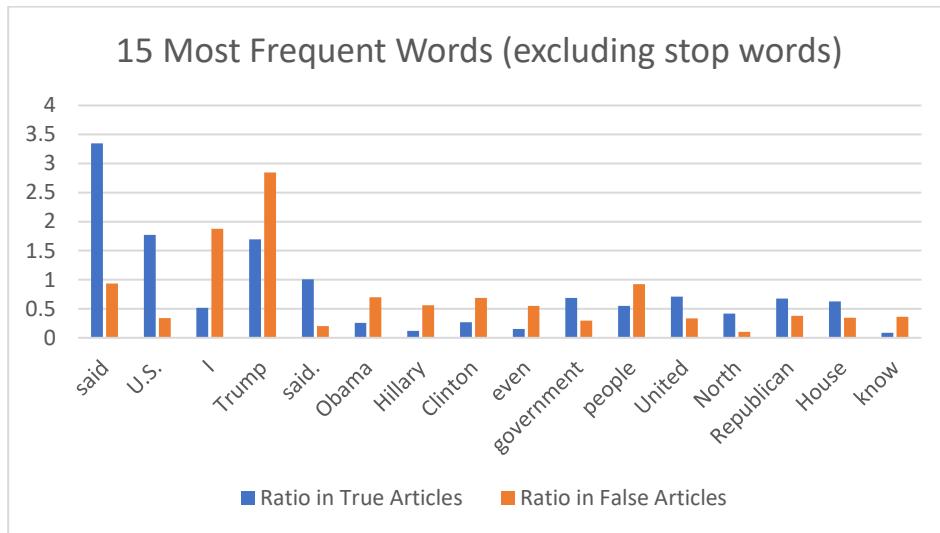


Figure 1 – 15 Most Frequent Words (excluding stop words).

- **Named Entity Recognition.** To classify each unique word, Named Entity Recognition (NER) [15] was applied to locate and classify words into predetermined sets: Cardinal, Date, Event, FAC, GPE, Language, Law, LOC, Money, Norp, Ordinal, Org, Percent, Person, Product, Quantity, Time, Work of Art.

Table 4 – Description of each tag of Named Entity Recognition

Named Entity Recognition		
Tags	Description	Example(s)
Cardinal	Numerals that do not fall under another tag	6
Date	Date entity	2018 or annual
Event	Named hurricanes, battles, wars, sports events, etc.	World Cup
Fac	Facility	Aviara Resort
GPE	Geopolitical Entity (Countries, cities, states)	California
Language	Language entity	English
Law	Named documents made into laws	Copyright Law of US
Loc	Location entity (Non-GPE)	North America
Money	Monetary value	\$78,230
Norp	Nationalities, religious, or political groups	American
Ordinal	Ordinal entities	Second, 4th
Org	Organization entity	Amazon
Percent	Percentage with '%' or 'percent'	40%
Person	Proper noun	George
Product	Vehicles, weapons, foods, etc. (Not services)	Honda SUV
Quantity	Quantity of things such as weight, distance	5 miles away
Time	Times smaller than a day	An hour
Work of Art	Titles of books, songs, etc.	Alchemist

The frequency of each category within an article is extracted as a feature and is similarly inputted into the same models as above.

4. Evaluations and Results

Table 5 - Classification models used for evaluating features.

Classifier	Abbreviation
Multiple Linear Regression	MLR
Logistic Regression	LR
eXtreme Gradient Boosting	XGB
Long Short-Term Memory	LSTM
Recurrent Neural Network	RNN

Accuracy is the percentage of correct predictions over total predictions.

Feature Importance quantifies the usefulness of a feature in predicting a target variable.

Table 6 - Readability Evaluation

Readability Scores	(MLR) Accuracy (%)	(XGB) Feature Importance
Gunning Fog score	49.95	0.03
Gunning Fog grade level	NaN	NaN
Flesch Kincaid score	49.44	0.05
Flesch Kincaid grade level	49.53	0.00
Flesch score	41.37	0.14
Flesch ease	NaN	NaN
Flesch grade level	NaN	NaN
Coleman Liau score	55.95	0.07
Coleman Liau grade level	55.99	0.00
Dale Chall score	60.78	0.16
Dale Chall grade level	NaN	NaN
Ari score	56.45	0.28
Ari grade level	NaN	NaN
Ari ages	NaN	NaN
Linsear Write score	55.96	0.08
Linsear Write grade level	55.90	0.00
Spache score	52.29	0.18
Spache grade level	52.01	0.00
Smog score	NaN*	NaN
Smog grade level	NaN*	NaN
All	--	(XGB) Accuracy (%) 75.96

* Texts do not satisfy the 30-sentence prerequisite

Table 7 - Readability Evaluation

Readability Scores	(XGB) Feature Importance
Dale Chall score	0.31
Ari scores	0.36
Spache score	0.32
All	(XGB) Accuracy (%) 78.03

The calculated readability scores are evaluated above. Most features did not have a substantial impact with the exception of Dale Chall, Ari, and Spache scores (bolded in Table 6). These features are separately reevaluated with XgBoost in Table 7 and yield a marginally higher accuracy.

Table 8 - Emotion Feature Evaluation

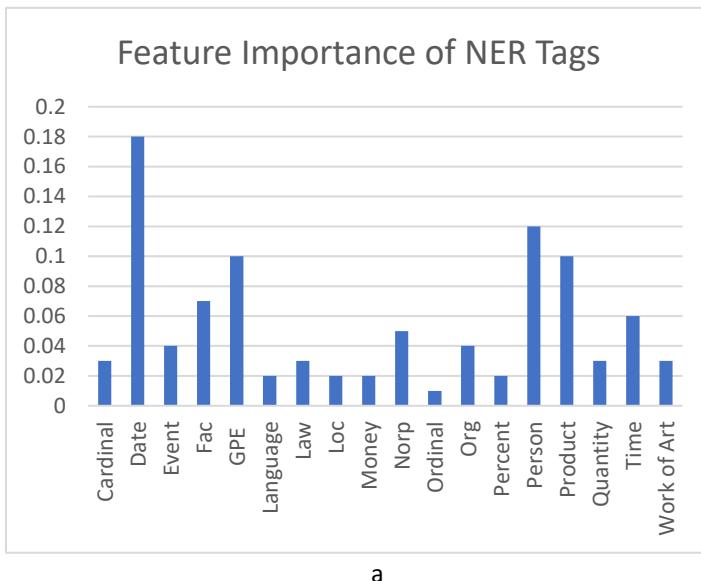
Emotion	Accuracy (%)		
	LR	XGB	XGB (without stop words)
Happy	52.34	50.80	51.57
Angry	52.48	56.49	57.73
Surprise	52.41	51.60	49.15
Sad	52.91	56.29	53.50
Fear	56.86	57.08	56.28
All	59.05	62.04	61.56

In addition to the evaluations of Table 8, when applied to the combination of all emotion features, MLR yields a 58.97% accuracy and a 58.49% accuracy to the text without stop words.

Table 9 - General Feature Evaluation

General Features	(MLR) Accuracy (%)	(XGB) Accuracy (%)	(XGB) Feature Importance
Stop Word Count Ratio	65.67	67.17	0.54
Average Sentence Length	57.21	65.34	0.46
All	69.38	76.44	--

Table 9 depicts the general features that, together, produce at 76.44% accuracy. This exceeds any emotion feature accuracy in Table 8 and the accuracy of the cumulative readability scores in Table 6. However, once the top three weighted readability scores are evaluated (separate from the others), they yield a higher 78.03% accuracy, indicating that there were various counterproductive scores in Table 6, meaning that they were harming the prediction.



Named Entity Recognition	(XGB) Feature Importance
Date	0.32
Person	0.19
Product	0.16
GPE	0.18
Time	0.15
All	(XGB) Accuracy (%) 73.70

b

Named Entity Recognition	(XGB) Feature Importance
Date	0.59
Person	0.28
Product	0.13
All	(XGB) Accuracy (%) 69.35

c

Figure 2 – a) Feature Importance of all NER Tag. b) top 5 (based on feature importance of 2a) tags of NER. c) top 3(based on feature importance of 2a) tags of NER.

Figure 2 depicts Date, Person, Product, GPE, and Time being among the most important features of all NER tags. After applying XgBoost and Multiple Linear Regression models to the cumulative NER features, accuracies of 78.34 % and 73.75%, respectively, were reached. By extracting and modeling the most important features, specifically the top five and three, with XgBoost, decreased accuracies 73.70% and 69.35 % accuracies, respectively, resulted. Thus, the top NER tags aren't as precise as the cumulative of all tags, even those with lesser importance.

Table 10 – Top, Intercategorical Features

Best Features	(XGB) Feature Importance
Date	0.14
Person	0.11
Product	0.08
GPE	0.08
Time	0.06
Stop Word Count Ratio	0.18
Average Sentence Length	0.06
Dale Chall score	0.06
Ari score	0.17
Spache score	0.07
All:	(XGB) Accuracy (%) 88.23

After combing the top yielding features of each category (NER, General Features, and Readability Scores) a new highest accuracy was achieved via XgBoost at approximately 88.23% (Table 10). The most important features are bolded. Additional hyperparameter tuning (of the XgBoost model) concluded at the following parameter quantities: max_depth = 7 and alpha = 5.

5. Analysis of Results

Since pronouns are exempt from the “person” label given by NER, if every name was changed to its respective pronoun (he, she, they), it would still be grammatically correct even if the text loses some clarity. To quantify the impact of this feature, an publicly available, fully developed Long Short-Term Memory (LSTM) deep learning model specifically for classifying fake news was tested [16]. To create a “no-name” dataset on a large scale, all entities labeled “person” in each article were extracted from the original, preprocessed dataset and replaced with the “they” pronoun. On the LSTM model, both the normal and “no-name” datasets were inputted to observe the impact of this one feature on a DL model. The LSTM model has a sense of memory meaning it can learn order dependence and context required to make predictions. Therefore, the only minimal decline in accuracy was expected when testing between the original and “no-name” datasets. However, the LSTM did show a notable change in the prediction distributions, shown in Figure 3 below. While the total averages of predictions from original and the “no-name” test data were similar (0.4738337 to 0.4736906), the plotted distributions of each prediction depict less certainty with the “no-name” test data.

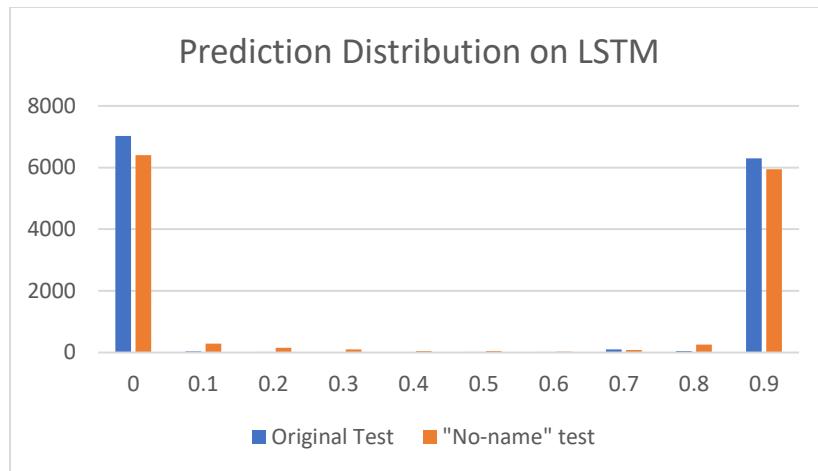


Figure 3 – Distributions of LSTM predictions using original and “no-name” dataset

Another publicly available model developed to classify fake news was tested with the “no-name” test dataset [17]. It uses a Recurrent Neural Network algorithm, similar to the prior LSTM model in that it is also Deep Learning based. Again, there are no explicit features that the model is trained on, leaving little insight into the physical behavior of the system. When tested on randomly selected texts, the model was successful tricked into thinking the input was true when it was, in fact, fake. Specifically, a fake, original article was inputted into the RNN model and achieved a correct, false prediction. Then, when the same article, taken from the “no-name” dataset, was inputted, an opposite prediction resulted: the model classified it as “true.”

In Figure 1, there are isolated words disproportionately frequent in either true or false classes. For instance, there is an abundance of names in fake articles: “Donald”, “Trump”, “Hillary”, “Clinton”, “Obama”. This could be traced back to the fact that a overwhelming majority of fake news is generated for political motives, such as the positive or negative hyperbolization of a politician’s actions

6. Discussion:

Despite the original speculation that emotion is a significant aspect of misleading news, the emotion features produced the lowest prediction accuracies of all other features. However, the readability scores features were significantly reliable features which, when considered, is plausible because misleading news is generated with motives to deceive, not the usual purpose to educate and inform. Therefore, lower readability scores such as less clarity and more primitive vocabulary, sentence structure, and language are likely characteristics of fake news - which is reflected in the results. As for general features such as average sentence length and stop word count ratio, the idea that misleading news often tries to conceal itself as informative could justify the high accuracies. This follows the tendency of fake news articles to have longer, more informative titles. It is an effort of concentrating as much information as possible to distract from the fabricated parts. Since stop words provide little meaning, it can be reasoned that fake news authors naturally avoid stop words as they are more concerned with the details providing instead of coherency and clarity that most news sources prioritize. However, this may not necessarily apply to misleading news spread through social media, which is still a valid type of fake news. NER provides insight into the most distinctive categories between fake and true, including Date, Person, Product, GPE, and Time. Most significantly, the person tag can be explained by the tendency to exaggerate politician’s actions. This is very commonly seen among the two most prominent political parties of the United States, Republican and Democratic. Thus, Figure 1 highlights words such as “Trump”, “Hillary”, “Clinton”, “Obama”, and “Republican”. These are in accordance with the results achieved with the NER features.

When applying the “no-name” test datasets to other LSTM and RNN models, the very slight change in results with LSTM is explained by its prominent “memory” aspect. This capability allows it to quickly

adapt to new types of data. Since RNN does not have this feature, it is subject to overlooking this variation in data and is not able to adapt as well as the LSTM model. Thus, individual deceptions were successful in the RNN. However, the LSTM was still affected in its distribution of predictions. There are increased amounts between the 0.2-0.3 and 0.7-0.8 ranges as pictured in Figure 3. This shows more uncertainty regarding predicting “no-name” test datasets rather than the original test datasets. It is still mostly grammatically correct, yet the effect of the “no-name” dataset is an instance of an unforeseen text that the model was not trained for. Especially in such a digital and rapidly changing world, creating a versatile model is required to tackle the fake news epidemic. This is best achieved by returning to feature engineering, because the context it provides is crucial to comprehending the different underlying motives, which will allow for a greater understanding of how to best solve the issue.

7. Conclusion

The subjectivity of fake news makes it much more difficult to classify than medical data, for instance. Therefore, the heavy reliance on data-driven models using deep learning is a flawed approach to this issue. The deep learning models are not explainable in nature, making the feature engineering aspect of traditional machine learning a more appealing route. As per the technique described in this paper, understanding the origins of the feature-dependent models is vital to contextualizing the results and tackling this urgent issue effectively.

References:

- [1] Igielnik, R., Keeter, S., & Hartig, H. (2021). Behind Biden's 2020 Victory. *Pew Research Center, Washington, D.C.* [Behind Biden's 2020 Victory | Pew Research Center](#)
- [2] Mitchell, A., Gottfried, J., Kiley, J., & Matsu, K. (2014). Political Polarization & Media Habits. *Pew Research Center, Washington, D.C.* [Political Polarization & Media Habits | Pew Research Center](#)
- [3] Horne, B., & Adali, S., (2017) This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News. *arXiv.* [1703.09398.pdf \(arxiv.org\)](#)
- [4] Beer, D., & Matthee, M. (2020). Approaches to Identify Fake News: A Systematic Literature Review. *Integrated Science in Digital Age 2020*, 136, 13–22. [https://doi.org/10.1007/978-3-030-49264-9_2](#)
- [5] Stahl, K. (2018). Fake news detection in social media. *California State University Stanislaus.* [02_stahl.pdf \(csustan.edu\)](#)
- [6] Kapetanois, E., & Alshahrani, S., (2016). Are Deep Learning Approaches Suitable for Natural Language Processing?. *9612.* [http://dx.doi.org/10.1007/978-3-319-41754-7_33](#)
- [7] Alameri, S., & Mohd, M., (2021). Comparison of Fake News Detection using Machine Learning and Deep Learning Techniques. *2021 3rd International Cyber Resilience Conference (CRC)* pp. 1-6, [10.1109/CRC50527.2021.9392458](#)
- [8] Drif, A., Hamida. Z., & Giordano, S., (2019). Fake News Detection Method Based on Text-Features. *The Ninth International Conference on Advances in Information Mining and Management.* [https://www.researchgate.net/publication/335110255_Fake_News_Detection_Method_Based_on_Text-Features](#)
- [9] Nasir, J., Khan. O., & Varlamis, I., (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights, Volume 1, Issue 1.* [https://doi.org/10.1016/j.ijimei.2020.100007.](#)
- [10] Ahmed, H., Traore. I., & Saad, S., (2018). Detecting opinion spams and fake news using text classification. *Journal of Security and Privacy, Volume 1, Issue 1.* [https://doi.org/10.1002/spy2.9](#)
- [11] Ahmed, H., Traoré, I., & Saad, S. (2017). Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques. *ISDDC.* [https://doi.org/10.1007/978-3-319-69155-8_9.](#)
- [12] Wikipedia contributors. (2021). List of fake news websites. In *Wikipedia, The Free Encyclopedia.* [https://en.wikipedia.org/w/index.php?title=List_of_fake_news_websites&oldid=1051356433](#)
- [13] Khanna, C. (2021). *Text preprocessing: Stop words removal / Chetna / Towards Data Science.* Medium. [https://towardsdatascience.com/text-pre-processing-stop-words-removal-using-different-libraries-f20bac19929a](#)

- [14] Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive research: principles and implications*, 5(1), 47. <https://doi.org/10.1186/s41235-020-00252-3>.
- [15] Terry-Jack, M. (2019). NLP: Pretrained Named Entity Recognition (NER). Medium. [NLP: Pretrained Named Entity Recognition \(NER\) | by Mohammed Terry-Jack | Medium](#)
- [16] Makhijani, C., (2020). Fake And Real News Detection. GitHub Repository. <https://github.com/charumakhijani/fake-and-real-news-detection.git>
- [17] Maheshkar, S., (2020). Fake News Classification. GitHub Repository. <https://github.com/SauravMaheshkar/Fake-News-Classification>