FIGURATIVE LANGUAGE DETECTION USING DEEP AND CONTEXTUAL FEATURES

By

MD SAIFULLAH BIN RAZALI

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

July 2023

FSKTM 2023 10

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

FIGURATIVE LANGUAGE DETECTION USING DEEP AND CONTEXTUAL FEATURES

By

MD SAIFULLAH BIN RAZALI

July 2023

Chairman : Alfian bin Abdul Halin, PhD
Faculty : Computer Science and Information Technology

This thesis addresses the scarcity of research focused on deciphering the contextual meaning behind instances of Figurative Language (FL). Existing approaches often neglect the intricate contextual nuances by either relying solely on features extracted through deep learning architectures, abandoning the contextual essence, or resorting to manually extracted features through rigorous processes, with limited exploration of combinatory methods.

The research identifies a critical gap in the literature concerning the application of well-established Machine Learning classification models, such as Support Vector Machine, K-Nearest Neighbor, Logistic Regression, Decision Tree, and Linear Discriminant Analysis, in the context of FL detection tasks. This study aims to bridge this gap by conducting an in-depth exploration of the effectiveness of these models in discerning Figurative Language instances.

Furthermore, the thesis critiques prior works employing manually crafted features for Figurative Language detection, noting the lack of precision in identifying the most crucial features. The research introduces a novel approach by combining features extracted from a Convolutional Neural Network (CNN) model with manually extracted features obtained from well-known lexicons. This integration aims to enhance the robustness and accuracy of Figurative Language detection by leveraging the strengths of both deep learning and traditional feature extraction methods.

The experimental design involves the use of a word-embedding technique, a CNN model, and various well-known machine learning classification techniques. The study not only investigates the efficiency of the proposed methodology but also delves into the importance of individual features, providing precise insights and discussions on the significance of lexicons used in the process. The findings of this research contribute to the advancement of Figurative Language detection methods, offering a more nuanced understanding of contextual meanings and paving the way for future research in this domain.

Keywords: Machine Learning, Deep Learning, Figurative Language, Sarcasm Detection, Metaphor Detection, Satire Detection

SDG: Industry, Innovation and Infrastructure, Reduced Inequalities, Sustainable Cities and Communities, Responsible Consumption and Production, Peace, Justice and Strong Institutions

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

PENGESANAN BAHASA KIASAN MENGGUNAKAN CIRI-CIRI MENDALAM DAN KONTEKSTUAL

Oleh

MD SAIFULLAH BIN RAZALI

Julai 2023

Pengerusi : Alfian bin Abdul Halin, PhD
Fakulti : Sains Komputer dan Teknologi Maklumat

Tesis ini membincangkan kajian yang difokuskan pada makna kontekstual di sebalik Figurative Language ataupun Bahasa Kiasan. Pendekatan sedia ada sering mengabaikan nuansa kontekstual yang rumit dengan hanya bergantung kepada ciri-ciri yang diekstrak melalui kaedah Deep Learning, meninggalkan nuansa kontekstual, atau menggunakan ciri-ciri yang diekstrak secara manual melalui proses yang rumit dan memakan masa.

Kajian ini mengenal pasti jurang kritikal dalam literatur dan juga mengaplikasikan model klasifikasi Machine Learning, seperti Support Vector Machine, K-Nearest Neighbour, Logistic Regression, Decision Tree, dan Linear Discriminant Analysis, dalam konteks Bahasa Kiasan. Kajian ini bertujuan untuk menyelesaikan jurang ini dengan menjalankan kaedah penerokaan yang mendalam.

Selain itu, tesis ini mengkritik kerja-kerja terdahulu yang menggunakan ciri-ciri yang dibentuk secara manual untuk pengesanan Bahasa Kiasan dan mencatat kekurangan ketepatan dalam mengenal pasti ciri-ciri yang paling penting. Kajian ini memperkenalkan pendekatan baru dengan menggabungkan ciri-ciri yang diekstrak dari model Convlutional Neural Network (CNN) dengan ciri-ciri yang diekstrak secara manual dari leksikon-leksikon. Integrasi ini bertujuan untuk meningkatkan kebolehtahan dan ketepatan pengesanan Bahasa Kiasan dengan memanfaatkan kelebihan kedua-dua kaedah Deep Learning dan kaedah ekstraksi ciri tradisional.

Reka bentuk eksperimen melibatkan penggunaan teknik word-embedding, model CNN, dan pelbagai teknik klasifikasi pembelajaran mesin yang tradisional. Kajian ini tidak hanya menyelidiki kecekapan metodologi yang dicadangkan, tetapi juga merinci kepentingan ciri-ciri individu, menyediakan pandangan yang tepat dan perbincangan mengenai kepentingan leksikon yang digunakan dalam proses tersebut. Hasil kajian ini

menyumbang kepada kemajuan kaedah pengesanan Bahasa Kiasan, menawarkan pemahaman kontekstual yang lebih halus, dan membuka jalan bagi penyelidikan masa depan dalam domain ini.

Kata Kunci: Pembelajaran Mesin, Pembelajaran Mendalam, Bahasa Kiasan, Pengesanan Sindiran (Sarcasm Detection), Pengesanan Metafora (Metaphor Detection), Pengesanan Satira (Satire Detection)

SDG: Industri, Inovasi dan Infrastruktur, Mengurangkan Ketidaksamaan Bandar dan Komuniti Mampan, Penggunaan dan Pengeluaran Bertanggungjawab, Kedamaian, Keadilan dan Institusi yang Kuat

ACKNOWLEDGEMENTS

This thesis would not be possible without the grace and guidance of Allah SWT and Prophet Muhammad SAW. Hereby, I would also like to acknowledge the roles of individuals for their invaluable time spent on me:

I would like to express my gratitude towards my supervisory committee; Dr. Alfian Abdul Halin, Dr. Casey Chow, Dr Noris Mohd Norowi and Dr. Shyamala A/P C. Doraisamy who have been attending to all my questions and concerns.

Special thanks to my support circle; my family especially my wife, Nor Afizah Mustafa, my father, Dr Razali Shahul Hameed, my mother, Noriyah Othman, and my mother-in-law, Rohani Abdul Manaf. Also, to all my friends.

All of you have kept me going.

Thank you.

This thesis was submitted to the Senate of the Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Alfian bin Abdul Halin, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Noris binti Mohd Norowi, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Shyamala a/p C. Doraisamy, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Yang-Wai Chow, PhD
Associate Professor
Faculty of Engineering and Information Sciences
University of Wollongong
Australia
(Member)

ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 18 April 2024

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

The size of data shared over the Internet today is tremendous. A big part of the whole bulk comes from short-text posts in social networking sites such as Twitter and Facebook. Some of it also comes from online news sites such as Cable News Network (CNN) and The Onion. This type of data is very good for data analysis since they are very personal (Ghosh and Veale, 2017). For years, researchers in the academia and various industries have been analysing this type of data. The purpose includes for product marketing, event monitoring and trend analysis (Aquino, 2012; Liu, 2012). This field is called sentiment analysis (Liu, 2012).

However, the writer of such posts has no obligation to stick to only literal language. This gives them freedom to also use figurative language (FL). Hence, online posts can be categorized into two: literal or figurative. Literal posts commonly contain traditional words that can be found in a dictionary with no other meaning that what it is intended to. On the contrary, figurative posts may contain words or phrases that carries a different meaning than the standard. This could flip the whole polarity of a given post. Consider the sentence "I am very sad to see a corrupt leader thrown to prison". If the sentence is taken as literal when it is not supposed to, all the evaluations of algorithms would decrease. This problem would be more severe when there are many instances like this in the dataset.

Due to this nature, it can jeopardize a sentiment analysis work that only focuses on polarity of posts. This makes the work of FL detection a non-trivial task and one of the biggest problems in sentiment analysis (Joshi et al., 2017b; Abulaish et al., 2020). Hence, detecting it would be crucial and significant. There are existing works that tried to tackle the problem of FL by specific types. Most of it is using rule-based, lexicon and seeding techniques. All of these ways are very rigorous, and did not even come close to the results yielded by the works using deep learning. However, deep learning is also having a big problem by being automated. It is that contextual meaning behind every FLs are left out.

In the real world, FL is used a lot in everyday conversations to convey ideas which is difficult to visualize without it (Roberts and Kreuz, 1994). Even with this realization, most researchers in this field is still focusing on literal language (Chakrabarty et al., 2022). This is because of all the reasons stated previously which also makes FL very implicit and nature (Shutova et al., 2010) and again the task of detecting it would be very important.

There are essentially seven commonly used FL categories: sarcasm, irony, simile, metaphor, satire, humor and hyperbole (Abulaish et al., 2020). However, some of these concepts are overlapping or the distinction between one another is not very clear.

For example, sarcasm is sometimes considered the same as irony (Filatova, 2012; Joshi et al., 2015). However, there is a slight difference. Sarcasm is proven as a contrast between positive sentiment and negative situation (Riloff et al., 2013). In the case of irony, it is proposed that the speaker of the utterance would pretend to be an unwise person (Clark and Gerrig, 1984). This theory is then supported by another work (Kumon-Nakamura et al., 2007).

Simile and metaphor have the exact same structure apart from the use of words like "like" and "as" (Qadir et al., 2016). For example the sentence "My office is like the Antarctica" would mean that "My office" is very cold since Antarctica is known for being a cold place. A metaphoric instance of the same sentiment would be use the same sentence but without the word "like".

In the case of humor, it is fully dependent on culture and language for it to be understood (Driessen, 2015). For example, the sentence "Your friend is green" could be perceived as funny in the Malaysian culture, but not anywhere else. Hyperbole, is just a sentence with exaggeration (McCarthy and Carter, 2004). Consider the sentence "There is millions of them". It would be agreed that this sentence does not carry the literal meaning for the word "millions". Instead, the speaker is just trying to portray high volume.

In comparison to the other FL types mentioned previously, satire is the most unique in terms of data collection. This is because satire does not only happen in short sentences, but also throughout paragraphs or books. It is a critiquing technique used upon a particular scenario or situation. For example, the book Gulliver's Travels is a critique on the writers of travelogues that are persistent on making their travels sound unique in their books (Orwell, 1967). In Gulliver's Travels, the author writes about his encounter with giants and very tiny people (Swift et al., 1959). These giants and very tiny people does not actually exist. All of the explanations of the FL types above bring this work to its main focus. It is to only detect three categories of FL: sarcasm, metaphor, and satire. There are existing works done to detect each of the FL types. The main difference between the works are the methods used. Some works used rule-based techniques and some works used deep learning techniques. It is very seldom that a work is using the combination of both. All of these existing methods are thoroughly explained in the literature review chapter of this work.

1.2 Research Motivation

Today, a large amount of data is accessible by basically anyone anywhere. This information covered various topics and sentiments. They could be very valuable to decision makers in the industry or the academia. They could even solve existing problems in these organizations. However, the data themselves can pose multiple problems. One of the biggest problems is when the data is using FL or having different meanings then what they are supposed to mean. For example, in the case of sentences or words used to convey sarcasm, metaphor or satire to the receiver, a machine could hardly detect the sender's real intention. The task to analyze and come up with solid verdict would be more challenging. This research aims to detect the three biggest categories of FL in text: sar- casm, metaphor, and satire. The reason being is that when they are detectable, they can be analyzed correctly. The information gathered from the analysis done on a well understood data would be much more valuable. Another inter- esting motivation is to investigate whether the contexts of all these FL categories would bring any points to the sheer detection of them.

1.3 Problem Statement

One of the biggest problems in sentiment analysis is FL detection, due to its ability to flip the polarity of instances. This would jeopardize the evaluations for each algorithms used in experiments. There are many categories of FL and differentiating one to another is not a trivial task. This is due to each of them being contextually unique.

Hence, several research in the domain of FL detection have been reviewed and summarized as below:

1. Sarcasm detection: It is found that most of recent works use supervised learning approach to tackle the issues in sarcasm (Poria et al., 2016; Misra and Arora, 2019; Yin et al., 2021). Although some of them uses semi- supervised (Davidov et al., 2010; Tsur et al., 2010; Lukin and Walker, 2017; Riloff et al., 2013; Ghosh and Veale, 2016; Bharti et al., 2015; Bouazizi and Ohtsuki, 2016) or rule-based approaches (Veale, 2012; May- nard and Greenwood, 2014; Bamman and Smith, 2015; Ghosh et al., 2015; Rajadesingan et al., 2015; Schifanella et al., 2016), the performance of these works are not as good as the ones using supervised learning works mentioned earlier. This makes the two latter approaches almost outdated. However, the context of sarcasm in each of the sentences in the dataset is abandoned in a given supervised learning setting.
2. Metaphor detection: Even though this domain has been studied a lot from the perspective of psychology, linguistics, sociology, anthropology, and computational linguistics (Mohler et al., 2013), metaphor detection remains one of the most challenging tasks. Metaphor is deemed by most researchers as very challenging to understand (Shutova et al., 2010). There have been works done on building a metaphor dataset with seeding tech- nique (Shutova et al., 2010; Mason, 2004). They collected metaphoric instances manually and used these seeds to create a larger corpus while preserving the same sentence structure. However, the seeding technique is very rigorous. No work has been done to fully automate the detection or collection of metaphor instances. This is partly because there is not much work done on building models to understand the contextual meaning behind metaphorical instances.
3. Satire detection: Satire in computational perspective are rare (Abulaish et al., 2019). Apart from being used as an indicator in the works of fake news detection (Barbieri et al., 2015b; Rubin et al., 2016; Guibon et al., 2019), there are hardly credible works on satire detection. It is mentioned that the acceptance of satire among the writers and the readers is based on commonality, a lot like sarcasm (Frye, 1944). The main issue faced in satire is also the same as sarcasm, which is the detection of context and intent for each of the instances in the datasets. However, the contexts of sarcasm and satire remain different. Even though the same framework can be used to detect these categories, different

datasets have to be used to train the models to find the contexts.

As a result of the thorough reviews done on existing works, the problem statements for this thesis are found and finalized as below:

1. The existing works either only use features that are extracted by deep learning architectures that abandon the contextual meanings of instances or use manually extracted features through rigorous processes. Combina- tory works are largely lacking.
2. Exploratory work on using well-known Machine Learning classification models such as Support vector machine K-Nearest Neighbor, Logistic Re- gression, Decision Tree, and Linear Discriminant Analysis in relation to FL detection tasks is predominantly absent.
3. Current approaches that used manually hand-crafted features for Figura- tive Language detection did not focus on on which features are the most important and which features are not, with preciseness and discussions on lexicons used.

1.4 Research Objectives

FL detection is an actively studied domain where several approaches and tech- niques have been used to support the task. However, it is done on each category respectively i.e. sarcasm detection, metaphor detection and satire detection. In this work, the general objective is to devise the detection approach for three categories of FL. To achieve this general objective, four objectives have been outlined:

1. To propose the most essential and justified contextual-based feature sets for each of the FL category detection tasks (sarcasm detection, metaphor detection, satire detection).
2. To design effective combinations of features extracted by a deep learning architecture and manually extracted features.
3. To prove that there are contextually important features of each FL type, even while combining with features extracted from a deep learning archi- tecture.

1.5 Research Scope

The research scope is as the following:

1. Only text modality is covered in this work.
2. Only three FL is chosen for this work: Sarcasm, Metaphor, Satire. This is due to the fact that Sarcasm and Irony are very similar, apart from that Sarcasm is more clearly defined in literature. Metaphor and Simile is also very similar, apart from that Simile is only adding the words “like” and “as” to instances of Metaphor. Hyperbole only mean sentences that uses exaggeration such as

multiple exclamation marks and Humor is very dependent on the culture it is used in. For these reasons this work is not experimenting on Irony, Simile, Hyperbole and Humor.

3. Only data that is available online is used. This is to ease the process of comparing the performance of this work with the existing works in the same domain.
4. There are over 7000 languages that exists in the world today. Each of the languages has their own way of conveying messages from speakers to listeners or writers to readers. Since English is the most spoken language with 1132 million speakers, it is the only language chosen for this work.

1.6 Research Contributions

This thesis has made the following contributions:

1. Feature extraction methods that are based on the contextual justifications of each FLs.
2. A proof that using the combination of a deep learning architecture with carefully crafted contextual features can optimize the performance of nat- ural language processing (NLP) tasks.
3. A thorough comparison of the performances of machine learning classifiers post-experimentation on each FL detection tasks.
4. A thorough findings of the most to the least important feature sets for every FL category.

1.7 Thesis Organization

This thesis contains seven chapters. Chapter two will describe in detail the literature review. The background of existing research in FL detection: i. sar- casm detection, ii. metaphor detection, iii. satire detection and techniques that have been used will be identified and described in detail. This chapter will also highlight important criteria that an FL detection work should possess. Finally, this chapter also discuss the issues posed by the relevant literature.

Next, chapter three will describe the research methodology that has been used throughout the work. Generally, this chapter will explain all the activities involved in defining the significant components, classification, and evaluation of the framework.

The fourth chapter will discuss a new framework for FL detection using deep learning with contextual features. In this chapter, the focus is on sarcasm detection. The fifth chapter focuses on metaphor detection. The sixth chapter focuses on satire detection.

Finally, the seventh chapter will present the conclusion of this research and future works recommended.

1.8 Summary

This research is an integration of figurative language detection with its contex- tual features as well as machine learning. It highly contributes to identify the useful methods for the specific task of the detection. This chapter presents the essence of the thesis; issues faced, the motivation for this work and the main contributions. In the next chapters, more details will be given to the techniques, models and the experimental analysis.

REFERENCES

Abulaish, M. and Kamal, A. (2018). Self-deprecating sarcasm detection: an amalgamation of rule-based and machine learning approach. In *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pages 574–579. IEEE.

Abulaish, M., Kamal, A., and Zaki, M. J. (2020). A survey of figurative lan- guage and its computational detection in online social networks. *ACM Transactions on the Web (TWEB)*, 14(1):1–52.

Abulaish, M., Kumari, N., Fazil, M., and Singh, B. (2019). A graph- theoretic embedding-based approach for rumor detection in twitter. In *IEEE/WIC/ACM International Conference on Web Intelligence*, pages 466–470.

Adel, H. and Schu¨tze, H. (2016). Exploring different dimensions of attention for uncertainty detection. *arXiv preprint arXiv:1612.06549*.

Agarap, A. F. (2018). Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*.

Al Bataineh, A. and Kaur, D. (2018). A comparative study of different curve fit- ting algorithms in artificial neural network using housing dataset. In *NAE- CON 2018-IEEE National Aerospace and Electronics Conference*, pages 174–178. IEEE.

Aljadaan, N. (2018). Understanding hyperbole. *Arab World English Journal (October, 2018), Theses ID*, 212.

Amir, S., Wallace, B. C., Lyu, H., and Silva, P. C. M. J. (2016). Modelling context with user embeddings for sarcasm detection in social media. *arXiv preprint arXiv:1607.00976*.

Aquino, J. (2012). Transforming social media data into predictive analytics. *CRM Magazine*, 16(11):38–42.

Attardo, S., Eisterhold, J., Hay, J., and Poggi, I. (2003). Multimodal markers of irony and sarcasm. *Humor*, 16(2):243–260.

Bamman, D. and Smith, N. (2015). Contextualized sarcasm detection on twit- ter. In *proceedings of the international AAAI conference on web and social media*, volume 9, pages 574–577.

Barbieri, F., Ronzano, F., and Saggion, H. (2015a). Do we criticise (and laugh) in the same way? automatic detection of multi-lingual satirical news in twitter. In *Twenty-Fourth International Joint Conference on Artificial In- telligence*.

Barbieri, F., Ronzano, F., and Saggion, H. (2015b). Is this tweet satirical? a computational approach for satire detection in spanish. *Procesamiento del Lenguaje Natural*, (55):135–142.

Barbieri, F., Saggion, H., and Ronzano, F. (2014). Modelling sarcasm in twitter, a novel approach. In *Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 50–58.

Baruah, A., Das, K., Barbhuiya, F., and Dey, K. (2020). Context-aware sarcasm detection using bert. In *Proceedings of the Second Workshop on Figurative Language Processing*, pages 83–87.

Baumgartner, J. C. and Lockerbie, B. (2018). Maybe it is more than a joke: Satire, mobilization, and political participation. *Social Science Quarterly*, 99(3):1060–1074.

Beardsley, M. C. (1962). The metaphorical twist. *Philosophy and phenomeno- logical research*, pages 293–307.

Bharti, S. K., Babu, K. S., and Jena, S. K. (2015). Parsing-based sarcasm sen- timent recognition in twitter data. In *2015 IEEE/ACM International Con- ference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 1373–1380. IEEE.

Bharti, S. K., Vachha, B., Pradhan, R., Babu, K. S., and Jena, S. K. (2016). Sarcastic sentiment detection in tweets streamed in real time: a big data approach. *Digital Communications and Networks*, 2(3):108–121.

Bigi, A., Plangger, K., Bonera, M., and Campbell, C. L. (2011). When satire is serious: how political cartoons impact a country's brand. *Journal of Public Affairs*, 11(3):148–155.

Biles, Z. P. (2002). Intertextual biography in the rivalry of cratinus and aristo- phanes. *American journal of philology*, 123(2):169–204.

Biship, C. M. (2007). Pattern recognition and machine learning (information science and statistics).

Bizzoni, Y. and Ghanimifard, M. (2018). Bigrams and bilstms two neural net- works for sequential metaphor detection. In *Proceedings of the Workshop on Figurative Language Processing*, pages 91–101.

Black, M. et al. (1979). More about metaphor. *Metaphor and thought*, 2:19–41. Bogel, F. V. (2019). *The difference satire makes*. Cornell University Press.

Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146.

Booker, M. K. (1995). *Flann O'Brien, Bakhtin, and Menippean Satire*. Syracuse University Press.

Bouazizi, M. and Ohtsuki, T. O. (2016). A pattern-based approach for sarcasm detection on twitter. *IEEE Access*, 4:5477–5488.

Brant, W. (2012). Critique of sarcastic reason: the epistemology of the cognitive neurological ability called "theory-of-mind" and deceptive reasoning.

Burfoot, C. and Baldwin, T. (2009). Automatic satire detection: Are you having a laugh? In *Proceedings of the ACL-IJCNLP 2009 conference short papers*, pages 161–164.

Cabitza, F. and Dal Seno, B. (2005). Djess-a knowledge-sharing middleware to deploy distributed inference systems. In *WEC (2)*, pages 66–69. Citeseer.

Campbell, J. D. (2012). Investigating the necessary components of a sarcastic context.

Campbell, J. D. and Katz, A. N. (2012). Are there necessary conditions for inducing a sense of sarcastic irony? *Discourse Processes*, 49(6):459–480.

Cano Mora, L. (2003). At the risk of exaggerating: how do listeners react to hyperbole?

Carlyle, T. (2020). Note on the text. In *Sartor Resartus*, pages XCV–CXLIV. University of California Press.

Carrol, L. (1939). *Alice in Wonderland and Through the Looking-glass*. Lulu. com.

Cason, H. (1930). Methods of preventing and eliminating annoyances. *The Journal of Abnormal and Social Psychology*, 25(1):40.

Chakrabarty, T., Choi, Y., and Shwartz, V. (2022). It's not rocket science: In- terpreting figurative language in narratives. *Transactions of the Association for Computational Linguistics*, 10:589–606.

Chatwin, B. (2012). *In Patagonia*. Random House.

Cheang, H. S. and Pell, M. D. (2009). Acoustic markers of sarcasm in cantonese and english. *The Journal of the Acoustical Society of America*, 126(3):1394– 1405.

Chen, A. and Boves, L. (2018). What's in a word: Sounding sarcastic in british english. *Journal of the International Phonetic Association*, 48(1):57–76.

Christmann, A. (1994). Least median of weighted squares in logistic regression with large strata. *Biometrika*, 81(2):413–417.

Clark, H. H. and Gerrig, R. J. (1984). On the pretense theory of irony.

Cohen, G., Afshar, S., Tapson, J., and Van Schaik, A. (2017). Emnist: Extend- ing mnist to handwritten letters. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 2921–2926. IEEE.

Condren, C. (2012). Satire and definition. *Humor*, 25(4):375–399.

Dauphin, Y. N., Fan, A., Auli, M., and Grangier, D. (2017). Language modeling with gated convolutional networks. In *International conference on machine learning*, pages 933–941. PMLR.

Davidov, D., Tsur, O., and Rappoport, A. (2010). Semi-supervised recogni- tion of sarcasm in twitter and amazon. In *Proceedings of the fourteenth conference on computational natural language learning*, pages 107–116.

Day, A. and Thompson, E. (2012). Live from new york, it's the fake news! sat- urday night live and the (non) politics of parody. *Popular Communication*, 10(1-2):170–182.

De Cervantes, M. (2016). *Don Quixote*. Lulu. com.

del Pilar Salas-Z´arate, M., Paredes-Valverde, M. A., Rodriguez-Garc´ıa, M. A´., Valencia-Garc´ıa, R., and Alor-Hern´andez, G. (2017). Automatic detection of satire in twitter: A psycholinguistic-based approach. *Knowledge-Based Systems*, 128:20–33.

Devitt, M. and Hanley, R. (2008). *The Blackwell guide to the philosophy of language*. John Wiley & Sons.

Diehl, N. (2013). Satire, analogy, and moral philosophy. *The Journal of Aes- thetics and Art Criticism*, 71(4):311–321.

Do Dinh, E.-L. and Gurevych, I. (2016). Token-level metaphor detection using neural networks. In *Proceedings of the Fourth Workshop on Metaphor in NLP*, pages 28–33.

Dong, S., Wang, P., and Abbas, K. (2021). A survey on deep learning and its applications. *Computer Science Review*, 40:100379.

Driessen, H. (2015). Humor, anthropology of. *Wright, JD (ed.), International encyclopedia of the social & behavioral sciences (2nd ed.)*, pages 416–419.

Ebrahimi, M., Yazdavar, A. H., and Sheth, A. (2017). Challenges of sentiment analysis for dynamic events. *IEEE Intelligent Systems*, 32(5):70–75.

Elman, J. L. (1990). Finding structure in time. *Cognitive science*, 14(2):179–211. englishclub.com (2021a). Iverb. https://www.englishclub.com/vocabulary/irregular-verbs-list.htm. Accessed: 2021-02-21. englishclub.com (2021b). Verb. https://www.englishclub.com/vocabulary/regular- verbs-list.htm. Accessed: 2021-02-21.

Erwin, T. (2008). Menippean satire reconsidered: From antiquity to the eigh- teenth century.

Expectans, R. (1906). Mr. winston churchill and democracy. *Westminster re- view, Jan. 1852-Jan. 1914*, 165(1):15–21.

Fellbaum, C. (2010). Wordnet. In *Theory and applications of ontology: computer applications*, pages 231–243. Springer.

Fife, J. (2016). Peeling the onion: Satire and the complexity of audience re- sponse. *Rhetoric Review*, 35(4):322–334.

Filatova, E. (2012). Irony and sarcasm: Corpus generation and analysis using crowdsourcing. In *Lrec*, pages 392–398. Citeseer.

Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, 27(2):83–85.

Frye, N. (1944). The nature of satire. *University of Toronto Quarterly*, 14(1):75– 89.

Garc´ıa-D´ıaz, J. A., Jim´enez-Zafra, S. M., Garc´ıa-Cumbreras, M. A., and Valencia-Garc´ıa, R. (2023). Evaluating feature combination strategies for hate-speech detection in spanish using linguistic features and transformers. *Complex & Intelligent Systems*, 9(3):2893–2914.

Garson, G. D. (2008). Discriminant function analysis. statnotes: topics in mul- tivariate analysis. *Retrieved March*, 29:2010.

Gehler, P. and Nowozin, S. (2009). On feature combination for multiclass object classification. In *2009 IEEE 12th International Conference on Computer Vision*, pages 221–228. IEEE.

Ghosh, A., Li, G., Veale, T., Rosso, P., Shutova, E., Barnden, J., and Reyes, A. (2015). Semeval-2015 task 11: Sentiment analysis of figurative language in twitter. In *Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015)*, pages 470–478.

Ghosh, A. and Veale, T. (2016). Fracking sarcasm using neural network. In *Proceedings of the 7th workshop on computational approaches to subjectivity, sentiment and social media analysis*, pages 161–169.

Ghosh, A. and Veale, T. (2017). Magnets for sarcasm: Making sarcasm detection timely, contextual and very personal. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 482–491.

Golbeck, J., Mauriello, M., Auxier, B., Bhanushali, K. H., Bonk, C., Bouza- ghrane, M. A., Buntain, C., Chanduka, R., Cheakalos, P., Everett, J. B., et al. (2018). Fake news vs satire: A dataset and analysis. In *Proceedings of the 10th ACM Conference on Web Science*, pages 17–21.

Gonz´alez-Ib´anez, R., Muresan, S., and Wacholder, N. (2011). Identifying sar- casm in twitter: a closer look. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Tech- nologies*, pages 581–586.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press. Grosan, C. and Abraham, A. (2011). *Intelligent systems*. Springer.

Guibon, G., Ermakova, L., Seffih, H., Firsov, A., and Le No´e-Bienvenu, G. (2019). Multilingual fake news detection with satire. In *CICLing: In- ternational Conference on Computational Linguistics and Intelligent Text Processing*.

Hays, P. L. (1966). Tennessee williams' use of myth in" sweet bird of youth".*Educational Theatre Journal*, pages 255–258.

Hazarika, D., Poria, S., Gorantla, S., Cambria, E., Zimmermann, R., and Mi- halcea, R. (2018). Cascade: Contextual sarcasm detection in online discussion forums. *arXiv preprint arXiv:1805.06413*.

Hepburn, A. D. (1875). *Manual of English rhetoric*. Wilson, Hinkle & Com- pany.

Hochreiter, S. and Schmidhuber, J. (1997a). Long short-term memory. *Neural computation*, 9(8):1735–1780.

Hochreiter, S. and Schmidhuber, J. (1997b). Lstm can solve hard long time lag problems. *Advances in neural information processing systems*, pages 473–479.

Hou, J., Gao, H., Xia, Q., and Qi, N. (2015). Feature combination and the knn framework in object classification. *IEEE transactions on neural networks and learning systems*, 27(6):1368–1378.

Housley, S. (1896). The evolution of parliament. *Strand magazine: an illus- trated monthly*, 11:104–112.

Houston, K. (2013). *Shady characters: The secret life of punctuation, symbols, and other typographical marks*. WW Norton & Company.

Huang, J. and Ling, C. X. (2005). Using auc and accuracy in evaluating learning algorithms. *IEEE Transactions on knowledge and Data Engineering*, 17(3):299–310.

Ili´c, S., Marrese-Taylor, E., Balazs, J. A., and Matsuo, Y. (2018). Deep contex- tualized word representations for detecting sarcasm and irony. *arXiv preprint arXiv:1809.09795*.

Ismail, S. S., Mansour, R. F., El-Aziz, A., Rasha, M., Taloba, A. I., et al. (2022). Efficient e-mail spam detection strategy using genetic decision tree processing with nlp features. *Computational Intelligence and Neuroscience*, 2022.

Itti, L. and Koch, C. (1999). Comparison of feature combination strategies for saliency-based visual attention systems. In *Human vision and electronic imaging IV*, volume 3644, pages 473–482. SPIE.

Ivanko, S. L. and Pexman, P. M. (2003). Context incongruity and irony pro- cessing. *Discourse Processes*, 35(3):241–279.

Jang, H., Jo, Y., Shen, Q., Miller, M., Moon, S., and Rose, C. (2016). Metaphor detection with topic transition, emotion and cognition in context. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 216–225.

Jang, H., Moon, S., Jo, Y., and Rose, C. (2015). Metaphor detection in dis- course. In *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 384–392.

Jaynes, J. (2000). *The origin of consciousness in the breakdown of the bicameral mind*. Houghton Mifflin Harcourt.

Joshi, A., Agrawal, S., Bhattacharyya, P., and Carman, M. J. (2017a). Expect the unexpected: Harnessing sentence completion for sarcasm detection. In *International Conference of the Pacific Association for Computational Linguistics*, pages 275–287. Springer.

Joshi, A., Bhattacharyya, P., and Carman, M. J. (2017b). Automatic sarcasm detection: A survey. *ACM Computing Surveys (CSUR)*, 50(5):1–22.

Joshi, A., Sharma, V., and Bhattacharyya, P. (2015). Harnessing context incongruity for sarcasm detection. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 757–762.

Joy, S. (2009). Lost in translation: Emotion and expression through technology.

Kalchbrenner, N., Grefenstette, E., and Blunsom, P. (2014). A convolutional neural network for modelling sentences. *arXiv preprint arXiv:1404.2188*.

Kamin´ski, B., Jakubczyk, M., and Szufel, P. (2018). A framework for sensitivity analysis of decision trees. *Central European journal of operations research*, 26(1):135–159.

Kim, Y. (2014). Convolutional neural networks for sentence classification.

Kirkpatrick, P. K. (2016). 2016: The 22nd conference of the australasian humour studies network (ahsn). In *26th Australasian Humour Studies Network Conference*.

Klebanov, B. B., Leong, B., Heilman, M., and Flor, M. (2014). Different texts, same metaphors: Unigrams and beyond. In *Proceedings of the Second Workshop on Metaphor in NLP*, pages 11–17.

Klingner, B. (2015). Respond cautiously to north korean engagement offers. *Heritage Foundation*.

Knight, C. A. (2004). *The literature of satire*. Cambridge University Press.

Kozareva, Z. (2013). Multilingual affect polarity and valence prediction in metaphor-rich texts. In *Proceedings of the 51st Annual Meeting of the Associ- ation for Computational Linguistics (Volume 1: Long Papers)*, pages 682–691.

Kreuz, R. and Caucci, G. (2007). Lexical influences on the perception of sar- casm. In *Proceedings of the Workshop on computational approaches to Figura- tive Language*, pages 1–4.

Kruger, J., Epley, N., Parker, J., and Ng, Z.-W. (2005). Egocentrism over e-mail: Can we communicate as well as we think? *Journal of personality and social psychology*, 89(6):925.

Kumon-Nakamura, S., Glucksberg, S., and Brown, M. (2007). How about another piece of pie: The allusional pretense theory of discourse irony. *Irony in language and thought*, pages 57–96.

Kunneman, F., Liebrecht, C., Van Mulken, M., and Van den Bosch, A. (2015). Signaling sarcasm: From hyperbole to hashtag. *Information Processing & Management*, 51(4):500–509.

LeCun, Y., Bengio, Y., et al. (1995). Convolutional networks forimages, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995.

Leslau, W. (1995). *Reference grammar of Amharic*. Otto Harrassowitz Verlag.

Liberman, A. (2009). *Word origins... and how we know them: Etymology for everyone*. OUP USA.

Liebeschuetz, W. (1965). Beast and man in the third book of virgil's georgics.*Greece & Rome*, 12(1):64–77.

Liebrecht, C., Kunneman, F., and van Den Bosch, A. (2013). The perfect solution for detecting sarcasm in tweets# not.

Lin, Z., Feng, M., Santos, C. N. d., Yu, M., Xiang, B., Zhou, B., and Bengio Y. (2017). A structured self-attentive sentence embedding.*arXiv preprint arXiv:1703.03130*.

Lincoln, A. (2008). *The collected works of Abraham Lincoln*, volume 2. Wildside Press LLC.

Littlefield, H. M. (1964). The wizard of oz: Parable on populism. *American Quarterly*, 16(1):47–58.

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167.

Liu, B., Hu, M., and Cheng, J. (2005). Opinion observer: analyzing and com- paring opinions on the web. In *Proceedings of the 14th international conference on World Wide Web*, pages 342–351.

Liu, B. and Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In *Mining text data*, pages 415–463. Springer.

Lukin, S. and Walker, M. (2017). Really? well. apparently bootstrapping improves the performance of sarcasm and nastiness classifiers for online dialogue. *arXiv preprint arXiv:1708.08572*.

Mack, M. (1952). The world of hamlet. *Yale Review*, 41(502):23.

Majumder, N., Poria, S., Peng, H., Chhaya, N., Cambria, E., and Gelbukh, A. (2019). Sentiment and sarcasm classification with multitask learning. *IEEE Intelligent Systems*, 34(3):38–43.

Mao, R., Lin, C., and Guerin, F. (2018). Word embedding and wordnet based metaphor identification and interpretation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Pa- pers)*, pages 1222–1231.

Mason, Z. J. (2004). Cormet: a computational, corpus-based conventional metaphor extraction system. *Computational linguistics*, 30(1):23–44.

Matthiessen, P. (2008). *The snow leopard*. Penguin.

Maynard, D. G. and Greenwood, M. A. (2014). Who cares about sarcastic tweets? investigating the impact of sarcasm on sentiment analysis. In *Lrec 2014 proceedings*. ELRA.

McArthur, T., Lam-McArthur, J., and Fontaine, L. (2018). *Oxford companion to the English language*. Oxford University Press.

McCarthy, M. and Carter, R. (2004). "there's millions of them": hyperbole in everyday conversation. *Journal of pragmatics*, 36(2):149–184.

McHardy, R., Adel, H., and Klinger, R. (2019). Adversarial training for satire detection: Controlling for confounding variables. *arXiv preprint arXiv:1902.11145*.

Meier, B. P. and Robinson, M. D. (2005). The metaphorical representation of affect. *Metaphor and symbol*, 20(4):239–257.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Misra, R. and Arora, P. (2019). Sarcasm detection using hybrid neural network.*arXiv preprint arXiv:1908.07414*.

Mohammad, S., Shutova, E., and Turney, P. (2016). Metaphor as a medium for emotion: An empirical study. In *Proceedings of the Fifth Joint Conference on Lexical and Computational Semantics*, pages 23–33.

Mohanty, M. N. and Palo, H. K. (2019). Segment based emotion recognition using combined reduced features. *International Journal of Speech Technology*, 22(4):865–884.

Mohanty, M. N. and Palo, H. K. (2020). Child emotion recognition using probabilistic neural network with effective features. *Measurement*, 152:107369.

Mohebali, B., Tahmassebi, A., Meyer-Baese, A., and Gandomi, A. H. (2020). Probabilistic neural networks: a brief overview of theory, implementation, and application. *Handbook of probabilistic models*, pages 347–367.

Mohler, M., Bracewell, D., Tomlinson, M., and Hinote, D. (2013). Semantic signatures for example-based linguistic metaphor detection. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 27–35.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European sociological review*, 26(1):67–82.

Musgrave, D. (2014). *Grotesque Anatomies: Menippean Satire Since the Re- naissance*. Cambridge Scholars Publishing.

Orwell, G. (1967). Politics vs. literature: an examination of gulliver's travels. In *Fair Liberty was all his Cry*, pages 166–185. Springer.

Orwell, G. (1984). *The Orwell reader: Fiction, essays, and reportage*. Houghton Mifflin Harcourt.

Orwell, G. and Heath, A. (2003). *Animal farm and 1984*. Houghton Mifflin Harcourt.

Palo, H. K. and Mohanty, M. N. (2018). Wavelet based feature combination for recognition of emotions. *Ain shams engineering journal*, 9(4):1799–1806.

Panda, A. (2020). *Kim Jong Un and the bomb: Survival and deterrence in North Korea*. Oxford University Press.

Peled, L. and Reichart, R. (2017). Sarcasm sign: Interpreting sarcasm with sen- timent based monolingual machine translation. *arXiv preprint arXiv:1704.06836*.

Pennebaker, J. W., Francis, M. E., and Booth, R. J. (2001). Linguistic in- quiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001.

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., and Zettlemoyer, L. (2018). Deep contextualized word representations. *arXiv preprint arXiv:1802.05365*.

Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, 4(2):1883.

Poria, S., Cambria, E., Hazarika, D., and Vij, P. (2016). A deeper look into sarcastic tweets using deep convolutional neural networks. *arXiv preprint arXiv:1610.08815*.

Potamias, R.-A., Siolas, G., and Stafylopatis, A. (2019). A robust deep en- semble classifier for figurative language detection. In *Engineering Applications of Neural Networks: 20th International Conference, EANN 2019, Xersonisos, Crete, Greece, May 24-26, 2019, Proceedings 20*, pages 164–175. Springer.

Pt´aˇcek, T., Habernal, I., and Hong, J. (2014). Sarcasm detection on czech and english twitter. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 213–223.

Qadir, A., Riloff, E., and Walker, M. A. (2016). *Automatically inferring implicit properties in similes*.

Rajadesingan, A., Zafarani, R., and Liu, H. (2015). Sarcasm detection on twitter: A behavioral modeling approach. In *Proceedings of the eighth ACM international conference on web search and data mining*, pages 97–106.

Ramakrishnan, N. (2013). *Reading Gandhi in the Twenty-First Century*. Springer.

Ramteke, A., Malu, A., Bhattacharyya, P., and Nath, J. S. (2013). Detect- ing turnarounds in sentiment analysis: Thwarting. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 860–865.

Reagan, R. (1986). *Public Papers of the Presidents of the United States: Ronald Reagan, 1984*. Best Books on. really-learn english.com (2021). Pronoun. https://www.really-learn-english.com/list- of-pronouns.html. Accessed: 2021-02-21.

Recupero, D. R., Alam, M., Buscaldi, D., Grezka, A., and Tavazoee, F. (2019). Frame-based detection of figurative language in tweets [application notes]. *IEEE Computational Intelligence Magazine*, 14(4):77–88.

Reganti, A. N., Maheshwari, T., Kumar, U., Das, A., and Bajpai, R. (2016). Modeling satire in english text for automatic detection. In *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, pages 970– 977. IEEE.

Rei, M., Bulat, L., Kiela, D., and Shutova, E. (2017). Grasping the finer point: A supervised similarity network for metaphor detection. *arXiv preprint arXiv:1709.00575*. relatedwords.org (2021). Distance words. https://relatedwords.org/relatedto/distance. Accessed: 2021-05-25.

Reyes, A., Rosso, P., and Buscaldi, D. (2012). From humor recognition to irony detection: The figurative language of social media. *Data & Knowledge Engineering*, 74:1–12.

Richards, I. A. and Constable, J. (2018). *The philosophy of rhetoric*. Routledge.

Riloff, E., Qadir, A., Surve, P., De Silva, L., Gilbert, N., and Huang, R. (2013). Sarcasm as contrast between a positive sentiment and negative situation. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, pages 704–714.

Rivera, A. T., Oliver, A., Climent, S., and Coll-Florit, M. (2020). Neural metaphor detection with a residual bilstm-crf model. In *Proceedings of the Second Workshop on Figurative Language Processing*, pages 197–203.

Roberts, R. M. and Kreuz, R. J. (1994). Why do people use figurative language? *Psychological science*, 5(3):159–163.

Rockwell, P. (2003). Empathy and the expression and recognition of sarcasm by close relations or strangers. *Perceptual and motor skills*, 97(1):251–256.

Rockwell, P. A. (2006). *Sarcasm and other mixed messages: The ambiguous ways people use language*. Edwin Mellen Press.

Rubin, V. L., Conroy, N., Chen, Y., and Cornwell, S. (2016). Fake news or truth? using satirical cues to detect potentially misleading news. In *Proceed- ings of the second workshop on computational approaches to deception detection*, pages

7–17.

Russel, S. and Norvig, P. (1995). Artificial intelligence: A modern approach, 2003. *EUA: Prentice Hall*, 178.

Sam, G. and Catrinel, H. (2006). On the relation between metaphor and simile: When comparison fails. *Mind & Language*, 21(3):360–378.

Sanders, C. (1971). *The scope of satire*. Scott, Foresman.

Santurkar, S., Tsipras, D., Ilyas, A., and Madry, A. (2018). How does batch normalization help optimization? In *Proceedings of the 32nd international conference on neural information processing systems*, pages 2488–2498.

Schifanella, R., de Juan, P., Tetreault, J., and Cao, L. (2016). Detecting sarcasm in multimodal social platforms. In *Proceedings of the 24th ACM international conference on Multimedia*, pages 1136–1145.

Schuler, K. K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon*. University of Pennsylvania.

Shakespeare, W. (2007). *The complete works of William Shakespeare*. Wordsworth Editions.

Shamay-Tsoory, S. G., Tomer, R., and Aharon-Peretz, J. (2005). The neu- roanatomical basis of understanding sarcasm and its relationship to social cog- nition. *Neuropsychology*, 19(3):288.

Shero, L. R. (1923). The cena in roman satire. *Classical Philology*, 18(2):126– 143.

Shmueli, B., Ku, L.-W., and Ray, S. (2020). Reactive supervision: A new method for collecting sarcasm data. *arXiv preprint arXiv:2009.13080*.

Shutova, E., Sun, L., and Korhonen, A. (2010). Metaphor identification using verb and noun clustering. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, pages 1002–1010.

sightwordsgame.com (2021a). Directional words. https://www.sightwordsgame.com/directional- positional-words/.

sightwordsgame.com (2021b). Positional words. https://www.sightwordsgame.com/directional- positional-words/. Accessed: 2021-05-25.

sightwordsgame.com (2021c). Size words. https://www.sightwordsgame.com/parts- of-speech/adjectives/size/. Accessed: 2021-05-25.

Simon, T. (2007). *Jupiter's travels*. Penguin UK.

Simpson, R. (1878). *The School of Shakespeare...*, volume 2. JW Bouton.

Sin-Wai, C. (2015). The routledge encyclopedia of translation technology.*Abingdon:*

Routledge.

Su, C., Fukumoto, F., Huang, X., Li, J., Wang, R., and Chen, Z. (2020). Deep- met: A reading comprehension paradigm for token-level metaphor detection. In *Proceedings of the Second Workshop on Figurative Language Processing*, pages 30–39.

Swarnkar, K. and Singh, A. K. (2018). Di-lstm contrast: A deep neural network for metaphor detection. In *Proceedings of the Workshop on Figurative Language Processing*, pages 115–120.

Swift, J., Davis, H., and Williams, H. (1959). *Gulliver's travels: 1726*. Black- well. talkenglish.com (2021). Noun. https://www.talkenglish.com/vocabulary/top-1500-nouns.aspx. Accessed: 2021-02-21.

Tan, J., Yang, J., Wu, S., Chen, G., and Zhao, J. (2021). A critical look at the current train/test split in machine learning. *arXiv preprint arXiv:2106.04525*.

Thorogood, J. (2016). Satire and geopolitics: Vulgarity, ambiguity and the body grotesque in south park. *Geopolitics*, 21(1):215–235.

Thu, P. P. and New, N. (2017). Implementation of emotional features on satire detection. In *2017 18th IEEE/ACIS International Conference on Soft- ware Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, pages 149–154. IEEE.

Troiano, E., Strapparava, C., O¨ zbal, G., and Tekiro˘glu, S. S. (2018). A compu- tational exploration of exaggeration. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3296–3304.

Tsur, O., Davidov, D., and Rappoport, A. (2010). Icwsm—a great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews. In *fourth international AAAI conference on weblogs and social media*.

Tsvetkov, Y., Boytsov, L., Gershman, A., Nyberg, E., and Dyer, C. (2014). Metaphor detection with cross-lingual model transfer. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 248–258.

Turney, P., Neuman, Y., Assaf, D., and Cohen, Y. (2011). Literal and metaphor- ical sense identification through concrete and abstract context. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 680–690.

Ullman, B. L. (1913). Satura and satire. *Classical Philology*, 8(2):172–194.

Vapnik, V. and Izmailov, R. (2017). Knowledge transfer in svm and neural networks. *Annals of Mathematics and Artificial Intelligence*, 81(1):3–19.

Veale, T. (2012). *Exploding the creativity myth: The computational foundations of linguistic creativity*. A&C Black.

Vu, N. T., Adel, H., Gupta, P., and Schu¨tze, H. (2016). Combining recurrent and convolutional neural networks for relation classification. *arXiv preprint arXiv:1605.07333*.

Wang, F., Liu, H., and Cheng, J. (2018). Visualizing deep neural network by alternately image blurring and deblurring. *Neural Networks*, 97:162–172.

Wang, H., Li, Y., Khan, S. A., and Luo, Y. (2020). Prediction of breast can- cer distant recurrence using natural language processing and knowledge-guided convolutional neural network. *Artificial intelligence in medicine*, 110:101977.

Weinbrot, H. D. (2005). *Menippean satire reconsidered: From antiquity to the eighteenth century*. JHU Press.

Wilcock, S. (2013). *A comparative analysis of fansubbing and professional DVD subtitling*. PhD thesis, University of Johannesburg.

Wilks, Y., Dalton, A., Allen, J., and Galescu, L. (2013). Automatic metaphor detection using large-scale lexical resources and conventional metaphor extrac- tion. In *Proceedings of the First Workshop on Metaphor in NLP*, pages 36–44.

Yang, F., Mukherjee, A., and Dragut, E. (2017). Satirical news detection and analysis using attention mechanism and linguistic features. *arXiv preprint arXiv:1709.01189*.

Yih, W.-t., Richardson, M., Meek, C., Chang, M.-W., and Suh, J. (2016). The value of semantic parse labeling for knowledge base question answering. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 201–206.

Yin, C., Chen, Y., and Zuo, W. (2021). Multi-task deep neural networks for joint sarcasm detection and sentiment analysis. *Pattern Recognition and Image Analysis*, 31(1):103–108.

Yin, W., Schu¨tze, H., Xiang, B., and Zhou, B. (2016). Abcnn: Attention-based convolutional neural network for modeling sentence pairs. *Transactions of the Association for Computational Linguistics*, 4:259–272.

yourdictionary.com (2021). Temporal words. https://grammar.yourdictionary.com/style- and-usage/list-transition-words.html. Accessed: 2021-02-21.

Zayed, O., McCrae, J. P., and Buitelaar, P. (2018). Phrase-level metaphor identification using distributed representations of word meaning. In *Proceedings of the Workshop on Figurative Language Processing*, pages 81–90.

Zhang, Y. and Wallace, B. (2016). A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification.

Zhang, Y., Yang, F., Zhang, Y., Dragut, E., and Mukherjee, A. (2020). Birds of a feather flock together: Satirical news detection via language model differ- entiation. *arXiv preprint arXiv:2007.02164*.

Zhang, Z. and Luo, L. (2019). Hate speech detection: A solved problem? the challenging case of long tail on twitter. *Semantic Web*, 10(5):925–945.

Zolnay, A., Schluter, R., and Ney, H. (2005). Acoustic feature combination for robust speech recognition. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 1, pages I–457. IEEE.