
Faculty of Language and Communication

# ANALYSING HYPERBOLIC FEATURES OF SARCASTIC TWEETS OF DIFFERENT SENTIMENT POLARITIES

**NOR EISYA SHABILA BINTI ISMAIL**
**(61809)**

**Bachelor of Arts with Honours (Linguistics)**
**Universiti Malaysia Sarawak**

**2020**

# ANALYSING HYPERBOLIC FEATURES OF SARCASTIC TWEETS OF DIFFERENT SENTIMENT POLARITIES

by

**NOR EISYA SHABILA BINTI ISMAIL**
**(61809)**

This final year project is submitted in partial fulfilment of the requirements for the
Degree of Bachelor of Arts with Honours (Linguistics),
Faculty of Language and Communication,
Universiti Malaysia Sarawak

2020

**UNIVERSITI MALAYSIA SARAWAK**

Grade: _____

**Please tick ( )**

Final Year Project ☐

Report Masters ☐

PhD ☐

**DECLARATION OF ORIGINAL WORK**

This declaration is made on the …….. day of ..... (year).

**Student's Declaration**:

I ---------- ----
(PLEASE INDICATE STUDENT'S NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled, ----------is my original
work. I have not copied from any other students' work or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

_____ _____

Date submitted Name of the student (Matric No.)

**Supervisor's Declaration:**

I---------- (SUPERVISOR'S NAME) hereby certifies that the work entitled, ---------- ----------
(TITLE) was prepared by the above named student, and was submitted to the "FACULTY" as a * partial/full fulfilment for the conferment of ----------
--
(PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work

Received for examination by:

(Name of the supervisor)

Date:

I declare this Project/Thesis is classified as (Please tick (√)):

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*****
☐ **RESTRICTED** (Contains restricted information as specified by the organisation where research was done)*****
☐ **OPEN ACCESS**

**Validation of Project/Thesis**

I therefore duly affirmed with free consent and willingness declared that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abide interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies for the purpose of academic and research only and not for other purpose.
- The Centre for Academic Information Services has the lawful right to digitise the content to for the Local Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic exchange between Higher Learning Institute.
- No dispute or any claim shall arise from the student itself neither third party on this Project/Thesis once it becomes sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student except with UNIMAS permission.

Student's signature_____ Supervisor's signature:

_____

(Date) (Date)

Current Address:

_____

_____

Notes: ***** If the Project/Thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach together as annexure a letter from the organisation with the period and reasons of confidentiality and restriction.

[The instrument was duly prepared by The Centre for Academic Information Services]

## ACKNOWLEDGEMENT

First and foremost, praises and thanks to the God, the Almighty, for His showers of blessings throughout my research work to complete the research successfully.

I would like to express my deep and sincere gratitude to my research supervisor, Mr Chuah Kee Man, Deputy Dean (Teaching Excellent), Centre for Applied Learning and Multimedia, for giving me the opportunity to do research and providing invaluable guidance throughout this research. His dynamism, vision, sincerity and motivation have deeply inspired me. He has taught me the methodology to carry out the research and to present the research works as clearly as possible. It was a great privilege and honor to work and study under his guidance. I am extremely grateful for what he has offered me. He is indeed my mentor and I am always thrilled with your spirits.

Not to mention my course mates, my friends and my fellow juniors of Language and Communication Faculty for giving me hands and motivational support throughout my journey. We had a wonderful time together, from knowing and memorizing each other names, I never thought we could become a team that struggling together for our future. Furthermore, I am extremely grateful to my parents for their love, prayers, caring and sacrifices for educating and preparing me for my future. My family had provided enough financial and moral support for me to complete this research. I really would like to pay them back with my successful live. For that, I owe them big time. One day, I hope I can make them proud by sending me to UNIMAS.

## ABSTRACT

## ANALYSING HYPERBOLIC FEATURES OF SARCASTIC TWEETS OF DIFFERENT SENTIMENT POLARITIES

~~NOR EISYA SHABILA BINTI ISMAIL~~

The increasing use of Twitter as a medium of communication has led to more linguistic analysis of its content. One area which is crucial but lacking in research is sarcasm usage in t~~T~~weets. This study aims to analyse hyperbolic features of sarcastic t~~T~~weets from different sentiment polarities. This research applied~~s~~ qualitative methods to analyse and investigate the sarcasm usage found among the users in the context of an asynchronous social network platform (Twitter). This research implements few methods for strengthening the methodology which includes data corpus, and content analysis using two main instruments, known as SummarizeBot tool, and Antconc. software. The tweets gathered are mainly according to the sarcastic remarks or hashtags, which include #sarcasm #sarcastic, #not~~really~~, and #justsaying. In this study, 500 tweets were selected for analysis with approximately 11, 000 words counts. The findings show that the highly used sentiment polarity in sarcastic tweets are positive sentiment. In addition, this study ~~found out~~reveals that hyperbolic features ~~also~~were present in ~~detected~~ more than 50% of the corpus. Based on the findings, sarcasm can be identified as one of the human's creativity in language use. However, these findings ~~are remained uncertain~~may not be conclusive ~~if it will be the same for the future~~ because the use of sarcasm can change from time to time based on their purpose of creating the true meaning of the sentences through sarcasm. Thus, further studies are needed to identify other linguistic features of sarcastic t~~T~~weets to ease communication.



**Keywords** – Twitter; hyperbolic features; sentiment polarity; sarcasm


**Formatted:** Font: Bold

**ABSTRAK**

***MENGANALISIS CIRI-CIRI HYPERBOLIK DALAM "TWEETS" SARKASTIK BERDASARKAN DUA POLARITI SENTIMEN YANG BERBEZA***

~~NOR EISYA SHABILA BINTI ISMAIL~~

~~<~~ *Peningkatan penggunaan Twitter sebagai media komunikasi menyebabkan analisis kandungan linguistik telah berkembang dan menjadi lebih banyak. Salah satu bidang yang sangat penting namun masih kurang dalam penyelidikan adalah kajian mengenai penggunaan sarkasme dalam tweets".~~~~translate the added lines in Engl~~ ~~ish>~~

***Keywords*** *– Twitter; ciri-ciri hiperbolik; polarity sentimen; sarkasme*

**TABLE OF CONTENT**



Formatted Table

## LIST OF TABLES

## LIST OF FIGURES



Formatted: Centered

# CHAPTER 1

# INTRODUCTION

## 1.1 Background

Social media have gained massive prevalence among the citizens of all over the world. With the appearance of smart mobile devices that accompany the fast web, the users are currently able to surf and connect with social media and experiencing the "online life" administrations. The increasing number of social media also affecting the numbers of users on social media sites such as Twitter, Facebook, Instagram, and so on. Chaffey (2016) stated that the statistics shows a number of 17% yearly increase in mobile users with the total number of smartphone users almost 3.7 billion people, according to the Global WebIndex. Social media had been one of the initiatives for users to express their feeling and opinion on a daily basis. Moreover, social media also gained popularity because it became a platform where users can interact and connect with other people worldwide. Moreover, Facebook had asserted that they have 1.59 billion dynamic users monthly while Twitter professes to have in excess of 500 million users, which around 332 million of dynamic users (Chaffey, 2016).

Sarcasm can be interpreted in different ways, however recognizing sarcasm is very important for the Natural Language Processing (NLP) to abstain from confusing sarcastic statements as literal statement. For instance, sentiment analysis can be easily misled by the nearness of words that have solid polarity yet are used sarcastically, which implies that the contrary of polarity was intended. Sarcasm is a mode or way of communication where literal and intended meanings are in opposition. According to Huang, Gino, and Galinsky (2015),

sarcasm includes the exposure to contradictions among expressed and intended meanings. Therefore, sarcasm is regularly used to pass on negative messages by utilizing positive words. Gibbs (2000) stated that the most typical form of verbal irony often used to amusingly pass on not at all subtle dissatisfaction, hatred, and contempt, as on account of sarcastic reactions. Therefore, sarcasm was widely used in social media or websites, for example on Twitter. Hence, the art approaches of sentiment analysis and opinion tend to have lower performances when analysing data collected by these websites. According to Maynard and Greenwood (2014), when sarcasm within sarcastic statements are identified, the sentiment analysis might be highly enhanced.

In this study, the researcher choose the hyperbolic features of interjections, punctuation marks and quotes according to several requirement, the interjection must have an abruption remarks, such as "Yay, Well, Oh, Wow, etc.". The punctuation marks features were selected when the tweets contained repeated use of marks or symbol such as the repeated full stop marks, punctuation marks, and question marks. Lastly, the quotes features were selected if the tweets have the quotation marks (e.g: "hate"). Bharti et al. (2016) stated that a hyperbolic text contains one of the text properties such as intensifier (e.g., Adjective or Adverb), interjection (e.g., wow, aha), quotes (e.g.," " or ' '), punctuation mark (e.g., ????, !!!!). In addition, Kreuz et al. (2007) also stated that the other hyperbolic terms such as interjection and punctuation and

# CHAPTER 1

# INTRODUCTION

## 1.1 Background

Social media have gained massive prevalence among the citizens of all over the world. With the appearance of smart mobile devices that accompany the fast web, the users are currently able to surf and connect with social media and experiencing the "online life" administrations. The increasing number of social media also affecting the numbers of users on social media sites such as Twitter, Facebook, Instagram, and so on. Chaffey (2016) stated that the statistics shows a number of 17% yearly increase in mobile users with the total number of smartphone users almost 3.7 billion people, according to the Global WebIndex. Social media had been one of the initiatives for users to express their feeling and opinion on a daily basis. Moreover, social media also gained popularity because it became a platform where users can interact and connect with other people worldwide. Moreover, Facebook had asserted that they have 1.59 billion dynamic users monthly while Twitter professes to have in excess of 500 million users, which around 332 million of dynamic users (Chaffey, 2016).

Sarcasm can be interpreted in different ways, however recognizing sarcasm is very important for the Natural Language Processing (NLP) to abstain from confusing sarcastic statements as literal statement. For instance, sentiment analysis can be easily misled by the nearness of words that have solid polarity yet are used sarcastically, which implies that the

contrary of polarity was intended. Sarcasm is a mode or way of communication where literal and intended meanings are in opposition. According to Huang, Gino, and Galinsky (2015), sarcasm includes the exposure to contradictions among expressed and intended meanings. Therefore, sarcasm is regularly used to pass on negative messages by utilizing positive words. Gibbs (2000) stated that the most typical form of verbal irony often used to amusingly pass on not at all subtle dissatisfaction, hatred, and contempt, as on account of sarcastic reactions. Therefore, sarcasm was widely used in social media or websites, for example on Twitter. Hence, the art approaches of sentiment analysis and opinion tend to have lower performances when analysing data collected by these websites. According to Maynard and Greenwood (2014), when sarcasm within sarcastic statements are identified, the sentiment analysis might be highly enhanced.

In addition, according to Bharti et al. (2017), sarcastic tweets usually contain hyperbolic features either intensifier or interjection. The usage of hyperbolic features in language is very common and can be frequently found, either in texts or conversations. However, these features have only occasionally been idea deserving of genuine diagnostic consideration. Overstatement is typically treated as a minor figure of speech which has a place with either of the two prevailing allegorical employments of language, metaphor and irony. Hyperbole have very important role in identifying sarcasm, even more in textual data. The combination of some text properties such as intensifier, interjections, quotes, punctuation marks and more in texts are considered as hyperbole. Kunneman et al. (2015) also stipulated that if texts contain interjection words hence it is most likely and have higher tendency to be considered as sarcastic. Therefore, we can conclude that hyperbole is one of the common patterns that can be seen in sarcastic texts. Even though this research already covered about the hyperbolic features, however it is never be done according to its sentiment polarities.

## 1.2 Problem Statement

However, because of the casual language utilized on this web-based life locales, social media sites, and the limitation in terms of characters which is only 280 characters per tweet, it is difficult to understand the opinions or messages that the other users wanted to convey to others. In addition, the presence of sarcasm in sentences makes it more difficult to understand the real meaning of the sentences. According to Liebrecht et al. (2013), a sense of sarcasm can be the polarity-switcher of tweets meanwhile Maynard and Greenwood (2014) proposed a set of rules to set the polarity of the tweets, either it is positive or negative when sarcasm detected.

Sometimes, people tend to be more sarcastic from one another, however, generally sarcasm is so common and much harder to detect. The usage of sarcasm in daily life mainly because people are trying to be humorous as possible and making jokes from one another. However, sarcasm also used when people wanted to criticize something, such as any ideas, persons and more. Since browsing social medias had become one of the daily routines in life, therefore people always express their opinion though these platforms, hence sarcasm tends to be widely used in social media or websites, for example on Twitter, Facebook and more. The widely used sarcasm in daily life had cause confusion within the society because of the positive words used instead of negative words. Sarcasm is known as a positive way to convey negative meaning- This problem had caused more misconception that leads to miscommunication among the speakers itself. Hence, the knowledge about types of frequent used positive words in conveying sarcastic messages need to be delivered in order to create awareness among the society and community itself.

Polarity classifation can be considered as one of the most mainstream in sentiment analysis tasks. According to Pang et al. (2002), polarity classification is the field which deciding the texts is considered as positive, negative, or neutral. Although deciding the polarities of texts may seem like a simple yet easy work, however it did not easy as it seems because in deciding

the polarities, it requires the reader to understand the context of the words itself. This is due to the words did not always express the same sentiment, where it always changes according to its context. For example, words that clearly convey positive polarity such as "happy" can be neutral or negative according to its context. Hence, sentiment analysis should also include semantic-level analysis, in order to solve any words ambiguity and deciding its polarities correctly. Also, complex linguistic processing is needed to deal with problems such as the effect of negations and informal language. Moreover, understanding the sentimental meaning of the different textual units is important to accurately determine the overall polarity of a text.

### 1.3 Aim and objectives

The aim of this study is to analyse hyperbolic features of sarcastic tweets from different sentiment polarities. This research applies qualitative methods to analyse and investigate the sarcasm usage found among the users in the context of an asynchronous social network platform (Twitter). This research focuses more on the linguistic features that are classified under the words' polarity and sentiment analysis. The objectives of this research include:

i. to identify the sentiment polarity of sentences used in sarcastic tweets.
ii. to identify the frequency of words and hyperbolic features of sarcastic tweets.
iii. to compare hyperbolic features of sarcastic tweets with positive sentiment and negative sentiment.

### 1.4 Operational definition of terms

In this study, there are several specific 'terms' are used in order to avoid any misleading information that might be out of the context. This section will explain briefly some specific terms used throughout this research:

### 1.4.1 Sentiment polarities

Polarity is a relationship between two opposite characteristics or tendencies. Polarity also literally refer to a positive or negative of something measurable. Less literally, it indicates something with two opposing but related qualities. Positive and negative language is found sprinkled all through even the most unremarkable of expressions. Over a more drawn out bit of content, positive and negative assumptions will in general normal out towards impartial and not slant measurable inductions excessively far. When investigating short substance, (for example, Tweets), in any case, it's imperative to isolate authentic assessment from fundamental, true regular expressions.

Sentiment polarity is a particular feature of any textual data. It is usually divided into three which are positive, neutral and negative. Positive polarity is when the texts or sentiments contain more positive words than other polarities, neutral polarity is when the texts did not belongs to either positive and negative, or in a condition where positive polarity and negative polarity is in equal, and negative polarity is when most of the words in the texts convey negative messages. Hence, a document that containing more than one opinionated statement would have a mixture of the polarities, so the final polarity of the sentiments will be decided by identifying the most presence polarities in the sentiment. Contextual polarity refers to the polarity of a word when considered in its context, as opposed to prior polarity when it stands alone.

In this study, the researcher will look in both sentences and words used in sarcastic tweets. The sentences polarity definitely depends on its context because in order to identify the polarity, the meaning should also be taken as important aspects in textual data. As for the words' polarity, researcher would like to identify most frequent used words polarity in tweets, either positive or negative types of words.

### 1.4.2 Sarcastic tweets

Tweets is the short massages with maximum of 280 characters per tweets from an online social networking service known as Twitter. Sarcastic is an act using sarcasm element. In other words, when an individual inserting few sarcasm elements in their texts or utterances while trying to be sarcastic toward others is also known as sarcastic. In the other hand, sarcasm is also known as the activity of saying or writing the opposite of what you mean, or the way of speaking that intended to make someone else feel stupid or show them that you are angry. The definition of sarcasm according to the Collins Cobuild English Language Dictionary (1987), sarcasm is a speech or writing which means the opposite of what it seems to be saying and which is usually intended to mock or insult someone. Sarcasm naturally involves some sort of incongruity between what is said and the situation in which sarcasm is used. Whenever an individual intending to be sarcastic, they will express, often superficially positive sentiment although they are referring to the opposite of the sentiment. Lee and Katz (1998) reported that sarcasm is generally appears in the context of interpersonal communication, either through speech or written texts.

Based on the definitions and explanations that already explained above, the researcher decided to consider the definition of sarcasm is "the activity of saying or writing the opposite of what you mean, or of speaking in a way intended to make someone else feel stupid or show them that you are angry" to be the main definition in selecting the data for the corpus. On the other hand, in this study, it refers to a sarcastic statement as one where the opposite meaning is intended and gathered through the use of hashtags, for example #sarcasm #sarcastic #justsaying and #not.

### 1.4.3 Hyperbolic features

Wales (2001) describes that hyperbole is the common trope or figure in literature and aiming to exaggerate or over-statement expression in language. However, hyperbole and telling lies to others are different. Besides, hyperbole also is an articulation which surpasses the strict articulation and furthermore contains not many semantic characteristics of communicating something. Hyperbole is frequently used in daily life, hence it might occur in our everyday conversation. This highlight, as a rule, hold the capacity of underscoring or considered as an entertaining impact in everyday discussion. Christodoulidou (2011) describes that there are two types of hyperbole which the overstatement of numbers and quantity, and the incomprehensible depiction, which the exaggeration is the point at which the speaker will, in general, exaggerate quantities of times, years, events, while the outlandish portrayal resembles dreams which are created such that they are unrealistic to happen. In this way, it must be taken metaphorically or logically.

There are a lot of features can be identified from hyperbole. A hyperbolic text contains one of the text properties such as intensifier (e.g., Adjective or Adverb), interjection (e.g., wow, aha) , quotes (e.g.," " or ' '), punctuation mark (e.g., ????, !!!!) etc (Bharti, 2016). Kreuz et al. (2007) also discussed that the other hyperbolic terms such as interjection and punctuation and showed how hyperbole is useful in sarcasm detections. In this study, these four hyperbolic features in sarcastic Tweets will be coded and analysed in order to compare the positive and negative sentiment. From those definitions we can conclude that hyperbole is a figure of speech which emphasizes a meaning with exaggeration of statement.

## 1.5 Significance of the Study

Looking at the demanding and vast evolution of delivering messages, sarcasm is being used as one of the most influential bases on giving other people's view on the intended meaning using opposite words from the actual meaning. Theoretically this study will contribute greater insight of sarcasm analysis using the proposed framework by Bharti et al. (2016) especially in terms of hyperbolic features of sarcastic tweets. More features can be discovered to the framework based on the outcome from this study.

As for the practical significances, the findings from this study will benefit agencies who are interested in advertisement and marketing to produce suitable content for specific field in order to reach out the target audiences. For example, the usage of sarcastic words in advertisement can be effective applied to promote their products. Besides, it will rebound to the benefit in considering that sarcasm plays an important role in daily life either spoken or written. The greater demand for having good communication skills justifies the need for more life-changing knowledge approaches. Social influencers, academic researcher and people who are interested in finding something interesting in studying trending language without realizing they use it in their daily conversation among themselves. People tend to use positive words when conveying sarcastic conversation, however it is quite difficult to differentiate and get the real meaning in sarcastic text or conversation intended by the speaker. By knowing certain words that frequently used in sarcastic conversation, it can create awareness among the society about sarcastic words used, and further problems such as miscommunication can be avoided. By conducting this research, the researcher would like to enhance the knowledge of society about the ways of sarcastic tweets being used and affected the community. Sometimes, people cannot understand what is the messages that the sentences want to convey. This will lead to misunderstanding occur among the users that involve in the conversations. By understanding the real meaning of the sarcasm would like to convey, this problem can be avoided.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Introduction

This chapter aims to review the related work in the following manner. Firstly, the researcher examined the research that has been conducted into identifying sarcasm in any fields and methods used. Next, the researcher looks into the study that had been conducted in detecting the sarcasm usage on media socials. After that, the researcher also examined some of the past studies that look for word polarities and hyperbolic features in their study.

## 2.2 Sarcasm

Tepperman et al. (2006) take a shot at sarcasm identification in discourse and reach a resolution that the expression "yeah right" is ordered whether it is utilized sarcastically or not as per the "prosodic, spectral, and contextual cues". Sarcasm likewise one of the components that make dialects remarkable. At the point when an individual is wry or unexpected, the person in question is communicating a supposition that for the most part in constructive words yet aiming the contrary significance of the words or setting. In the exploration announced here, sarcasm is a type of irony that by and large shows up with regards to relational correspondence, regardless of whether through discourse, composed exchange, or PC intervened talk; and is all the more intently connected with criticism of a particular individual (Lee & Katz, 1998). The utilization of sarcasm in a setting at times may reduce or upgrades the antagonism of the snide articulation.

The study conducted by Bowes and Katz (2011) 'When Sarcasm Stings' wanted to examine the relationship of family members between shyness and the socioemotional functioning has the effect on irony and sarcasm. In this study, they introduced a "pipeline" model, which can be considered as the modern-day irony and sarcasm detection. Besides, as a broad vary of fantastic aspects that cover exceptional facets of affect exploiting information from a number sentiment and emotion lexical assets for English reachable to the community, possibly referring to special psychological models of effect. Understanding sarcasm requires considering social and cultural factors, which are regularly left out in models of language. This article focuses on sarcasm, for which the definitions have frequently been free and confusing, integrating it into the thought of the irony. Although lookup has examined the perception of figurative language such as metaphor in some depth, sarcasm has been much less studied.

Wang (2013) in "#Irony or #Sarcasm- A Quantitative and Qualitative Study Based on Twitter" did a study to identify the similarities and the distinctions among irony and sarcasm based on twitter data. this study used the mixture of qualitative and quantitative for its data analysis, where using sentiment analysis to gather the quantitative data and content analysis for qualitative data. This study revealed that there are few elements of aggressiveness in sarcasm and two sense owns by irony, which is equal to aggressive. The first sense is more used to attack a specific target while the second sense is used to classify an event in order to be more ironic. This finding revealed that there are negative elements in sarcastic statements.

Next, there is also a study conducted by Skalicky and Crossley (2018) in "Linguistic Features of Sarcasm and Metaphor Production Quality" wanted to assess either the linguistic features could help in terms of quality of the figurative languages. This study used a large corpus of metaphors and sarcastic responses that had been collected from human subjects in order to seek for a deeper insight of figurative languages. This study found out that the

linguistic features were able to predict a small amount of variance, either in metaphors or in sarcastic production quality.

A study conducted by Ezzah 'Afifah Nasrudin (2019) in "*Persepsi Golongan Remaja Terhadap Penggunaaan Bahasa Sarkastik*" wanted to identify the perceptions of teenagers at all levels and the effect of using sarcastic language. The methodology used in this study is survey design using quantitative and qualitative data collected through questionnaire and interview forms. Then, the data were measure the level of language usage sarcastic questionnaire using the interpretation method suggested by Ratanawalee Wimolmas (2013). This study indicates that the use of sarcastic language has been conveyed clearly or otherwise. Besides, the effect of the use of sarcastic language could give an emotional impact on individuals. Sarcastic languages also able to help one's saying something nice (e.g., convince and praises words) even though in state of anger. Ezzah (2019) revealed that the use of sarcastic language really works to mock and insulting. This statement was supported by Gibbs (2000) who stated that sarcastic language has some elements or sentiments that can lead to negative messages. This research helps in proving the importance of knowledge about sarcastic languages.

In conclusion, sarcasm likewise one of the components that make dialects remarkable. At the point when an individual is being sarcastic, the person in question is communicating a supposition that for the most part in constructive words yet aiming the contrary significance of the words or setting. The utilization of sarcasm in a setting at times may reduce or upgrades the antagonism of the snide articulation. The use of sarcasm is concentrated to gather a definition and demonstrate a couple of characteristics of sarcasm. The current research proposes and tests a novel theoretical model in which both the construction and interpretation of sarcasm lead to greater creativity because they activate abstract thinking. The researcher

employed three different creativity measures and a well-established measure of abstract thinking, then manipulated sarcasm via a simulated conversation task and a recall task.

### 2.3 Sarcasm detection on Twitter

There are also studies that had been done in detecting sarcasm on one of social media platforms, Twitter. Ellen et al. (2013) familiarize a procedure with recognizing sarcasm in tweets that rise out of intricacy between a positive decision contrasting to a negative situation. In order to learn phrases contrasting with positive inclinations and negative conditions, this uses a bootstrapping figuring that keeps the cycle between two phases. The second step is learning the positive supposition communicates that happen close to negative situation phrases. After different accentuation shapes, the got once-over of negative conditions and positive inclination phrases are used to see the sarcasm in tweets by perceiving settings that contain a positive supposition in closeness to a negative situation express. This system relies upon the supposition that various mean tweets contain the going with structure: [+VERB PHRASE] [-SITUATION PHRASE]. In any case, the system has a couple of hindrances since it can't perceive sarcasm over different sentences. Coreference objectives are an errand in customary language getting ready to perceive various words or articulations that insinuate a comparative component, for instance, individual, spot or thing.

Bieberi, Saggion, and Ronzano (2014) aiming to assign a polarity value of exploiting, which also known as the irony-awareness features. Hence, in their study entitled “Modelling Sarcasm on Twitter: A Novel Approach”, they introduced a pipeline model that were believed in helping them to detect irony or sarcasm on social medias. This pipeline model uses a wide range of the effective features, which also covers the different facets and its effect of exploiting information from various emotion lexical and sentiment resources for English. This features also available for the community, which refers to the different psychological models of effect.

The findings of this study include that while sentiment affective representations yield best results on datasets comprising of short length texts such as tweets, richer representations derived from fine-grained emotions are more suitable for detecting sarcasm from longer length documents such as product reviews and discussion forum posts. The experimental analysis indicated that topic-enriched word embedding schemes utilized in conjunction with conventional feature sets can yield promising results for sarcasm identification. In this paper, the researcher employed and evaluated classification models using two different representations based on bag-of-words and fastest to address the problem of sentiment analysis over emojis/emoticons for Turkish positive, negative and neutral tweets. The ability to reliably identify sarcasm and irony in a text can improve the performance of many Natural Language Processing (NLP) systems including summarization, sentiment analysis, etc. This study reveals that the usage of emoticons could also have intended of being sarcastic.

Next, the study conducted by Sharma and Chakravety (2018) which entitled "Sarcasm Detection in Online Review Text" was conducted to investigate the possibility of classifying sarcasm in text and its process to distinctive textual features that are essentials for sarcasm detection. In this study, they used different categories of text features in detecting sarcasm, which lexical features, pragmatic features, linguistic incongruity and context incongruity. These four categories help to analyse the tweets from different views, for example, the lexical features used unigram, pragmatic features analyse the numbers of capital letters, emoticons, slangs and also the usage of punctuation marks. For linguistic incongruity, they analyse the tweets according to its polarity and lastly, context incongruity analyses the similarity between words. The findings of this study found out that models trained can be used in order to detect sarcasm in online text. This study reveals that pragmatic features can be considered as one of the methods for sarcasm detection.

In conclusion, the two works delineate higher than a machine is examined all through a controlled setting compare irony to a restricted set of alternative topics, whereas took from the unlabelled take a look at set a sample of product evaluations with fiftieth of the sentences classified as black. In distinction, we have the tendency to practice an educated irony detector to a real-world take a look at set representing a realistically huge sample of tweets to denote on a precise day of that the overwhelming majority is not black. Detection irony in social media is arguably and it is only real looking to take a look at a system in the context of a common distribution of sarcasm in tweets. After extraordinary accentuation shapes, the got once-over of terrible stipulations and positive inclination phrases are used to see the sarcasm in tweets by using perceiving settings that contain a high-quality supposition in closeness happening neighbouring to a poor scenario express.

**2.4 Sentiment Polarity**

Sentiment polarity is a particular feature of the text. It is usually dichotomized into two either it is positive or negative, but polarity can also be thought of as a range. A major obstacle for automatically determining the polarity of a short text constructs in which the literal meaning of the text is not the intended meaning of the sender, as many systems for the detection of polarity primarily lean on positive and negative words as markers. Sentiment in Bahasa Melayu (BM) intensified following the sentiment data is growing on social media. Therefore, the development of the lexicon of sentiment in BM is necessary. However, it is difficult to develop a lexicon of sentiment in BM because of its digital resources required is limited. Thus, various approaches and methods are used to generate a lexicon of sentiment. In “Generating a Malay Sentiment Lexicon Based on Wordnet” studied by Alexander and Omar (2017), the study was aiming to develop algorithms for generating a lexicon of sentiment in WordNet-based BM. In this study, the approach used was to match WordNet Language to WordNet Language English

to get the offset value of the initial word set. The generation of initial word sets is done through semantic relationships of synonyms and semantics found in WordNet English. The test results have shown the effectiveness of the proposed algorithm for generating the lexicon-based sentiment in English WordNet.

A study entitled "Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis" written by Maynard and Greenwood (2014) would like to identify the usage of sarcasm in tweets and its effect on sentiment analysis. In this study, they perform an analysis of the effect of sarcasm in sentiment analysis. They also analyse hashtags. In this study, the hashtags are used as the tokenization, therefore they collected the tweets by using the #notreally, #irony #sarcasm. As the first step, they simply reversed the polarity of the tweets, mean that the contrast from the actual text was identified as being sarcastic. As the results, they revealed that if there are positive hashtags followed by a sarcasm indicator, and the polarity of the tweet is either positive or neutral, hence the polarity of the sentiment needs to be flipped in order to analyse its true sentiment.

Next, Rillof et al. (2013) in "Sarcasm as a Contrast of Positive and Negative Situation" aiming to identify sarcasm that arises from contrast between a positive sentiment referring to a negative situation. In this study, the researcher identified the negative situation phrases that follows a positive sentiment (e.g., love). Next, they identified the positive sentiment phrases that occur near a negative situation phrase. Use the learned lists of sentiment and situation phrases to recognize sarcasm in tweets by identifying its context. The results find out the examples of positives sentiment (love, enjoy, awesome, can't wait) contrasting with a negative situation. Tweets that incorrectly labelled as sarcastic by the contrast method were due to overlay general negative situation phrases such as "I love working there" was labelled as sarcastic. Rillof et al. (2013) revealed that the usage of positive words referring to negative

situation mostly due to the intention of being sarcastic. This study helps in understanding that the sentiment polarities play a big role in daily communication.

González-Ibánez et al. (2011) in their study entitled "Identifying sarcasm in Twitter: A Closer Look" would like to distinguish sarcasm from positive and negative sentiments expressed in Twitter messages using lexical and pragmatic factors. They collect tweets that include hashtags that express sarcasm (#sarcasm, #sarcastic), direct positive sentiment (e.g., #happy, #joy, #lucky), and direct negative sentiment (e.g., #sadness, #angry, #frustrated). Next, they filtered all tweets where the hashtags of interest were not located at the very end of the message. Lastly, they eliminate all the hashtag and analyse only the massages in the tweets. The results revealed that negative emotion, positive emotion, auxiliary verbs, and punctuation marks and indications are the possible dependence among factors that could differentiate sarcasm from both positive and negative tweets. Through this study, Gonzalez-Ibanez et al. (2011) explored the contribution of linguistic and pragmatic features of tweets to the automatic separation of sarcastic messages from positive and negative ones. This research helps to filter the data for analysis.

The BoosTexter AdaBoost machine learning algorithm, Schapire and Singer (2000) was then implemented in two experiments. Firstly, it was used to classify instances as contextually neutral or polar, and secondly, to assert the positive or negative polarity of those instances that were not neutral. They compared the performance of the classifier when using (i) just prior polarity (ii) a small feature set of word and modification features (iii) a complex feature set that also contained sentence and structure features, for both experiments. In both tasks, the larger feature set performed best managing an accuracy of 75.9% for neutral-polar classification and 65.7% for polarity classification. This research suggests that contextual polarity is a complex phenomenon that must be tackled with sophisticated feature sets. However, we note that while a feature set must have the ability to express the complexities of

the classification target, a focus on expanding the feature set brings with it a tendency towards models that do not generalize well. Therefore, there is an important balance to be struck when comparing potential features for the supervised algorithms to be implemented.

As a conclusion, sentiment polarity is a particular feature of the text. It is usually dichotomized into two either it is positive or negative, but polarity can also be thought of as a range. Furthermore, a distinction must be made between the polarity of sentiment and of its strength. The polarity of sentiment is when deciding whether it is positive, negative or neutral, is a crucial feature for humour detection and therefore correctly ascertaining the contextual polarity of terms in subjective content forms a critical component of a number of pieces of research. Contextual polarity refers to the polarity of a word when considered in its context, as opposed to prior polarity when it stands alone. A major obstacle for automatically determining the polarity of a short text constructs in which the literal meaning of the text is not the intended meaning of the sender, as many systems for the detection of polarity primarily lean on positive and negative words as markers.

### 2.5 Hyperbolic features

As the framework of this study, a study conducted by Bharti et al. (2016) entitled "Sarcastic Sentiment Detection in Tweets Streamed in Real Time: A Big Data Approach" was aiming to identify the sarcasm detection from textual data by analysing its key features. The methods used in this study are first, they detect tweets that start with interjection words such as "aha, wow, nah, uh, etc" using few keywords with the presence of hashtags #sarcasm and #sarcastic. Next, the selected tweets will be check thoroughly if the interjection and adjective or adverb are classified as sarcastic in the tweets. Then, they categorize the sentiment analysis on sentiment phrases into positive and negative phrases. In this study, they also analyse hyperbolic features like a pattern that commonly found in sarcastic tweets. This finding found out that

sarcasm detection and analysis in social media provides invaluable insight into the current public opinion on trends and events in real-time. Through this study, Bharti et al. (2016) revealed that the usage of hyperbolic features in tweets can be considered as sarcastic. This research helps this study to analyse the hyperbolic features in tweets.

Kunneman et al. (2015) conducted a study to investigate the usefulness of a sarcastic hashtag to train sarcasm detection in other language regions in "Signalling Sarcasm: From Hyperbole to Hashtag". In this study, they collected tweets that contained few criteria such as contained the selected hashtags '#sarcasm', '#irony', '#cynicism', and '#not'. Then, they cleaned up the dataset by only including tweets in which the given hashtag was placed at the end or exclusively followed by other hashtags or a 'url'. The data was then be analysed according to its ranked of the strength of their connection to the label of sarcasm. This study found that tweets explicitly marked with the '#sarcasm' or its pseudo-synonyms are identified as sarcastic. Most tweets contain a literally positive message and contain four types of markers for sarcasm: intensified as well as un-intensified evaluative words, exclamations, and non-sarcastic hashtags. Through this study, Kunneman et al. (2014) revealed a substantial set of positive exclamations are found by the classifier as strong predictors. Exclamations are another means to make an utterance hyperbolic and thereby sarcastic. This study helps in narrow down the scope for data collecting by searching through the hashtags.

Hyperbole is widely used in daily conversation. However, this feature seems to be a bit obscure whether it is belongs to irony or metaphor. In "Hyperbolic Language and its Relation to Metaphor and Irony" written by Carston and Wearing (2015) would like to examine the range of hyperbole in order to identify in which tropes did it belongs to. In this study, the researcher used the two focal tropes, which are metaphor and irony. In order to examine the nature within these two, hence the researcher looked based on its plain particular properties. This study found that although hyperbole has the features that is quite similar with every one

of these figures, it does not adjust any longer with one than the other and it varies from both in significant manners. This study reveals that hyperbole does not align closely with either metaphor or irony but is a distinctive figure of speech on its own.

Sarcasm known as one of the difficult tasks in Natural Language Processing (NLP), hence a research conducted by Bharti et al. (2017) entitled "Hyperbolic Feature-based Sarcasm Detection in Tweets: A Machine Learning Approach" was trying to propose a hyperbolic feature-based sarcasm detector for twitter data. In this study, the researcher used an automated system to detect sarcastic sentiment to collect the data such as bigram and trigram hyperbolic phrases as one of the features. The findings found out that sarcastic tweets usually contain hyperbolic features either intensifier or interjection. There is also presence of adjective and adverbs detected, which always act as an intensifier in the textual data. The findings of this study revealed that those hyperbolic features such as intensifier, interjection, adjective and adverbs seems to have higher tendency to presence in sarcastic tweets. Although this study had been done according to machine learning based, however it is important to know the common pattern in the tweets. Hence, this study discussed the pattern that usually occur in sarcastic tweets, which the presence of hyperbolic features.

In conclusion, hyperbolic features can be identified as one of the uniqueness in languages. People tend to use these features to emphasize something, however it also can be considered as being sarcastic depending on its context. The widely and frequently used hyperbolic features in textual data everyday life has leaded to becoming a pattern that can be commonly found and seen in any textual data. Although hyperbolic features did not belong to both metaphor or irony, however it stands on its own way and make the language more unique itself.

# CHAPTER 3

# METHODOLOGY

## 3.1 Introduction

This study was carried out by analysing the presence of sarcasm that is found in social interactions on social media, Twitter among the users and how it affects the polarity of the sentences in tweets. Besides, this study also identified the frequency of words used and the usage of hyperbolic features was analysed accordingly. Qualitative approaches were used to analyse the collected data and more than one method was used for this research in order to get authentic data for better findings.

## 3.2 Research Design

In this study, qualitative research design was used where it is mainly based on content analysis. This method was chosen mainly due to the nature of the research questions, which deal with textual data collected in one of the social media platforms, Twitter. Content analysis was used to allow a closeness to text which can alternate between specific categories and relationships which also statistically analyses the coded form of the text. This qualitative research included the content analysis based on the sentiment polarity, hyperbolic features and thematic analysis.

### 3.3 Selection of corpus

This research was conducted by the researcher using a systematic sampling method to collect the data. Systematic sampling is a type of probability sampling method in which sample members from a larger population are selected according to a random starting point but with a fixed, periodic interval.

The researcher had chosen approximately five hundred (500) tweets on Twitter (about 10,000 to 12,000 words accumulatively) and then filtered the collected tweets by some criteria such as all the tweets must be from active users that use certain hashtags such as #Sarcasm, #Sarcastic, #Not, and #Justsaying. Next, based on the tweets, the researcher also filtered the word count in each tweet by choosing the tweets that contain a minimum of five (5) words. This action is to ease the researcher to analyse the tweets according to sentiment polarity also in order to ease the interpretation of meaning. Besides, in order to analyse the data, the chosen tweets must only contain texts, hence the graphics or videos were removed. In order to avoid any biased behaviour and actions occur during the sampling and conducting the research, the researcher also decided to limit the tweets chosen from the same user is not more than three (3) tweets per user. The tweets selected were mainly in the English language. Also, the use of slangs and dialects also was included as long it is understandable by the researcher. Finally, after the total of five hundred (500) tweets successfully collected by the researcher, the researcher checked each of the tweets manually by erasing the tweets that did not follow all the criteria.

### 3.4 Instruments

In this study, the researcher used two instruments in helping the analysis of the selected tweets. However, the researcher did not depend fully on these instruments, which means that the researcher also had manually go through the selected data as a step of making sure that the results given by the instruments had been labelled correctly according to its polarities and words counts.

#### 3.4.1 SummarizeBot

SummarizeBot is an application that can help to analysis the sentiment analysis of words. The researcher used this application to identify the polarities of the words used in the tweets. Below are the examples on how these tools was used in this study:

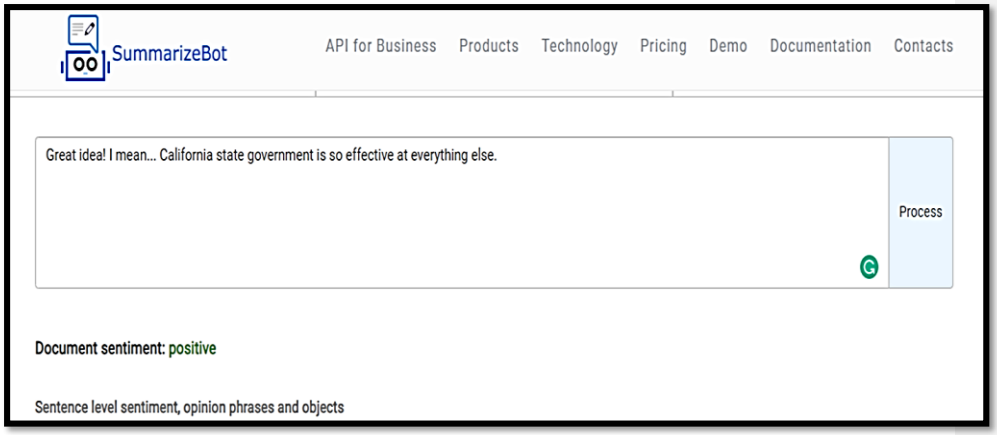



Figure 3.1: *The example of SummarizeBot tools*

#### 3.4.2 AntConc

AntConc is a program for analysing electronic texts (that is, corpus linguistics) in order to find and reveal patterns in language. This software is one of the most well designed and easy to use corpus tools.

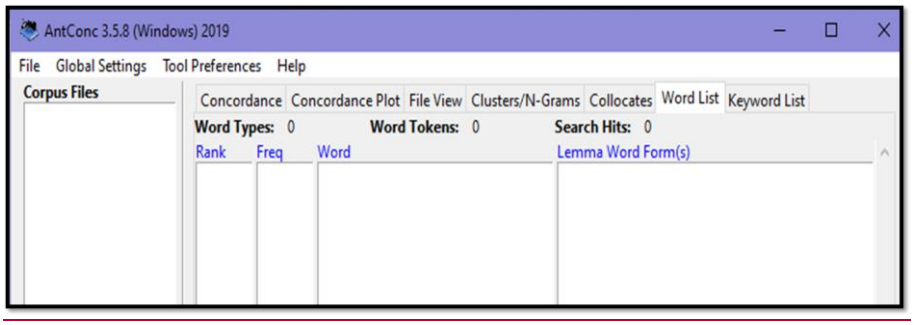



Figure 3.2: *The example of AntConc software*

## 3.5 Data collection procedures

The data is collected by adopting a corpus of tweets that had been divided equally by using the hashtags such as #Sarcasm, #Sarcastic, #Not, and #Justsaying by searching the tweets using the search tools in Twitter. The list of tweets that appear for each hashtag then had been filtered based on the criteria that had been decided. Next, the researcher also copied only the texts in the tweets, which excluded all the graphics, images, emojis and more into text format that will also include their username (e.g: @blackrose_ etc.). During this data collection procedure, the hashtags that used to select the data also had been eliminated. This action is to address the concern of Davidov et al. (2010), which he stated that tweets with #hashtags are "noisy", hence all tweets will be filtered where the hashtags of interest were eliminated during analysis process.

Table 3.1

*Table of frequency linked between the polarity and perceived tweets*

| **Polarity** | **Number of tweets** |
|---|---|
| **Positive** | …. |
| **Negative** | …. |

Table 3.1 is used to tabulate the final aspect of the study, which is to determine which polarity of sentiment that have the highest number of counts according to SummarizeBot. Based on this frequency table, the researcher was be able to see which one of the polarities that frequently being used by the users on Twitter. Based on the tweets that have sarcasm elements, the researcher decides to set a minimum of fifty (50) tweets per hashtags, if the minimum numbers do not follow or do not have the equal numbers for each hashtags, the researcher then tried to collect the data again until it reaches the expected numbers.

### 3.6 Data Analysis Procedures

For the data analysis procedures, the researcher adapted the procedures that had been proposed in the framework by Bharti et al. (2016) in their research entitled "Sarcastic Sentiment Detection in Tweets Streamed in Real Time: A Big Data Approach". To address to the objectives of this study and also adapting the procedures from the framework, there are four parts that had been focused on in this study, which firstly the researcher analysed the sentiment polarities of the tweets, next analysed the words frequency for each polarities, analysed the hyperbolic features found in the tweets and lastly made a comparison between the two polarities. Below are the adapted procedures from the framework:

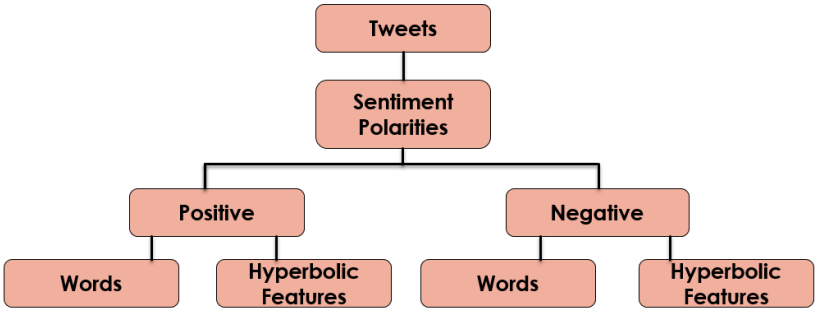



Figure 3.3: *Adapted from the framework proposed by Bharti et al. (2016) in Sarcastic Sentiment Detection in Tweets Streamed in Real Time: A Big Data Approach.*

### 3.6.1 Identifying the Sentiment Polarity using SummarizeBot

The tweets selected firstly were analysed using SummarizeBot to identify the sentiment polarity either it is a positive sentence or negative sentence. However, the researcher only used this application to help fasten the time identifying the sentiment polarities, which the researcher also had manually go through and check all of the data given to make sure the polarities being labelled according to its context. These are the examples of the sentiment polarities using SummarizeBot:


Figure 3.4: *The examples of analysing polarities of sentiment on SummarizeBot*

### 3.6.2 Determine the frequency of words choice using AntConc

The tweets filtered by SummarizeBot according to its polarities then were analysed using AntConc software to gather the data of the highest words used in sarcastic tweets. By gathering this data, the researcher able to identify the intensifier either adjective or adverbs had often used in sarcastic tweets from the words count and frequency. Below are the example using AntConc software:

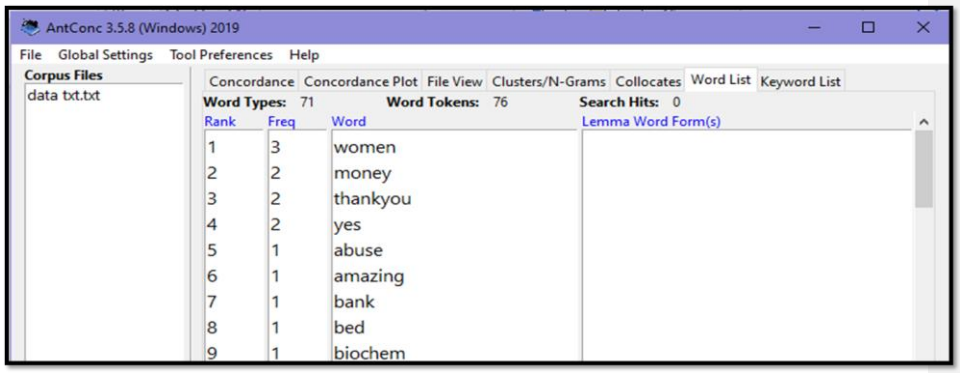



Figure 3.5: *The examples of analysing words counts using AntConc*

### 3.6.3 Determining the hyperbolic features

The tweets filtered by SummarizeBot also were analysed manually by the researcher to identify the hyperbolic features used in the sarcastic tweets. This is done based on the coding framework by Bharti et al. (2016). Below is the example of table that had been tabulated according to its hyperbolic features:

Table 3.2

*The examples of selected tweets tabulated according to its Hyperbolic Features.*

| Tweets containing Hyperbolic Features | | |
|---|---|---|
| Interjections (an abrupt remark, especially as an aside or interruption) | Punctuation marks (abuse of punctuation, repeated use) | Quotes (using quotation marks to indicate the intensity of sarcastic term) |
| **WOW!** Wish I could be perfect and order free talented like this guy. | Didnt notice you were gone, good for you**!!** lol | Nope. Nothing unusual about a Central California **"farmer"** arranging meetings halfway across the world. |

### 3.6.4 Comparing the hyperbolic features between two polarities

The researcher also gathered and tabulated the frequency of hyperbolic features that are often being used in different polarities for comparison and discussion. From this section, researcher able to tabulate the most frequent types of Hyperbolic Features used in both polarities. Below is the example of table tabulated:

Table 3.3

*The examples of table of Hyperbolic Features according to its polarities.*

| Hyperbolic Features | Number of Occurances | |
|---|---|---|
| | Tweets with Positive Sentiment | Tweets with Negative Sentiments |
| Interjection | | |
| Punctuation mark | | |
| Quotes | | |

In order to ensure the validity and reliability of the data analysis, the researcher also had made multiple checking of all the data that had been analysed. Next, the researcher also had double checking with the supervisor to validate the data and the analysis.

# CHAPTER 4

# RESULTS AND DISCUSSION

## 4.1 Introduction

As described in the methodology of this research, the data for this study were collected using three main instruments: the online social media Twitter, the Antconc software, and the SummarizeBot tool. This study then analyses the tweets collected to decide its polarity. This chapter will report on the analysis and explain the research design that had been used in this research. This chapter will also describe the difference in terms of polarities in sentiment, whether the sentiment is positive or negative. Besides, this chapter will explain the hyperbolic features that are found in the data selected. As described in the methodology of this research, the data of this research were collected and gathered using one main instrument, which is data corpus mainly on *Twitter*. The purpose of this study is to analyse the hyperbolic features of sarcastic tweets from different sentiment polarities, which is positive and negative sentiment polarities. The tweets are selected based on the frequently used hashtags to convey sarcastic statement, which is #sarcasm, #sarcastic, #not and #not. The hashtags were selected referring to the past research in filtering the corpus. Users tend to use the hashtags whenever saying something sarcastic.

## 4.2 Selection of Tweets

The selected Tweets were decided beforehand by the researcher in order to standardise the selection of the tweets. First, the tweets must be only in English language. Those tweets that have a mix of language or the language that cannot be identified by the researcher will not be chosen. This action is important for the researcher to analyse and understand the meaning of the tweets. Furthermore, the tweets must contain at least five words in order to ease the interpretation of the meaning. Lastly, the tweets selected also must not exceed three times from the same user. This is to avoid biased behaviour while analysing the corpus.

The data selected were analysed into differences parts. Firstly, the tweets were divided according to their sentiment polarities, whether the sentiment is positive, neutral, or negative. Positive sentiment usually contained more positive words, such as "love, happy, awesome etc.", neutral sentiment usually contained the mixture of positive and negative word or containing words that are neutral. In contrast, negative sentiment usually contains more negative word such as "sad, terrified, angry, etc.". However, this study focused on the positive and negative sentiments. In this research, a total of 500 tweets had been selected, including 100 of it collected from the Malaysian users. Within the 500 tweets, the researcher eliminated the data from neutral sentiment polarity, as this study wanted to see the differences between positive and negative sentiments. After deciding its polarity, the researcher then identified the most frequently used types of words found in sarcastic tweets. Thus, the researcher used the Antconc software to gather the data. This method helps the researcher to form a ranking of the frequency of words for each polarity. Next, to follow the framework of this study, the data thus had been analysed to identify the hyperbolic features found in this corpus. Lastly, a comparison of the hyperbolic features found had been explained. The results of this study are presented and discussed in detail and aligned with the research objectives mentioned above.

## 4.3 Sentiment Polarities

Deciding the sentiment polarities are important for opinion mining and content analysis. This is because the polarity of words can be the meaning-switcher according to its polarity. Sentiment Polarity is the expression that determines the sentimental aspect of an opinion. Polarity analysis takes into account the amount of positive or negative terms that appear in a given sentence. In this section, the results for the frequency of words choice found in sarcastic tweets are reported. This is to fulfill the first objectives of this study, which to identify sentiment polarities in sarcastic tweets. In this study, the researcher found out that positive polarity has the highest frequency in the corpus selected. Below is the table for the sentiment polarity for the corpus selected in this research:

Table 4.1

*The total number of Tweets based on sentiment polarity.*

| **Sentiment Polarity** | **Total** |
|---|---|
| **Positive** | 212 |
| **Negative** | 144 |

Based on Table 4.1, the highest number for sentiment polarity detected from the corpus of 500 tweets is positive polarity with a total number of 212 equal to 42.4% followed by neutral sentiment polarity and negative sentiment polarity with both total of 144 tweets equal to 28.8%. Based on the table 4.1 above, it is proven that people tend to use positive words when conveying negative message. According to Rillof et al. (2013), the usage of positive words referring to negative situation mostly due to the intention of being sarcastic. Sarcasm is known as a positive way to convey negative meaning. People tend to use positive words when

conveying sarcastic conversation, however it is quite difficult to differentiate and get the real meaning in sarcastic text or conversation intended by the speaker. Therefore, it is important to understand the context of a sentence before deciding its polarity, due to the context of the sentences, some words might convey different meaning. In order to identify the polarity of words, it is important to understand the context of the text first. Below is the example of tweets according to its polarities:

Table 4.2

*The examples of tweets according to its sentiment polarities.*

| **Sentiment Polarity** | **Tweets** |
|---|---|
| **Positive** | I **totally trust** that the FBI will give us the **truth** |
| | It's **okay** if someone **don't like** you. Not everyone has **good taste**. . |
| | Maybe a **good** reason America, why not everyone need to have weapons at home? |
| **Negative** | **Desperately wanting to have** a heart-to-heart with someone but being way **too awkward/anxious** to even try |
| | Strong people don't put others down. They lift them up and **slam** them on the ground for maximum **damage**. |
| | **Breaking** the law sound **stressful**. |

The examples in Table 4.2 shows that words can be the polarity switcher for the sentences, for example, the positive sentiment polarities considered as positive because it contains more positive words than negative words, therefore, the sentence is identified as positive. When a sentence has both positive and negative sentiment, it depends on the highest level of the word, for example, if the word has more tendency of positive sentiment, hence, the whole sentence

considered as positive sentence, and vice versa. Generally, sentiment polarity is a particular feature of any textual data. It is usually divided into three which are positive, neutral and negative. Positive polarity is when the texts or sentiments contain more positive words than other polarities, neutral polarity is when the texts did not belongs to either positive and negative, or in a condition where positive polarity and negative polarity is in equal, and negative polarity is when most of the words in the texts convey negative messages. In addition, a document that containing more than one opinionated statement would have a mixture of the polarities, so the final polarity of the sentiments will be decided by identifying the most presence polarities in the sentiment. Contextual polarity refers to the polarity of a word when considered in its context, as opposed to prior polarity when it stands alone. For example, the word "strong" is a positive word, however the positivity of this sentence had been decreased when the user also used the word "slam" and "damage" in the sentences. Thus, the negativity of the sentence is more to positivity, therefore the tweet is considered as negative.

### 4.4 Frequency of words

A test had been run to identify the most used or the frequency of words choice in sarcastic Tweet. This is corresponded to the second objective for this study which is to determine the frequency of words choice and hyperbolic features of sarcastic tweets. In this section, the results for the frequency of words choice found in sarcastic tweets will be reported. The researcher used a computer software, AntConc. to identify this pattern. The tweets that already filtered by SummarizeBot according to its polarities then had been analyse using Antconc. software to gather the data of the highest words used in sarcastic tweets. In this section, researcher only focus on adverbs and adjectives and nouns. The functional words were excluded hence it does not convey any contextual meaning. According to Bharthi et al. (2017), there are two other elements in hyperbolic features, which is either intensifier of interjection.

In a situation of the presence of an adjective or an adverb, it is usually act as an intensifier in the textual data. For example, "I am *extremely happy*", "I like it *so much*", and more. This study found out that the highest words count used is Adverbs. Below is the table tabulate the example of words choices for each polarity:

### 4.4.1 Positive words

Table 4.3

*The most frequent words choice for positive sentiment*

| Rank | Frequency | Word | Classes |
|---|---|---|---|
| 1 | 18 | Good | Adjective |
| 2 | 17 | Well | Adverb |
| 3 | 15 | Great | Adjective |
| 4 | 12 | Much | Adjective |
| 5 | 9 | Really | Adverb |
| 6 | 6 | Better | Adverb |
| 7 | 5 | Best | Adjective |
| 8 | 4 | Awesome | Adjective |
| 9 | 4 | Brilliant | Adjective |
| 10 | 3 | Actually | Adverb |

### 4.4.2 Negative words

Table 4.4

*The most frequent words choice for negative sentiment.*

| Rank | Frequency | Word | Classes |
|---|---|---|---|
| 1 | 19 | All | Adverb |
| 2 | 15 | Like | Adverb |
| 3 | 9 | More | Adjective |
| 4 | 6 | Some | Adverb |

| | | | |
|---|---|---|---|
| 5 | 5 | Bad | Adjective |
| 6 | 5 | Much | Adverb |
| 7 | 4 | Already | Adverb |
| 8 | 4 | Awesome | Adjective |
| 9 | 4 | Happy | Adjective |
| 10 | 3 | Good | Adjective |

The data in Table 4.3 and Table 4.4 above shows that the frequent words choices in sarcastic tweets are adverbs and adjectives. Based on the result above, each polarity for sarcastic tweets used adverbs and adjectives as its word choices while writing a sarcastic tweet. Utsumi and Akira (2000) discussed about the extreme adjective and adverb and how the presence of these two intensifies text. Most often, it provides an implicit way to display negative attitudes, for example, sarcasm. Sarcasm had been used widely either in social medias or real-life situations. According to Van Mulken and Schellens (2012), an intensifier is a linguistic element that can be removed or replaced while respecting the linguistic correctness of the sentence and context, however it is resulting in a weaker evaluation. A commonly used way to intensify utterances is by using word classes such as adverbs ('very') or adjectives ('fantastic' instead of 'good'). It may be that senders use such intensifiers in their tweets to make the utterance hyperbolic and thereby sarcastic, without using a linguistic marker such as '#sarcasm'. Sarcasm in daily life had cause confusion within the society because of the positive words used instead of negative words. Liebrecht et al. (2013) showed that sarcasm is often signaled by hyperbole, using intensifiers and exclamations; in contrast, nonhyperbolic sarcastic messages often receive an explicit marker.

## 4.5 Hyperbolic Features

Hyperbolic features are commonly the act of exaggerate or over-statement expression in language. In this study, four hyperbolic features had been selected to analyse the data, which is the intensifier, interjection, quotes and punctuation mark. However, this section will only report for interjections, punctuation marks and quotes, as the intensifiers had been discussed at previous sub-unit. This is to fulfil the following second objective for this study, to identify the hyperbolic features in sarcastic tweets. The tweets filtered by SummarizeBot then had been analysed manually by the researcher to identify the hyperbolic features used in the sarcastic tweets. This is done based on the coding framework by Bharti et al. (2016). Below is the diagram of the number of hyperbolic features found for each category:

Table 4.5

*The example of tweets according to the hyperbolic features.*

| **Tweets containing Hyperbolic features** | | |
|---|---|---|
| **Interjections (an abrupt remark, especially as an aside or interruption)** | **Punctuation marks (abuse of punctuation, repeated use)** | **Quotes (using quotation marks to indicate the intensity of sarcastic terms)** |
| **Wow**. What a difference a cup of coffee, nap, and a shower make | Tomorrow and Thursday are going to be SO MUCH FUN**!!!** | **"**The military and police said, `Sir, we can do it,**'"** (as we have seen since 2016) |
| **Ooh ya** may be thats why soviet and china in 60s were so RICH and currently sweden is having million dollars per capita incomes | Does giving up someone for lent actually count**!?!** This is brilliant! Now maybe I can keep coke and candy in my life! | Me and my sons are now **"successful"** factory workers, but we gotta lot of new neighbors that bother me! |
| **Gee**, that would be so shocking. | My crush rejected me last night **???¿¿** | Sick of saying **"**that's why we're champions.**"** |

| | | |
|---|---|---|
| **Oh**, I do apologise. I suppose you've never had an opinion about anything that doesn't directly affect your life then, oh sainted one? | You never hear of a transgender female participating in men's sports and winning**!!!** I wonder why that is? | Maybe we should be asking the Chinese Gov about this medicine that can move you from **"dead"** to **"critically ill"**. How can anything they tell us be trusted? |
| **Wow**. What a difference a cup of coffee, nap, and a shower make. | I don't know whether to laugh at this or feel pity over the system **!!!** | I mean, DAMN son, at least "pretend" you're not looking for me! |
| **Oh**, so Yun doesn't blur their faces? Do stans know that? | Humans proving again how decent they are**!!!** | Make myself the "freely elected dictator for life of the world." . |
| **Yeah**, he's well known as a rule of law kind of guy. | You look great**....** I mean it**….** | Quite surprised which professions are suddenly **"unskilled/low skilled**". |
| **Wow** hard to image all this bad economic news coming just AFTER the election! Guess nobody knows what's causing this | Angie, I don't have a job in libraryland, and I'm unqualified**...** so come on, where's my job**?!!** | "I have sacrificed so much for this job I am not qualified for, nepotism cost me everything" |
| **Yay** radio broken. Exactly what the gym needed. | I had no clue**!!!** What a dip-shit. | Finally saw **"Parasite"**. Wasn't expecting the spaceship... |
| **OMG**, I have no doubt that Trump was high on drugs during this press briefing! He's gacking SO BAD (watch his mouth)! Trump should be tested for cocaine, Adderall, OxyContin and Vicodin in his system. He needs to be removed ASAP for the sake of Our Nation! | Tech**...** pass. If technology companies are really honest about security, they should talk to banks. Web or ATM PINS aren't a mile long but they secure our financial records. oh right, these companies need subscriptions, more user data**...** sorry, my bad. | What a surprise. This is the Pinnochio that called traumatic brain injury a **"headache"** and he has grown by an inch since he became president. My dog knows more about COVID19 than the chief scientist, law enforcement officer, commander in chief and super businessman |

In this study, the researcher chose the hyperbolic features of interjections, punctuation marks and quotes according to several requirement, the interjection must have an abruption remarks, such as “Yay, Well, Oh, Wow, etc.”. The punctuation marks features were selected when the tweets contained repeated use of marks or symbol such as the repeated full stop marks, punctuation marks, and question marks. Lastly, the quotes features were selected if the tweets have the quotation marks (e.g: “hate”). Bharti et al. (2016) stated that a hyperbolic text contains one of the text properties such as intensifier (e.g., Adjective or Adverb), interjection (e.g., wow, aha), quotes (e.g.,” ” or ‘ ’), punctuation mark (e.g., ????, !!!!). In addition, Kreuz et al. (2007) also stated that the other hyperbolic terms such as interjection and punctuation and showed how hyperbole is useful in sarcasm detections. Based on Table 4.5, it is proven that hyperbolic features are detected in sarcastic tweets.

The result of this study found that the interjection mostly used in the beginning of a sentence. This action is due to the speaker as an act to interrupt a conversation or wanting to start their sentences with something insulting to make people feel uneasy about it. In addition, interjection also can be used to as a way to express their real emotion. Tweets is known with the nature of writing text as the spoken language, hence it has high tendency for the people used interjection as a way to express their real emotion. This study found some interjection such as “wow, oh, well, OMG, gee, and yeah”. The punctuation marks for this data is the repeated of punctuation marks, full stop marks, and question marks, however there are also few tweets that has the combination of marks, such as “!?!”, “?!!” and “???¿¿”. As for the quotes, there are few tweets that use the quotation marks (“”) to highlight certain words, however there are also tweets that quoting some text or speech from other people. Besides, instead of the tweets that contains either one of the hyperbolic features, there are also tweets that use the combination of two of the hyperbolic features or combine all the tree hyperbolic features. Below is the example of the tweets that have the combination of those hyperbolic features:

Table 4.6

*The example of tweets with the combination of hyperbolic features.*

| **Tweets** | **Hyperbolic Features** |
|---|---|
| **Yes**, THAT'S why so many are doing so much better under his administration, because **"**he only cares about himself**". Uh huh**. Sure. Ok. (<— for the idiots out there) | Interjection and quotes |
| **"Oh,** no! Why are people watching a video of kids? What are they pedos?**"** | Interjection and quotes |
| I mean well **...** **"**My Honesty is Loyal**"** | Punctuation marks and quotes. |
| I've always found. people kind of annoying. Always with the **"**That's an AWESOME idea**..."** and **"**I guess that next you'll tell me**..."**. No, thank you. | Punctuation marks and quotes. |
| **Yeah** let the Scientists in, they may find a cure for Coronavirus and not spread it **!!!** | Interjection and Punctuation marks |
| **Haishh** cik abg**...** I know your size is 3 times bigger than me**…** but it's not the reason for you to cut que I**..** | Interjection and Punctuation marks |
| **yeah…** of course, surely but **"**SLOWLY**" ...** | Interjection, Punctuation Marks, and Quotes |

Based on Table 4.6, the presence of the combination of hyperbolic features were detected in this study. At the point when exaggeration blends in with either metaphor, or irony, the outcome is certainly not another enveloping compound where one figure expands on another. Rather, the figure that is principally imparted is shaded with hyperbolic hints. The usage of more than one hyperbolic feature is consider as being extra overstatement in a

sentence. User might purposely use this combination to make sure people get the idea of him or her is being sarcastic. It is important to consider overstatement in hyperbolic figures similar to an interpretive impact, regulating the trademark impacts of the figure it blends in with. This is the reason overstatement is so flexible in blending in with a wide range of interesting expressions.

### 4.6 Comparison of the hyperbolic features between polarities

The researcher also gathered and tabulated the frequency of hyperbolic features that are often being used in different polarities for comparison and discussion. From this section, researcher able to tabulate the most frequent types of hyperbolic features used in both polarities.

Table 4.7

*The total number of occurrences for each hyperbolic feature.*

| Hyperbolic Features | Number of Occurrences | |
|---|---|---|
| | Tweets with Positive Sentiment | Tweets with Negative Sentiment |
| Interjection | 42 | 17 |
| Punctuation Mark | 58 | 44 |
| Quotes | 9 | 18 |

Table 4.7 shows the total number of occurrences for each hyperbolic feature according to its sentiment polarity. The highest number of occurrences for both sentiment polarity is punctuation marks, with a total number of 58 tweets in positive sentiment polarity, and 44 tweets in negative sentiment polarity. Although the pattern is quite similar, but the total number

of occurrences is differed for the second highest and the lowest. Positive sentiment polarity recorded interjections, however negative sentiment polarity recorded quotes as the second highest number of occurrences. The gap between the number of occurrences for quotes and interjections is quite large. Due to combination of hyperbolic features usually happened in sarcastic tweets. The table below tabulated the hyperbolic features according to its fraction of tweets detected for each interception:

### 4.6.1 Positive sentiment

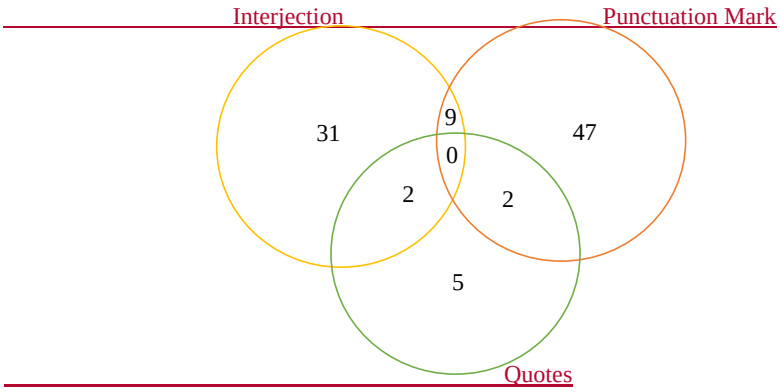



Figure 4.1: *The numbers of hyperbolic features for each category in positive sentiment.*

Based on Figure 4.1, the highest number of hyperbolic features found in positive sentiment is punctuation marks, with a total number of 47, followed by interjection with 31 tweets, and quotes with 5 tweets. The table above able to show the interception that occur in combining two or more hyperbolic features. The highest number of combinations occur in positive sentiment is between punctuation mark and interjection, with 9 number of occurrences. The number of interceptions between quotes and punctuation marks, and quotes and

interjection are the same, which is the total number of 2. In this section, there are no interception of combining all the three hyperbolic features.

### 4.6.2 Negative sentiment

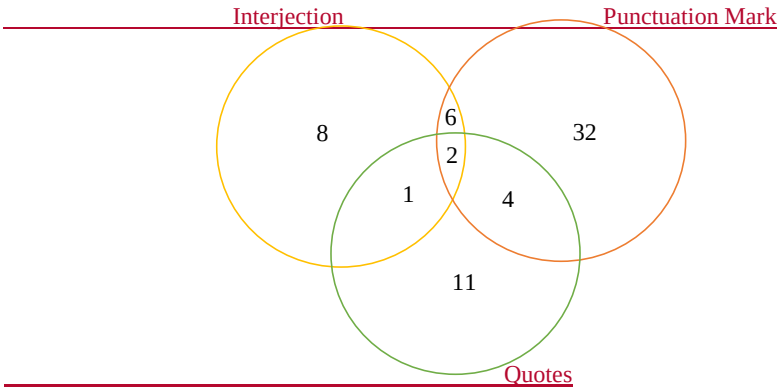



Figure 4.2: *The numbers of hyperbolic features for each category in negative sentiment.*

Figure 4.2 shows the fraction of hyperbolic features according to its interception. The highest number of hyperbolic features found in negative sentiment is punctuation marks, with 32 number of tweets, followed by quotes with 11 tweets and interjection with 8 tweets. The highest interception detected from negative sentiment polarity is also between punctuation marks and interjection, with 6 tweets. The second highest interception detected between quotes and punctuation marks with 4 tweets, followed by the combination of all the three hyperbolic features. The least combination occurs between quotes and interjection.

## 4.7 Discussion

Sentiment polarity is a feature of the text. It is usually dichotomised into three either it is positive, neutral, or negative, but polarity can also be thought of as a range. In this study, the result found out that most of sarcastic tweets used to convey negative messages by using positives words. Therefore, the highest polarity found with sarcastic remarks of hashtags is positive sentiment. Rillof et al. (2013) revealed that the usage of positive words referring to negative situation mostly due to the intention of being sarcastic. It is proven that people tend to use nice words such as praising, convincing, and joking when they are intended to say something sarcastic instead of uttering harsh and bad words. Sarcastic also can be considered as creativity in language, because when sarcastic element was added in a sentence, the chance of being ambiguous is higher, hence people will start guessing about the real meaning that the speaker wanted to convey. This also could arise confusion between the people who involve with the conversation. In addition, Ezzah (2019) revealed that the use of sarcastic language really works to mock and insulting. This statement was supported by Gibbs (2000) who stated that sarcastic language has some elements or sentiments that can lead to negative messages. Based on the result from the data collected, the researcher believes that most of the tweets were intended to mock and insulting others, this conclusion arise after trying to understand the meaning as well as the context of the tweets itself.

Besides, this study also found several patterns that always appear in sarcastic tweets, for example, the hyperbolic features such as interjections, quotation marks, punctuation marks, adverbs and adjectives. In a study by Bharti et al. (2016), they stated the usage of hyperbolic features in tweets also can be considered as sarcastic. In addition, a study by Bharti et al. (2017) also revealed that those hyperbolic features such as intensifier, interjection, adjective and adverbs seems to have higher tendency to presence in sarcastic tweets. Another study conducted by Abulaish and Akmal (2018) also stated that hyperbolic tweets usually employed

exaggeration and over-statement. Hence as an alternative to detect the sarcastic tweets, the researcher used a method by counting the frequency of interjections, question marks, exclamation marks, and any two intensifier-related hyperbolic features such as frequency count of adverbs and adjectives. In this study, the number of tweets that contains hyperbolic features such as interjection, quotation marks, and punctuation marks are 231 tweets over 500 tweets that had been analyses. After eliminated the neutral sentiment polarity of 144 tweets, the researcher found the hyperbolic features were detected 109 tweets out of 212 tweets for positive sentiment, and 79 times out of 144 tweets for negative sentiment. With a percentage of 51% and 55% for each category, this pattern can be considered as one of the common patterns used in sarcastic tweets. As a comparison with the previous studies, this study proved that hyperbolic features as one of the common patterns in sarcastic tweets. As mentioned by Bharti et al (2016), another hyperbolic features that usually found in sarcastic tweets is intensifier. Intensifier in hyperbolic features is an act of using lots of adverbs and adjectives to convey something sarcastically. In this study, the researcher found out that adverbs and adjectives are the most frequent used words in both sentiments. Hence, the presence of adverbs and adjectives as the most frequent used words in sarcastic tweets also strengthen the hypothesis, which adverbs and adjectives also as common features found in sarcastic tweets. As discussed by Utsumi and Akira (2000), a situation that uses an extreme adjectives and adverbs frequently provides an indirect way in exhibiting negative attitudes, such as sarcasm.

Although hyperbolic features had been discussed in other studies, however it is never been done according to its sentiment polarity. After conducting an observation on the hyperbolic features according to its sentiment polarity, the researcher can conclude that the usage of hyperbolic features detected more than 50% for both sentiment polarities. This situation proves the probability of people intended to say or convey something bad while inserting the hyperbolic features in order to make it sounds more insulting. As explained by

Liebrecht et al. (2013) using an example of saying "fantastic weather" when the actual weather is raining is classified as sarcastic by inserting the adjectives words, instead of uttering a non-hyperbolic utterance, "the weather is good". In addition, Abulaish and Kamal (2018) stated that the presence of these intensifiers and hyperbolic features adds another over-exaggeration component in a tweet. Therefore, it is proven that people use hyperbolic features to make their utterances sound more sarcastic.

On the whole, there is a difference in terms of how sentiment polarity affected the hyperbolic features from the data collected. Negative sentiment used hyperbolic features to make it sounds more annoying and insulting, however positive sentiment used hyperbolic features to express their true emotion in a nicer way.

# CHAPTER 5

# CONCLUSIONS

## 5.1 Introduction

This chapter summarises the present study as well concludes the results. It also provides recommendations for possible future research and implications of the findings. The chapter ends with a concluding remarks.

## 5.2 Summary of Findings

The aim of this study is to analyse hyperbolic features of sarcastic Tweets from different sentiment polarities. This research applies qualitative methods to analyse and investigate the sarcasm usage found among the users in the context of an asynchronous social network platform (Twitter). This research focuses more on the linguistic features that are classified under the words' polarity and sentiment analysis. The objectives of this research include:

iv. To identify the sentiment polarity of sentences used in sarcastic Tweets.

v. To identify the frequency of words and hyperbolic features of sarcastic Tweets.

vi. To compare hyperbolic features of sarcastic tweets with positive sentiment and negative sentiment.

The study used content analysis in identifying the sentiment polarities and the hyperbolic features that took place in asynchronous online platform, Twitter by adopting the framework proposed by Bharti et al. (2016) in Sarcastic Sentiment Detection in Tweets Streamed in Real Time: A Big Data Approach . The data from content analysis were supported by using two main instruments which are SummarizeBot, for identify sentiment polarities and Antconc. software to identify the frequency word choices. These two instruments were used as a main methodology to analyses the data collected.

The findings from the sentiment polarities showed that positives sentiment was more frequently detected compared to negative sentiment. The findings match those of Bharti et al. (2016). This is proven that people tend to use positive words to convey negative meanings, this resulting in the higher number of tweets detected from positive sentiments. In terms of hyperbolic features, the findings revealed that the frequency of words choices have more usages of adverbs and adjectives, as the intensifier. Intensifier is known as one of the hyperbolic features. On the other hand, the interjections, punctuation marks and quotes also highly detected in sarcastic tweets. The percentage of hyperbolic usage for positive sentiment is 51% and 55% for negative sentiment. As for punctuation marks features, it recorded the highest number of occurrences for both sentiment polarity.

In addition, the findings from the comparison between the hyperbolic features between positive and negative sentiment revealed that the combination of two or more hyperbolic features in a sentence was found in sarcastic tweets. Those combination show how the user tend to insert hyperbolic features in their text to make it sounds more awful and insulting.

## 5.3 Implications of Findings

Looking at the demanding and vast evolution of delivering messages, sarcasm is being used as one of the most influential bases on giving other people's view on the intended meaning using opposite words from the actual meaning. The findings from this study can benefit agencies who are interested in advertisement and marketing to produce suitable content for specific field in order to reach out the target audiences. According to Huang et al. (2015), sarcasm is an instigator of conflict but also a catalyst for creativity. General forms of sarcasm promote creativity through abstract thinking for both expressers and recipients. Therefore, by promoting their product with sarcastic elements might help their business growth and able to reach their target audience effectively.

Besides, it will rebound to the benefit in considering that sarcasm plays an important role in daily life either spoken or written. The greater demand for having good communication skills justifies the need for more life-changing knowledge approaches. Social influencers, academic researcher and people who are interested in finding something interesting in studying trending language without realising they use it in their daily conversation among themselves. By knowing certain words that frequently used in sarcastic conversation, it can create awareness among the society about sarcastic words used, and further problems such as miscommunication can be avoided.

People tend to get confuse when the other people in the same conversation being sarcastic. It can create ambiguous meaning and it some people may not understand the real meaning the speaker tried to convey, this can cause misunderstanding in the conversation itself. Therefore, it is important for people to know how people sarcastic was applied and frequently used in daily life. The result of this findings can give a clearer picture on how the sarcastic

language was used. For example, when a person uses lots of hyperbolic features in a text or conversation, it has high tendency of being sarcastic.

### 5.4 Direction for Future Research

The present study was conducted on a sample of 500 tweets, with 400 tweets were collected from abroad and 100 of them from Malaysian users. The tweets were collected according to several criteria, such as the tweets collected with the sarcastic remarks and hashtags #sarcasm #sarcastic #not #justsaying. The usage of sarcastic remarks had raised the assumption of the user intended to say something sarcastic. While the sample size is sufficient for statistical analysis, it may not be practical to generalise the findings to the whole population of the users on social media microblog, Twitter. Therefore, further studies should be carried out on a larger scale in other state or countries. This would provide richer data in understanding the phenomenon.

Also, future research may consider conducting a comparative study between the users from different social media platforms, such as Twitter and Facebook, Instagram and WeChat, and more. It should be noted that the content on Twitter are largely text-based and openness in terms of its "retweeting" feature would make it more convenient for its users to freely use all forms of language styles including sarcasm. Further study on such trend in other social media would be beneficial. This would provide a more focused and comprehensive view on the influence on how sarcastic language was used in their daily conversations. Moreover, the study also should be done considering more wide range in terms of their age and cultural background. This will encourage the data representation to be more valid since the numbers of participants from different ages or races affects their understanding in sarcastic languages.

Besides that, although there is a significant positive correlation between sentiment polarities and the hyperbolic features, the correlations generally do not have a really distinctive figure. Thus, this suggest that there may be other patterns or language features that contribute to usage of sarcastic languages. Moreover, this study only conducted using four sarcastic remarks, which #sarcasm, #sarcastic, #not, #justsaying. Hence it is advisible for the future study to use a combination of two hashtags according to specific topics, for example, #sarcasm #sports, #sarcastic #politics, #sarcasm #entertainment. This could provide more broad and wide point of view on how sarcasm was detected according to the topics. The future researcher also can make a comparison on how sarcasm was applied in politics field, sports field and entertainment fields.

In addition, since the tweets focused in this study was only in the aspect of the contextual meaning and the daily unformal conversation of Twitter, future studies could focus on comparing the gender in identifying the meaning making out of the sarcastic messages that brings more specific messages to a certain group of people because this studies could help provide clearer picture of which is more contribute to the participants' knowledge and their views on certain issues that are being conveyed using sarcastic languages.

Lastly, as with most of previous research, this present study shows that the sarcasm has rather used positives words instead of negatives words. However, to identify how hyperbolic features was used in sarcastic tweets, this study used a framework by Bharti et al (2016). As a matter of fact, there are limited study which is focusing particularly on the hyperbolic features, this study had made a new point of view because it analyses the data according to its sentiment polarities. Thus, the past studies had not been confined to a particular type of test.

## 5.5 Conclusion

Despite its limitations, the findings do suggest some valuable information which is the importance of sarcasm as one of the language features and how sarcasm can be considered as uniqueness in a language. This study is derived from the need to identify the sentiment polarity and hyperbolic features in sarcastic tweets. Sarcasm is frequently used in any situation, whether it is spoken or written. This issue has caught an eye for recent years, resulting many studies had been done in detecting sarcasm on social media. However, there were very minimal study done in identifying the hyperbolic features according to its sentiment polarity. This study is able to provide the comparison between positive and negative sentiment based on its hyperbolic features. On the other hand, this research has not only filled the existing gaps but also contributed to an increased understanding on how sarcasm was applied in daily conversations. As people be seeing sarcasm as an act to mock or insult other people, this study has created a better point of view on how people can use sarcasm

## REFERENCES


Abulaish, M., & Kamal, A. (2018). Self-Deprecating deprecating Sarcasm sarcasm Detectiondetection: An Amalgamation amalgamation of Rulerule-Based based and Machine machine Learning learning Approachapproach. In Proceedings of *IEEE/WIC/ACM International Conference on Web Intelligence*, 574-579. 10.1109/WI.2018.00-35.

Alexander, N. S., & Omar, N. (2017). Generating a Malay Sentiment sentiment Lexicon lexicon Based based on WordNet. *Asia-Pacific Journal of Information Technology and Multimedia*, *6*(1).

Barbieri, F., Saggion, H., & 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* (pp. 50-58).

Bharti, S. K., Vachha, B., Pradhan, R. K., Babu, K. S., & 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.

Bharti, S. K., Naidu, R., & Babu, K. S. (2017). Hyperbolic Feature-based Sarcasm Detection in Tweets: A Machine Learning Approach. In *2017 14th IEEE India Council International Conference (INDICON)* (pp. 1-6). IEEE.

Bowes, A., & Katz, A. (2011). When sarcasm stings. *Discourse Processes*, *48*(4), 215-236.

Carston, R., & Wearing, C. (2015). Hyperbolic language and its relation to metaphor and irony. *Journal of Pragmatics*, *79*, 79-92.

Chaffey, D. Global Social Media Research Summary (2016). Retrieved from URL (http://www.—smartinsights.com/social-media-marketing/social-media-strategy/new-glo bal-social-media-research/).

*Collins Cobuild English Language Dictionary*. (1987). London: HarperCollins. Douglas Colin Muecke. (1970). Irony and the ironic. *The Critical Idiom*, 13.

Christodoulidou, M. (-2011).”-Hyperbole in everyday conversations”. *Proceedings of the 19th International Symposium on Theoretical and Applied Linguistics*19: 143- 152.

Ellen, R., Ashequl, Q., Prafulla, S., Lalindra, S., Gilbert, D. S., Gilbert N., & Ruihong, H. (2013). Sarcasm as Contrast between a Positive Sentiment and Negative Situation *Proceedings of the Conference on Empirical Methods on Natural Language Processing* (EMNLP), pages 704–714, Seattle, Washington.

Gibbs, R. W. (2000). *Irony in talk among friends*. Metaphor and symbol, 15(1-2):5{27.


**Formatted:** Indent: First line: 0.5"

González-Ibánez, R., Muresan, S., & Wacholder, N. (2011, June). Identifying sarcasm in Twitter: a closer look. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2* (pp. 581-586). Association for Computational Linguistics.

Huang, L., Gino, F., & Galinsky, A. D. (2015). The highest form of intelligence: Sarcasm increases creativity for both expressers and recipients. *Organizational Behavior and Human Decision Processes*, *131*, 162-177

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

Kreuz, R. J., & Caucci, G. M. (2007, April). Lexical influences on the perception of sarcasm. In *Proceedings of the Workshop on computational approaches to Figurative Language* (pp. 1-4). Association for Computational Linguistics.

Liebrecht, C., Kunneman, F., & Bosch, A. (2013). The perfect solution for detecting sarcasm in tweets# not. WASSA *2013*, page 29.

Lee, C., & Katz, A. (1998). The differential role of ridiculeinridicule in sarcasm and irony. *Metaphor and Symbol*, 13, 1–15.

Maynard, D. & Greenwood, M. A. (2014). ‘‘Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis,’’ in *Proc. 9th Int. Conf. Lang. Resour. Eval.*, May 2014, pp. 4238–4243.

Nasrudin, E. A. (2019). Persepsi Golongan golongan Remaja remaja Terhadap terhadap Penggunaaan penggunaaan Bahasa bahasa Sarkastiksarkastik. *International Journal of Humanities, Philosophy, and Language*, 2 (7), 13-22.

Pang, B., & Lee, L. (2008). *Opinion mining and sentiment analysis*, Foundations and trends in information retrieval, 2.1-2, pp1-135.

Ratanawalee Wimolmas. (2013). A survey study of motivation in English language learning of first year undergraduate students at Sirindhorn International Institute of Technology (SIIT), Thammasat University: Thailand. Foreign Language Learning and Teaching. *Conference Proceedings by* LITU, 2(1)

Riloff, E., Qadir, A., Surve, P., De Silva, L., Gilbert, N., & Huang, R. (2013, October). Sarcasm as contrast between a positive sentiment and negative situation. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (pp. 704-714).

Schapire, & ~~and Y.~~ Singer, Y. (2000). *BoosTexter: A boosting-based system for text categorization,*
Machine learning 39.2-3, pp135-168.

Sharma, S., & Chakraverty, S. (2018). Sarcasm ~~Detection~~ detection in ~~Online~~ online ~~Review~~ review ~~Text~~text. *Division of Computer Engineering, Netaji Subhas Institute of Technology: India.*

Skalicky, S., & Crossley, S. (2018). Linguistic ~~Features~~ features of ~~Sarcasm~~ sarcasm and ~~Metaphor~~metaphor
Production Quality. In *Proceedings of the Workshop on Figurative Language Processing* (pp. 7-16).

Tepperman, J., Traum, D., & Narayanan, S. (2006). "Yeah right": Sarcasm recognition for
Spoken dialogue systems. In *Interspeech* (pp. 1838–1841). Pittsburgh: ICSLP.T.

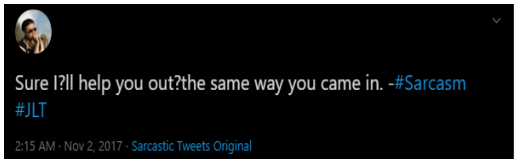
Utsumi, A. (2000) Verbal irony as implicit display of ironic environ-ment: Distinguishing ironic utterances from non-irony, *Journal of Pragmatics* 32 (12), pp. 1777–1806.

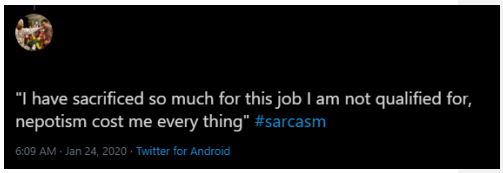
Van Mulken, M., & Schellens, P. J. (2012). Over lood-zware bassen en wapperende broekspijpen. gebruik enperceptie van taalintensiverende stijlmiddelen. *Tijd-schrift voor taalbeheersing*, 34(1):26–53.

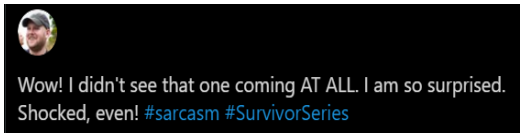
Wales, K~~atic~~. (2001). *A ~~Dictionary~~ dictionary of ~~Stylistics~~ dtylistics* (2nd ed.). London: Pearson Education
Limited.

Wang, P. Y. A. (2013, November). # Irony or #Sarcasm—A Quantitative and Qualitative Study Based on Twitter. In *Proceedings of the 27th Pacific Asia Conference on Language, Information, and Computation (PACLIC 27)*. ~~(pp. 3~~

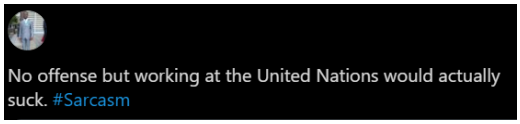
## APPENDICES

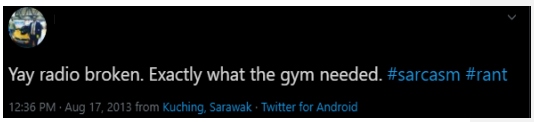


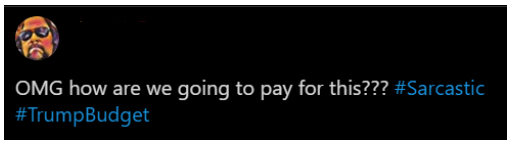



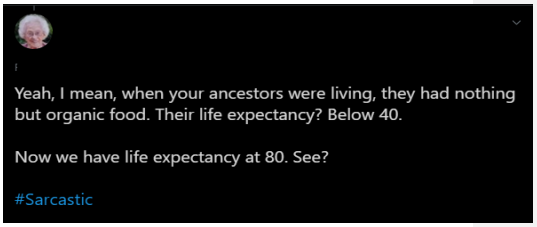





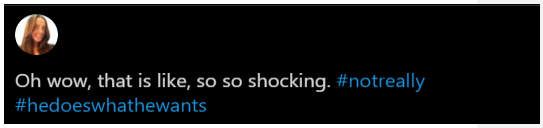


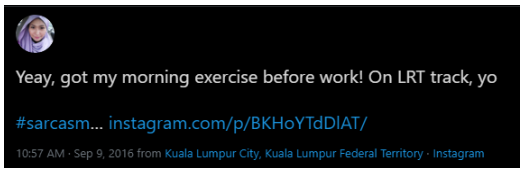



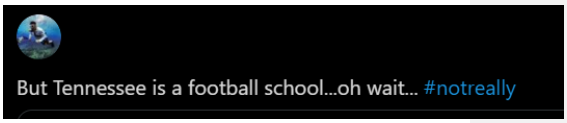



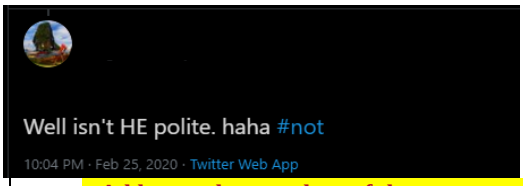



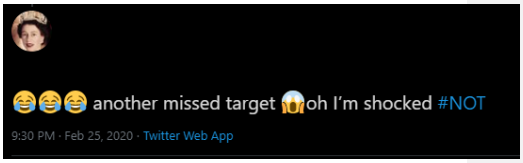





~~Add several screenshots of the tweets you collected here (hide the username)~~

~~Maybe about 4~~