7-7-2020

Multimodal fake news detection using a Cultural Algorithm with situational and normative knowledge

Priyanshi Virendra Shah
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Multimodal fake news detection using a Cultural Algorithm with situational and normative knowledge

by

Priyanshi Shah

A Thesis

Submitted to the Faculty of Graduate Studies

through the School of Computer Science

in Partial Fulfillment of the Requirements for

the Degree of Master of Science at the

University of Windsor

Windsor, Ontario, Canada

2020

Multimodal fake news detection using a Cultural Algorithm with situational and normative knowledge

by

Priyanshi Virendra Shah

APPROVED BY:

M. Baki

Odette School of Business

S. Saad

School of Computer Science

Z. Kobti Advisor

School of Computer Science

May 20, 2020

Declaration of Co-Authorship/Previous Publication

I. Co-Authorship

I hereby declare that this thesis incorporates material that is a result of research conducted under the supervision of Dr. Ziad Kobti (Advisor), Dr. Sherif Saad and Dr. Mohammed Fazel Baki contributed in revising the publication.In all cases, the key ideas, primary contributions, experimental designs, data analysis and interpretation, were performed by the author, and the contribution of co-authors was primarily through the proofreading of the published manuscripts.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

II. Previous Publication

This thesis includes one original paper that has been previously submitted for publication in peer reviewed journals as follows:

Section	Publication title/full citation	Publication status
1.2 , 4.1,4.2, 4.3, 4.4,4.5, 4.6, 5.5,6.2	Priyanshi Shah and Ziad Kobti. Multimodal fake news detection using a Cultural Algorithm situational and normative knowledge *IEEE Congress on Evolutionary Computation (CEC)* WCCI, 2020 conference.	Accepted

I certify that I have obtained a written permission from the copyright owner(s)

to include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as a graduate student at the University of Windsor.

III. General

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution

Abstract

The proliferation of fake news on social media sites is a serious problem with documented negative impacts on individuals and organizations. This makes detection of fake news an extremely important challenge. A fake news item is usually created by manipulating photos, text or videos that indicate the need for multimodal detection. Researchers are building detection algorithms with the aim of high accuracy as this will have a massive impact on the prevailing social and political issues. A shortcoming of existing strategies for identifying fake news is their inability to learn a feature representation of multimodal (textual+visual) information. In this thesis research, we present a novel approach using a Cultural Algorithm with situational and normative knowledge to detect fake news using both text and images. The proposed model's principal innovation is to use the power of natural language processing like sentiment analysis, segmentation process for feature extraction, and optimizing it with a Cultural algorithm. Then the representations from both modalities are fused, which is finally used for classification. An extensive set of experiments is carried out on real-world multimedia datasets collected from Weibo and Twitter. The proposed method outperforms the state-of-the-art methods for identifying fake news.

Dedication

I would like to dedicate this thesis to my parents, Virendra Shah and Sonal Shah without their support I would not have made it this far, my loving brother Jainam Shah, my family and friends.

Acknowledgements

I would like to take this opportunity to express my heartfelt gratitude to all those who not only helped me complete my research thesis but also made my experience at the University of Windsor unforgettable. Firstly, I would like to express my sincere gratitude to my advisor, Dr. Ziad Kobti, for his continued encouragement, inspiration, and tremendous expertise. He gave me the opportunity to choose my study area and discover the Online Social Network domain. His encouragement has supported me in the studies and writing of this thesis. Besides my advisor, I would like to thank my thesis committee, Dr. Mohammed Fazle Baki and Dr. Sherif Saad for their insightful comments and inspiration for the completion of this thesis.

I sincerely thank my parents for their love and unconditional care. They have always supported me through these years and motivate me. They have always believed more in me than myself. Also, special thanks to Jainam for always being there and motivating me.

I would also like to thank my lab mate Kalyani Selvarajah for helping me with her insights and her skills in writing papers. A special thanks to Shikhar Sanghvi for intriguing ideas and elaborate discussions.

Finally, life at Windsor University would not have been what it was without my friends' continued support. I would like to offer my sincere thanks to all my friends who have been vital support throughout my work.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1   Background

The rapid development in technology, the emergence of the internet and the digitization of media have reduced the difficulties in terms of geographic reach to the news source and quick information distribution. In addition to conventional news media, the online social network acts as the primary platform for sharing news and user opinions. The causes for this shift in user habits on social media platforms is because the distribution of content on social media is often more frequent and less costly. The news with images and videos can provide a stronger story-line and attract more attention from readers as compared to traditional text reporting. Owing to these advantages, many people tend to search for news from social media rather than classical news sources such as television or newspaper. The role of social networks in real-world events such as support, situational awareness and aid during disasters, crises, and emergencies is recorded in many studies [47]. But the creation and dissemination of fake information are harming the success of technical advances and digitization of news media.

Fake news is no new reality. The origin of it existed long ago in society, but the harm it caused on civilization made it a significant problem for the scientific community to address. This word has now become a jargon, but as compared to the previous studies [11], the way it is described is a little different. Previously, any type of distinct content like satires, hoaxes, news propaganda, and click baits were called fake news. However, a recent study describes fake news "to be news articles that are intentionally and verifiable false, and could mislead readers"[2]. Besides, such content is written to deceive anyone. Fake news which usually includes misinformation or even fake images, often takes advantage of this to mislead the users, damage a community or individual, create chaos, and gain financially or politically.

Recent evidence confirms that multimodal disinformation, i.e. disinformation that involves text and image-based content is more reliable than textual content alone. An example of the simplicity of image-based fake news generation is shown in Figure 1. Using the online service BreakYourOwnNews [1], a breaking news fake was produced that transported this misleading message.

Figure 1.1: An example of fake news which claims that the 2019 corona virus outbreak in China could be cured by cocaine consumption[13]

---

[1]https://breakyourownnews.com/

Unfortunately, fake news, which usually includes misinformation or even fake images, often takes advantage of this to mislead the users to damage a community or individual, create chaos, and gain financially or politically. For example, within the final three months of the 2016 U.S presidential election, the fake news generated to favor either of the two candidates was believed by many people and was shared by more than 37 million times on Facebook and Twitter [2].

The extensive spread of fake news can have a serious negative impact on individuals
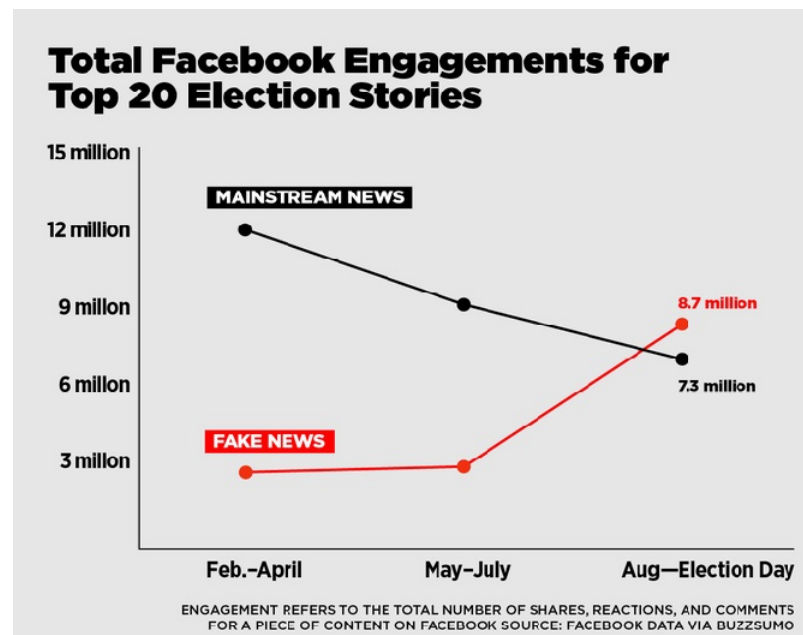
Figure 1.2: Buzz feed analysis shows how Viral Fake Election News Stories outperformed Real News on Facebook

and society like:

- Can break the authenticity balance of the news ecosystem

- It deliberately persuades consumers to accept biased or false beliefs.

- Affects the way in which people interpret and react to actual news.

This makes the task of detecting fake news a crucial one. Automated detection is in great need for reducing the serious negative shortcoming created by the fake news.

However, accurate fake news detection is still challenging due to the dynamic nature of the social media, the complexity and diversity of online communication data[7]. Designing reliable, automated, and applicable approaches to fake news detection online is significant.

1.2   Problem Definition

Detecting fake news on social media poses several challenging research problems. First, fake news is intentionally written to confuse viewers, which makes it nontrivial to identify simply based on news content. Thus, textual features only is not appropriate for the identification of fake news. Second, certain auxiliary information such as knowledge base and user social engagements, must also be added to improve detection [42]. But, the exploitation of this auxiliary information in fact leads to contributes to another important challenge of the data quality. Although information in different modalities can provide clues for fake detection, how to extract prominent features from each modality and effectively fuse them together is challenging.

Formally, we define our problem as follow:
Given a set of m social news articles which includes text and image information, we can represent the data as a collection of text-image tuples with $A = (A_i^T, A_i^I)_i{}^m$.

In the problem of fake news detection we want to determine whether the news articles in A are fake news or not.We can represent label set as $Y=\{1 , 0\}$ where 1 denotes the fake news and 0 present the real news.

Meanwhile, based on the news articles, e.g. $A = (A_i^T, A_i^I)_i{}^m$, a set of features can be derived from both the available text and image information in the article, which

can be interpreted as X_i^T and X_i^I respectively. The objective of fake news detection problem is to create a model $F:\{X_i^T, X_i^I\} \; \mathcal{E} \; X \rightarrow Y$ to infer the potential labels of the news articles in A.

The task of fake news detection model is to predict whether the news article A is a fake news piece or not, i.e., $F(A) \rightarrow 0, 1$ such that,

$$f(A) = \begin{cases} 1, & \text{if } a \text{ is a piece of fake news} \\ 0, & \text{otherwise .} \end{cases}$$

1.3 Motivation

Digital social networks have allowed news articles to switch from conventional text-only news to picture and video content which can give a better narrative experience and allow more readers to engage. Recent fake news articles take advantage of this very switch to visually-supported reporting.Fake news articles may also contain photos that are misrepresented, meaningless, and distorted to mislead readers. This suggests the need for fake news identification to provide multimodal network.

The motivation to leverage multimodal information in fake news detection model is as follows:

- Various modalities show different aspects of a news.

- Information originating from different modalities supplement each other in detecting the credibility of news.

- Different sources exploit different modalities based on their skills (e.g., certain people have skill in exploiting fake news by manipulating images and others may be good in manipulating modalities such as text, audio and videos).

1.4 Thesis Statement

The majority of research is done using unimodal information but information can be derived from multiple modalities so it is necessary to consider both text and visual information for better performance of fake news detection. Though existing multimodal systems perform well in detecting fake news, the classifiers have always been trained in tandem with another classifier. This increases training and model size overhead, at times can also hinder the generalizability of the systems due to a lack of data for the secondary task.

We hypothesize that fake news detection is heavily dependent on the sub task and in absence of sub task training, the performance of fake news detection degrades by 10% on an average so we do not consider any other sub-task in the detection process in our model. Also, we believe that an important step in developing predictive models is determining the best features to be used for building the models so by using the evolutionary approach like a Cultural algorithm that uses a wider variety of cultural knowledge for finding optimal features , we expect to see reduce in the cost and also make it more efficient.

To solve such issues, we design a model of fake news detection using a cultural algorithm on two different modalities and distinguish a piece of news content into real and fake without considering any secondary tasks.

Note that we can define fake news detection as a binary classification problem for the following reason: fake news is essentially a distortion bias on writer-manipulated content. According to prior work on media bias theory [17], distortion bias is typically described as an issue of binary classification [42].

1.5   Thesis Contribution

The main contribution of this research can be summarized as follows:

- We design a multimodal framework for fake news detection using a Cultural algorithm. The proposed model is aimed to detect whether particularly given news is real or fake. It takes no further sub-task into account in the detection process.

- Paper's principal innovation is to use the power of natural language processing like sentiment analysis, segmentation process and optimizing it with an evolutionary Cultural algorithm that uses a wider variety of cultural knowledge. Then, the representations from both modalities are concatenated together to produce the desired vector of news, which is finally used for classification.

- Our proposed architecture provides a general model for identifying fake news. The designed multi-modal extractor can be replaced by various versions designed for extraction of features.

- We experimentally demonstrate that the proposed model can effectively classify fake news and outperform existing state-of-the-art models on two large-scale real-world datasets.

1.6   Thesis Organization

The rest of the thesis / research work is structured as follows.

In chapter II, we introduce background knowledge of Fake News Detection and its various data mining perspective. We also discuss its characterization, feature extraction, and model construction.

In chapter III, we discuss the related work/literature review in the field of fake news identification using different techniques, with a focus on experiments utilizing multimodal evidence.

In chapter IV, we explain the detailed description of the proposed method and introduced our architecture model for fake news detection.

In chapter V, we present the experimental setup and statistical analysis of different types of data set used for the experiment with its assumptions.

In chapter VI, we compare our work with other state-of-the-art methods and a detailed analysis of the observations is done.

In chapter VII, concludes the research, explaining the insights received during the work and setting up a wide range of opportunities for the future work.

Chapter 2

Fake News Detection

2.1 Fake News Characterization

To work on identifying fake news, it is important to understand what is fake news and how it is characterized. For constructing detection models, it is necessary to start by characterization, indeed, before attempting to detect them, it is necessary to understand what is fake news.

Online fake news today tends to be manipulative and diverse in topics, styles, and platforms. And creating a widely accepted definition for "fake news" is not easy. Shu k. et al[42] describes fake news as: "news articles that are intentionally and verifiable false and to mislead readers". Fake news, according to Wikipedia is: "a type of false journalism or propaganda that consists of deliberate misinformation or hoaxes spread through traditional print and broadcast news media or online social media."

According to Kai Shu et al.[42] the definition of fake news is made up of two parts: authenticity and intent. Authenticity includes fake news content false facts that can be checked as such, but a conspiracy theory is not included in fake news because in

most cases it is difficult to prove true or false. The second part, intent, means the false information was written with the intention to deceive the reader.
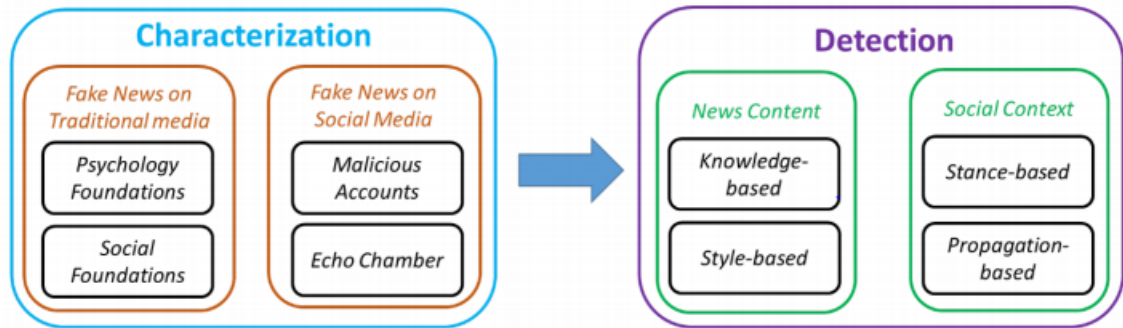
Figure 2.1: Fake news on social media: from characterization to detection [42]

2.1.1 Traditional News Media Fake News

Through time, the media ecology of fake news has evolved from newsprint to television to online news and social media. "Traditional fake news" is considered as the issue of fake news before social media had major effects on its creation and dissemination. Several psychological and social science foundations that define the impact of fake news on the ecosystem of person and social knowledge. Obviously, humans aren't very good at discriminating between real and false news. There are many theories in psychology and perception that can explain this phenomenon and the power of false information.

Traditional fake news targets consumers primarily by taking advantage of their vulnerabilities. Two big factors that make consumers vulnerable to false information are as follow:

- **Naïve Realism**: Consumers prefer to assume that their interpretation of truth is the only correct view, whereas those who disagree are considered uninformed, irrational, or biased [51].

- **Confirmation Bias**: Customers tend to obtain details confirming current views. Because of these cognitive biases that are inherent in human nature, consumers often perceive fake news as real [33] .

Considering the entire environment of news consumption we can also identify some of the social factors that lead to the dissemination of false news. As defined in the Theory of Social Identity [46] and the Theory of Normative Influence [5], this desire for social recognition and approval is central to the identity and self-esteem of an individual, making users likely to select socially secure options when accessing and disseminating news content, even if the news being shared is fake news.

2.1.2 Online Social Media Fake news

We will discuss some unique features of the fake news on social media in this subsection. In particular, we'll highlight the key features of fake news that social media enables. Notice that typical fake news features related to traditional often extend to social media.

While many social media users are legitimate, social media users can even be malicious, and in some cases, they are not even actual human beings. The low cost of social media account formation often facilitates malicious user accounts such as online bots, cyborg users and trolls. Social Bots are accounts on social media that are powered by a computer algorithm. Social bots offer a false impression that information is highly popular and endorsed by many, allowing the echo chamber effect to spread fake news [41].

For example, Facebook bots may become malicious entities intentionally designed to do harm, such as exploiting and disseminating fake news on social media. Studies show that social bots manipulated political debates on a wide scale in the 2016 U.S. presidential election [6], and that about 19 million bot accounts tweeted in favor of either Trump or Clinton in the week leading up to the election day.

2.2   Feature Extraction

Fake news identification on conventional news media depends mainly on news information like textual and visual content, while in social media, additional social information can be used to help detect fake news. Therefore, the specifics of how to derive and represent useful features from news material and social context will be discussed.
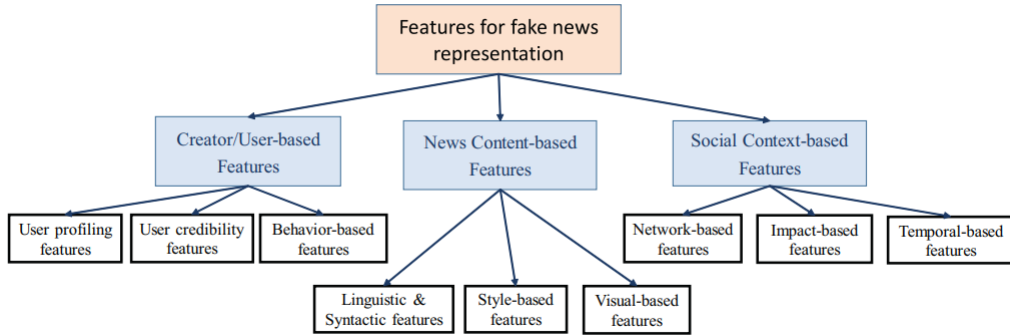
Figure 2.2: Feature representation for fake news

2.2.1   Textual content-based

Starting with the news content, it can be seen that it is composed of four main raw components:

- **Source**: Where the news comes from, who wrote it, is accurate or not this source.

- **Headline**: Brief description of the quality of news that is seeking to attract readers.

- **Body Text**: The actual substance of the news article.

- **Image / Video**: Visual content such as pictures, videos or audio is usually added to textual content.

The prevalent way to identify and detect fake news depends on the news articles content. The material of a news report can usually be classified into both textual and visual modalities. Among those modalities, with the help of textual modality, a significant portion of the news information is populated.

As mentioned earlier, false news is used to manipulate the audience, and for this, they also use a particular language to attract readers. On the other hand, non-fake news would usually move to another language list, being more formal. Such features may usually be divided into language features based on the attributes or structure-related language features.

Attribute-based language features

Attribute-based language features that characterize content style can be grouped into ten parallel dimensions: quantity, ambiguity, uncertainty, subjectivity, non-immediacy, sentiment, variety, informality, specificity, and readability [57]. Although attribute-based language features may be highly significant, explainable, and pre-

dictable, they are often ineffective compared to structure-based features in quantifying the style of deception. In particular, attributed features often require additional quantification or computational speeds, which can take time and attach greater importance to proper feature assessment and filtering for deception detection.

Structure-based language features

Structure-based language features define content style from (at least) four levels of language: (i) lexicon, (ii) syntax, (iii) semantic, and (iv) discourse. Also known as technique-oriented features are structure-related features since their quantification is largely based on advanced Natural Language Processing (NLP) techniques.

In particular, the key task at the level of the lexicon is to determine the frequency statistics of a letter(s), word(s), etc., which can be carried out accurately using n-gram models. At the level of syntax, shallow syntax tasks are performed by Part-Of-Speech (POS)-taggers which facilitate tagging and analysis of POS. Probabilistic Context-Free Grammars (PCFG) performs deep syntax level tasks with parse trees that allow for Context-Free Grammars (CFG) analysis. Linguistic Inquiry and Word Count (LIWC) is also used on the semantic level to provide semantic classes for semantic features.

2.2.2   Visual content-based

The advancement of multimedia technology facilitates the evolution of self-media news from text-based posts to picture or video multimedia posts, providing better storytelling, and attracting more reader's attention. Unfortunately, this advantage is also taken by fake news which usually contains misrepresented or even tampered

images to attract and mislead readers for rapid dissemination. As a result, visual content has become an important part of fake news that can not be neglected [37].

For example: In figure 2.3, A tampered image where Putin is spliced on the middle seat at G-20 to show that he is in the middle of an intense discussion between other world leaders.

Figure 2.3: Tampered images
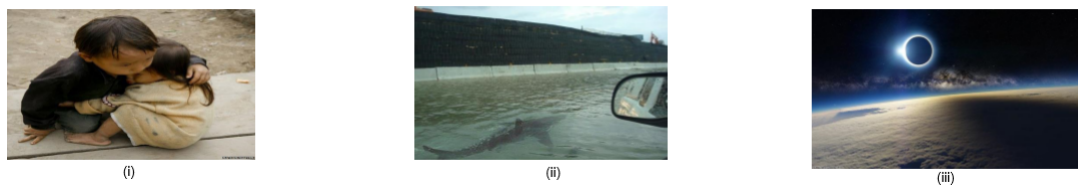
Figure 2.4: Misleading images

The above example reflects that pairs of visual and textual information can give better insights into fake news detection. (i) **Tweet**: "two Vietnamese siblings as being captured during the Nepal 2015 earthquakes". Whereas it is a real photo from the past re-posted as being associated with a current event.

(ii) **Tweet**: "the presence of sharks during Hurricane Sandy 2012". Whereas deep analysis on the image concludes that it was spliced to show fake sharks in the image.

(iii) **Tweet**: "picture of Solar Eclipse captured on March 20, 2015". However, the

image is an artwork done so beautifully that it is hard to distinguish from reality.

We can narrowly classify fake images into two groups: tampered images and misleading images. Tampered images mean fake news images that have been modified digitally as shown in the above figure, in our common sense, the same as fake images. Misleading images refer to fake news images that have not experienced any manipulation, but as described in a figure the content is misleading. Generally, these misleading photos come from artworks or obsolete pictures that are released at an early event [37].

Fake images in the real world may have significantly different characteristics from real-news images at both physical and semantic levels. Fake-news images can be of low quality at the physical level, which can be clearly shown in the frequency domain. In view of the fact that re-compressed and tampered images frequently display in the frequency domain periodicity which can be easily characterized by CNN which is capable of capturing spatial structural features. Fake images also display some distinct features in the pixel domain (also known as spatial domain) at the semantic level.

2.2.3 Social context-based

Multiple factors, such as user aspect, post aspect and community aspect, maybe taken into consideration when posting news on social media. For example, the actions of individual users can be analyzed and their metadata used to figure out whether a user is at risk of trusting or sharing false information. For example, this metadata can be its center of focus, its number of followers, or something that relates to it.

Social context-based features are designed to model the online news delivery trend,

and online user engagement. And the following three styles can be summarised: network-based features, distribution-based features, and temporal-based features.

- **Network-based features**: Network-based research intended to concentrate on a community of related online users in terms of various viewpoints, such as interest, topic, relations and behaviors. And network-dependent functions are selected and extracted based on a particular network and It can be used to research the specific characteristics of certain networks and the similarities and differences between different online accounts.

- **Distribution-based technologies** can help capture the distinct diffusion pattern of online news. Typically a the propagation tree may be built to promote the classification of a piece of news distribution [10]. And the features associated with the propagation tree include the degree of the root in a propagation tree, the maximum number of subtrees, the median / average degree and depth of the tree, and so on. Besides, some other features such as the number of retweets/reposts for the original tweet/post, the fraction of tweets/posts retweeted for an online account, the in-degree / out-degree of an online tweet/post [54].

- **Temporary-based features**: Temporary-based features can be used to define the posting activity of the online news producer in a time-series manner. They are a good attribute for detecting suspicious posting activities and can be used to show the false level of online news. Temporary-based features widely used include the distance between two posts, the frequency of posting, replying and commenting on a certain account, the day on which the original information is

posted/shared / commented, and the day on which the message was written.

2.3 Model Construction

In this section, we will discuss the details for several existing approaches of the model construction process. We categorise existing methods specifically based on their main sources of input such as: News Content Models and Social Context Models.

2.3.1 News content based

In this subsection, we focus on news content models which depend primarily on features of news content and established factual sources to identify fake news. Existing approaches can clearly be classified as knowledge-based and style-based approaches.

Knowledge-based

Since fake news attempts to spread false statements in news material, the best way to identify it is to test the truthfulness of major statements in a news article to assess the truthfulness of the news. Approaches based on knowledge seek to use external sources to verify reported statements in news reporting. Fact checking, initially developed in journalism, seeks to test the validity of news by matching information derived from news information to be checked (e.g., its claims or statements) with known facts (e.g., true information) Increasing focus has been paid to fact-checking, and several attempts have been made to build a practical automated fact-checking program. Broadly speaking, manual fact-checking can be divided into (I) expert-based and (II) crowd-sourced fact-checking (III) computational oriented.

	Topics Covered	Content Analyzed	Assessment Labels
PolitiFact	American politics	Statements	True; Mostly true; Half true; Mostly false; False; Pants on fire
The Washington Post Fact Checker	American politics	Statements and claims	One Pinocchio; Two Pinocchio; Three Pinocchio; Four Pinocchio; The Geppetto checkmark; An upside-down Pinocchio; Verdict pending
Fact Check	American politics	TV ads, debates, speeches, interviews and news	True; No evidence; False
Snopes	Politics and other social and topical issues	News articles and videos	True; Mostly true; Mixture; Mostly false; False; Unproven; Outdated; Miscaptioned; Correct attribution; Misattributed; Scam; Legend
TruthOrFiction	Politics, religion, nature, aviation, food, medical, etc	Email rumors	Truth; Fiction; etc
FullFact	Economy, health, education, crime, immigration, law	Articles	Ambiguity (no clear labels)
HoaxSlayer	Ambiguity	Articles and messages	Hoaxes, scams, malware, bogus warning, fake news, misleading, true, humour, spams, etc.

Table 2.1: A Comparison among Expert-based Fact-checking Websites

- **Expert-based fact-checking** : Expert-based fact-checking is based on domain-experts (i.e. fact-checkers) to verify the accuracy of the given news. Expert-based fact-checking is mostly carried out by a small group of highly reliable fact-checkers, is easy to handle, and contributes to reasonably accurate outcomes but is expensive and can not scale with the rise in the amount of news content to be reviewed.

Recently, several blogs have arisen to better serve the public by allowing for expert fact-checking. In Table 2.1 we mention and provide information about the well-known websites [57]. Some websites provide additional information, for example, PolitiFact offers "the PolitiFact score card", which presents data on the distribution of validity of all claims relating to a particular subject (see an example on Donald Trump, 45 [th] US President on below figure)

Although the website does not include specific definitions for these categories, its knowledge can theoretically be used for comparative studies of fake news as ground-truth. However, expert fact-checking is a time-consuming and intel-

Donald Trump's file

Republican from New York

Donald Trump was elected the 45th president of the United States on Nov. 8, 2016. He has been a real estate developer, entrepreneur and host of the NBC reality show, "The Apprentice." Trump's statements were awarded PolitiFact's 2015 Lie of the Year. Born and raised in New York City, Trump is married to Melania Trump, a former model from Slovenia. Trump has five children and eight grandchildren. Three of his children, Donald Jr., Ivanka, and Eric, serve as executive vice presidents of the Trump Organization.

The PolitiFact scorecard

True	25 (5%)
Mostly True	61 (11%)
Half True	83 (15%)
Mostly False	118 (22%)
False	173 (32%)
Pants on Fire	78 (14%)

Article Stats

Tag Distribution

- True / False
- Fallacy
- Descriptive
- Compliment

Most-Used Tags

Insightful	18
Well Researched	4
True	3
Overly Simplistic	2
Biased Wording	1

(a) (Expert-based) PolitiFact: the PolitiFact scorecard

(b) (Crowd-sourced) Fiskkit: the tag distribution

Figure 2.5: An illustration of fack-checking website

lectually challenging process which limits the potential for high efficiency and scalability.

- **Crowd-sourced fact-checking**

  Crowd-sourced fact-checking is based on a broad community of ordinary individuals who serve as fact-checkers (i.e., the collective intelligence). Compared with expert-based fact-checking, crowd-sourced fact-checking is more difficult to handle, less reliable and precise due to the fact-checker's political bias and contradictory annotations and has greater (although insufficient) scalability.Crowd sourcing-oriented fact-checking takes advantage of the "crowd wisdom" to encourage average people to annotate news content; these annotations are then aggregated to create an overall evaluation of the news veracity.

  Thus, in crowd-sourced fact-checking, one also needs (i) filtering non-credible users, and (ii) resolving contradictory fact-checking findings, both criteria becoming more important as the number of fact-checkers increases. Nevertheless, crowd-sourcing sites also encourage fact-checkers to provide more comprehensive input (e.g.their thoughts or positions), which can be discussed further in

fake news studies.

Like expert-based fact-checking, there is still early production of crowd-sourced fact-checking websites. An example is Fiskkit[1], where users can upload posts, provide sentence ratings inside posts and pick tags that better describe them.The sources given for the articles help (i) differentiate the content types (e.g., news versus non-news) and (ii) assess their credibility [57].The tags categorised into multiple dimensions allow one to analyse trends across fake and unfake news articles (for example, see Figure 2.5(b)).

Although there are not many crowd-sourced fact-checking websites, we believe that more crowd-sourced platforms or resources will emerge as major websites and social media have recognised the value of recognising fake news (e.g., Google, Facebook, Twitter, and Sina Weibo)

- **Computational oriented**

Computer-oriented fact-checking aims at offering an automated scalable framework for the detection of true and false statements. Previous computational-oriented fact-checking approaches seek to address two major issues: (i) recognising credible statements and (ii) distinguishing the veracity of statements of fact.To distinguish legitimate claims, factual arguments are extracted in news material, which express key statements and opinions, thus facilitating the subsequent fact-checking process [21]. To assess the truthfulness of a particular argument, fact-checking for specific statements relies primarily on outside resources. The open platform and the organised information network contain two common external outlets. Open web sources are used in terms of accuracy and frequency as standards that can be contrasted with particular statements.

---

[1]https://fiskkit.com/

As a structured network topology, knowledge graphs are extracted from the related open data, such as DBpedia and Google Relation Extraction Corpus. The purpose of fact-checking using a knowledge graph is to check if statements in news content can be inferred in the knowledge graph from actual facts [57].

Style-based

Fake news publishers also have malicious intent to disseminate and deceptive information and manipulate vast user populations, requiring different types of writing required to appeal to and convince a wide spectrum of customers that is not found in real news papers. Style-based strategies aim to identify fake news by identifying the news content manipulators in the writing format. There are primarily two traditional style-based approach categories: Deception-oriented and Objectivity-oriented.

- **Deception-oriented stylometric methods** catch the tricky statements or assertions from the substance of news. The inspiration for deception detection comes from forensic psychology (i.e., the Undeutsch Hypothesis)and numerous forensic techniques have been developed, including Criteria-based content analysis[49] and Scientific-based content analysis [30]. More recently, sophisticated models of natural language analysis are implemented from the following viewpoints to identify stages of deception: Deep syntax and rhetorical structure.

  Deep syntax models have been implemented using probabilistic context free grammars (PCFG) to turn sentences into rules defining the syntax structure.Different rules for detection of deception can be established based on the PCFG, such as unlexicalized/lexicalized development rules and grandparent rules[14]. The theory of rhetorical structure can be used to identify discrepancies between tricky

and true sentences.Deep network models such as convolutions neural networks (CNN) were also applied to assess the veracity of false news.

- **Objectivity-oriented** is a approaches capture style cues that could suggest a reduced objectivity of news content and thus the potential for misleading audiences, such as hyper-partisan types and yellow journalism. Hyper-partisan types reflect disproportionate actions against a single political group, which also coincides with a strong desire to produce false news.

  Hyper-partisan articles can be identified using linguistic-based features. Yellow-journalism describes those articles that do not contain well-researched news, but rely instead on eye-catching headlines (i.e., click bait) with a tendency to exaggerate, sensationalize, scaremonger, etc. News titles also summarise the main points of view of the article the author wishes to express, and thus inaccurate and deceptive clickbait titles can serve as a good indicator for identifying fake news articles

2.3.2   Social context based

Social media nature offers additional resources for researchers to supplement and improve the news content models. Social context models involve specific user social commitments in the study, collected from a variety of perspectives this auxiliary knowledge.We may divide current social context modelling methods into two categories: Stance-based, and Propagation-based.

Remember that very few current approaches to fake news identification have used model social context. Therefore, we also discuss related methods for detecting ru-

mours using social media, which have potential application for detecting fake news detection.

- **Stance-based methods** use the perspectives of users from related post material to infer the veracity of the original news articles. The stance of user posts can be interpreted either explicitly or implicitly. Explicit positions are direct expressions of emotion or belief, such as the reactions displayed in Facebook as "thumbs up" and "thumbs down". Implicit views can be automatically extracted from posts on social media [42].Stance detection is the task of deciding automatically from a post whether the user is in favour of, neutral towards or against any target person, event or idea.

- **Propagation-based** approaches to fake news detection trigger about the interrelationships between related social media posts to predict credibility of news.The basic idea is that a news event's credibility is strongly tied to the legitimacy of related social media articles. For the propagation process both homogeneous and heterogeneous networks of legitimacy can be created.

  Homogenous networks of reputation consist of a single category of entities, such as post or event. Heterogeneous networks of reputation include multiple types of entities, such as messages, sub-events, and events.

Chapter 3

Literature Review

In this section, we briefly discuss the work related to the proposed model. There are several activities involved in identifying false news, such as identifying rumours [24] and spam detection [39].We identify fake news as purposely fabricated news and can be checked as fake. The main challenge of fake news detection activity is how to identify news through features. Features can be derived from tweets, social context, and even attached images. Therefore, we evaluate existing work from the following two categories: single-modality based and multimodal based fake news detection.

3.1 Single Modality based Fake News Detection

3.1.1 Textual content based

Textual features are statistical or semantic features derived from the text content of messages, which were discussed in many fake news identification literature.Most previous research done at the news detection level depended heavily on features of text and user metadata. Textual feature extracts specific writing style [19] [15] [42] and emotional sensation [18] that commonly occurs in fake news content.

Attribute Type	Feature	[55]	[16]	[1]	[40]	[22]	[34]	[43]	[53]	[9]	[8]	[56]	[36]
Quantity	Character count			✓	✓								
	Word count	✓	✓	✓	✓	✓	✓	✓	✓				
	Noun count								✓				
	Verb count	✓	✓			✓		✓	✓				
	Number of noun phrases	✓											
	Sentence count	✓	✓	✓		✓		✓					
	Paragraph count											✓	
	Number of modifiers(e.g., adjectives and adverbs)	✓	✓	✓				✓	✓				
Complexity	Average number of clauses per sentence	✓						✓					
	Average number of words per sentence	✓	✓	✓	✓	✓		✓	✓				
	Average number of characters per sentence	✓	✓	✓	✓	✓		✓					
	Average number of punctuations per sentence	✓	✓	✓				✓					
Uncertainity	Percentage of modal verbs	✓	✓	✓		✓		✓					
	Percentage of centainty verbs	✓	✓	✓		✓		✓					✓
	Percentage of generalizing verbs		✓			✓							
	Percentage of tentative verbs		✓	✓		✓							✓
	Percentage of numbers and quantifiers			✓	✓	✓							✓
	Percentage of question marks			✓	✓								
Subjectivity	Percentage of subjective verbs	✓							✓	✓			
	Percentage of report verbs												✓
	Percentage of factive verbs												✓
	Percentage of imperative verbs												✓
Non-immediacy	Percentage of passive voice	✓	✓			✓							
	Percentage of rhetorical questions												✓
	Self reference: 1 person singular pronouns	✓	✓	✓		✓		✓	✓	✓			✓
	Group reference: 1 person plural pronouns	✓	✓	✓		✓		✓	✓	✓			✓
	Other reference: 2 and 3 person pronouns	✓	✓	✓		✓		✓	✓	✓			✓
	Number of quotations			✓	✓							✓	✓
Sentiment	Percentage of positive words	✓	✓	✓		✓	✓	✓	✓	✓	✓		✓
	Percentage of negative words	✓	✓	✓		✓	✓	✓	✓	✓	✓		✓
	Number of exclamation marks			✓	✓								
	Activation: the dynamics of emotional state	✓	✓										
Diversity	Lexical diversity: unique words or terms(%)	✓	✓	✓	✓	✓		✓	✓				
	Content word diversity: unique content words(%)	✓	✓			✓			✓				
	Redundancy: unique function words(%)	✓	✓	✓	✓	✓			✓				
Informality	Typographical error ratio: misspelled words(%)	✓				✓		✓	✓				
Specificity	Temporal ratio	✓	✓			✓						✓	
	Spatial ratio	✓	✓			✓						✓	
	Sensory ratio	✓	✓			✓	✓	✓				✓	
	Causation terms		✓			✓				✓			
	Exclusive terms		✓			✓							
Readability (e.g., Flesch-Kuncaid and Gunning-Fog index)				✓								✓	

Table 3.1: Attribute-based language features

Potthast et al. [19] demonstrated how to style analysis, network connection, and user reaction would contribute to fake news identification. Shu et al.[42] explained how the writing style of an author affects the views and opinions of people on reading these posts.In many fake news identification studies,emotion is considered a significant predictor and most of them use emotion primarily through user stances or simple statistical emotional features. In [15],they proposed a new dual emotion-based approach to detect fake news where it can learn from the content,user comment and representation of emotions from both publishers and users.Kim et al. [29] proposed a convolutions neural network model to identify fake news which can distinguish dif-

ferent granularity of text features with convolution filters.

3.1.2 Visual content based

Visual features extracted from visual elements (picture and video) have been shown to be a salient indicator of fake news detection [25][42]. However, the value of multimedia content on social media is being verified in very limited studies. Several visual and statistical features have recently been extracted for news prediction [25]. Fake images were recognized using a classification framework based on various user-level and tweet-level features [20]. Marra et al.[31]researched the efficiency of several fake image detectors toward image-to-image conversion using GANs. Though, these models are still hand-crafted and complex to represent visual content.

Methods of image manipulation include: retouching, copy-paste, image splicing and digital fingerprint falsification:

- **Retouching** : Image retouching is based on adding a series of filter to the original image to improve it in accordance with pre-defined goals while retaining similar properties. It involves selecting the area to be retouched, and applying modifiers to produce the most visually appealing image [52]. This technique of deception is widely used in marketing, films and media [26]

  For example, the covers and ads of fashion magazines usually go through some form of retouching to mask imperfections and thereby increase the degree of beauty in the photographs. The initial Figure 3.1b and retouched Figure 3.1a versions of a cover of a magazine called Nitro are shown in Figure 3.1. It displays a picture retouching illustration in which the image of a model has been

digitally altered.

(a) *Manipulated Image*   (b) *Original Image*

Figure 3.1: Manipulated cover of the Nitro magazine

- **Copy-paste** : The copy-paste technique involves copying a region into an image and pasting it from the same image over another region to hide the quality of the image or duplicate regions. It can also require additional post-processing methods, such as scaling, rotation, and application of other filters. These extra procedures make the process of detecting manipulation more costly.

Figure 3.2a gives an example of that technique. The first picture Figure 3.2a displays a historic photograph taken on Iran in 2008 and it portrays the successful launch of four missiles as reported by the news agency of Iran (Sepah News). The initial photo posted later reveals that there were only three missiles fired in Figure 3.2b. The updated picture also demonstrates the use of post-processing techniques on the smoke that the missile expels to mask the manipulation.

(a) Manipulated Image (b) Original Image

Figure 3.2: Manipulated photo of the Iranian missile launch

- **Image splicing** : Photo splicing is the act of copying and pasting the area of a given photo into a different image. This technique is commonly used in photo-montages where it combines two or more images to offer the impression of being one image. Image splicing detection is a difficult task compared to previous manipulation techniques since it is not possible to search for duplicate regions since the manipulated region comes from a different picture.

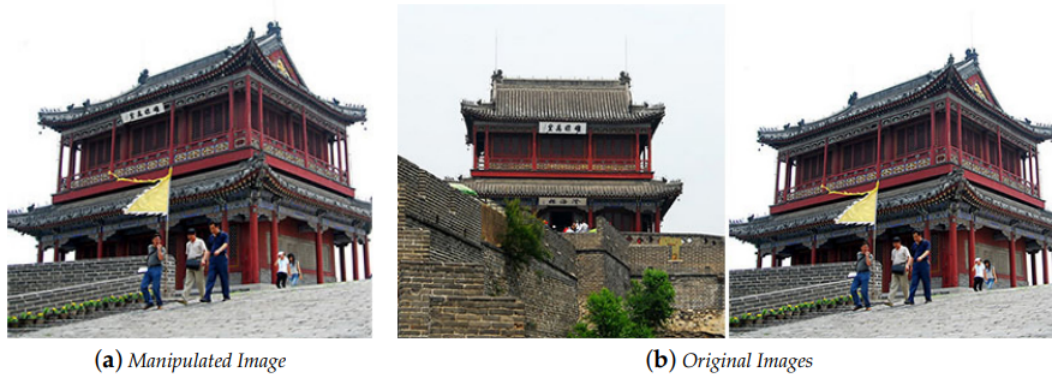(a) Manipulated Image (b) Original Images

Figure 3.3: Manipulated photo of temple

Figure 3.3 provides an example of the technique of splicing, and the two original images used in this process. The temple sign was cut from the first original image, Figure 3.3b, and put in the second original image, Figure 3.3b, to create the end result Figure 3.3a.

- **Digital Fingerprint Forgery** : These techniques do not include altering the image itself, but altering the relevant information (digital fingerprint) that comes from the sensor that captured the image. During the image acquisition process many sources of imperfections and noise are added. Also when a uniform and fully lighted photograph is taken, slight differences in intensity can be found between the individual pixels.

This occurs primarily for two reasons; firstly, there are random components such as readout noise or shot noise [4] and secondly due to pattern noise, which is a deterministic component of the sensor and stays approximately the same if many images are taken from the same scene. This pattern is useful for detecting an image's source of origin, since every system would have a different noise pattern [27].

3.1.3 Social context based

In addition to features directly related to the news articles content, auxiliary information can be extracted from user-driven social engagements of news on social media. In [31] , they proposed a novel approach to detect fake news based on news content using a knowledge graph. Ke Wu et al. [35] aims to detect propagation patterns and sentiments using graph-kernel based method. However, social context features are noisy, unstructured and labor-intensive to collect [44].

3.2   Multi-modal based Fake News Detection

Even though all of the above mentioned single-modality approaches have been able to deliver promising results, the unstructured nature of social networking data is always a difficulty in the extraction of knowledge. And so, the researchers began experimenting with features derived from different modalities (i.e. text and image) to address this constraint.

Deep neural networks have been successfully applied to various tasks to learn feature representations from multiple aspects. Wang et.al [50] created an end-to-end framework based on Adversarial Neural Network, which can derive event invariant features and can detect fake news on the newly arrived events. Their model have two components: for a textual part, it took word vector embedding as input and created text representation using text-CNN [29]. Image representation from VGG-19 platform pre-trained on Image Net [44] was extracted. Both components were fused together using a fully connected neural network classifier into event discriminator and fake news detector. Their proposed architecture is shown below.
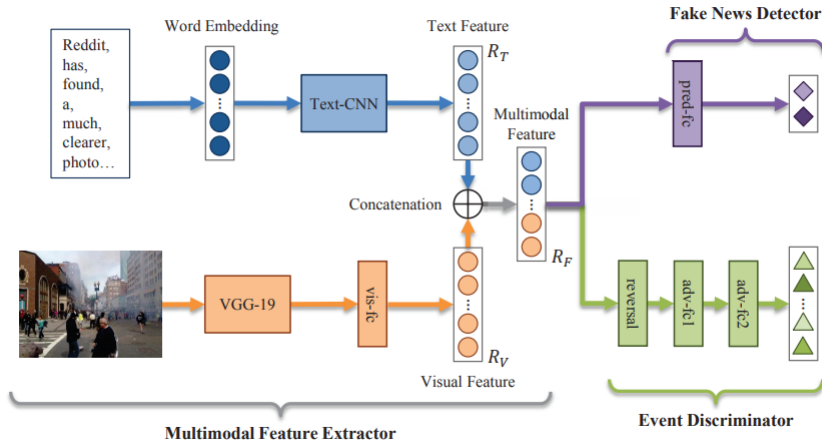

Figure 3.4: Event Adversarial Neural Network Architecture (EANN)

Shivangi et al. [45], introduced a model named SpotFake-A multimodal Framework for Fake News Detection. They used a language model like BERT to learn textual features and VGG-19 platform pre-trained on Image Net [44] to learn image features. Even though multi-modal based models do well in identifying fake news, consideration of secondary task with fake news detection problem decreases the efficiency performance.
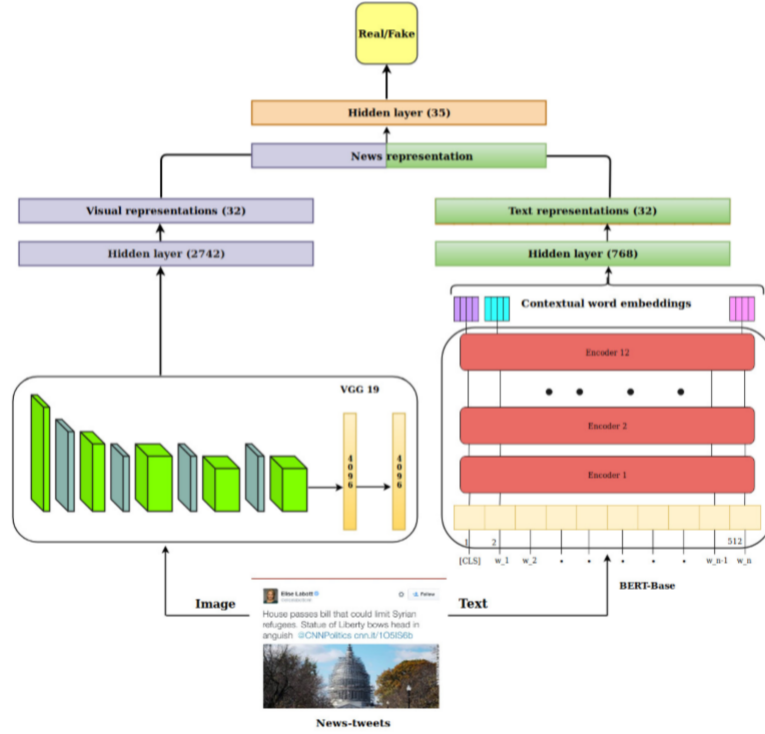
Figure 3.5: SpotFake-A multimodal framework

Author Khattar et al.[28] built a model named MVAE: Multimodal Variational Auto-encoder for Fake News Detection. It uses a bimodal vector autoencoder combined with a conditional classifier for fake news identification tasks. The platform comprises of three main components, an encoder, a decoder, and a fake news detector element. The variable autoencoder is capable of learning probabilistic latent

variable models by maximizing the marginal likelihood of observed data. They used bi-directional LSTMs and VGG-19 to extract text and image representation respectively.
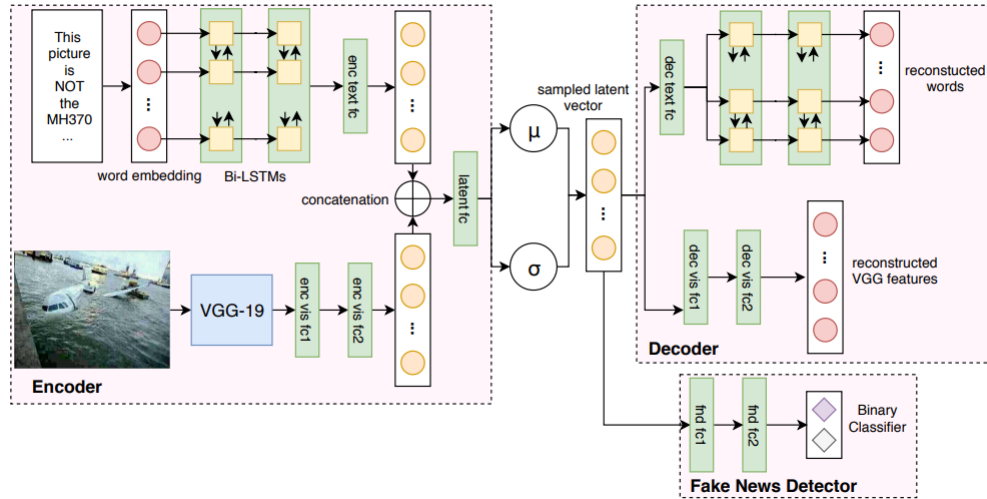
Figure 3.6: Multimodal Variational Auto-encoder (MVAE)

Chapter 4

Proposed Methodology

In this chapter, we give a brief overview and provide an elaborate description of the three components of the proposed model: the textual feature extractor , Visual feature extractor and fake news detector. Subsequent subsections describe importance of these three components to detect fake news.

4.1 Proposed Architecture

Here we introduce our proposed framework of multimodal fake news detection. An overview of our method can be seen in figure 4.1 .The basic idea behind our work is to detect fake news from both the modalities of given tweets independently without considering any other sub-task.

4.2 Proposed Model Overview

Proposed model's principal innovation is to use the power of natural language processing like sentiment analysis, segmentation process for feature extraction and optimizing
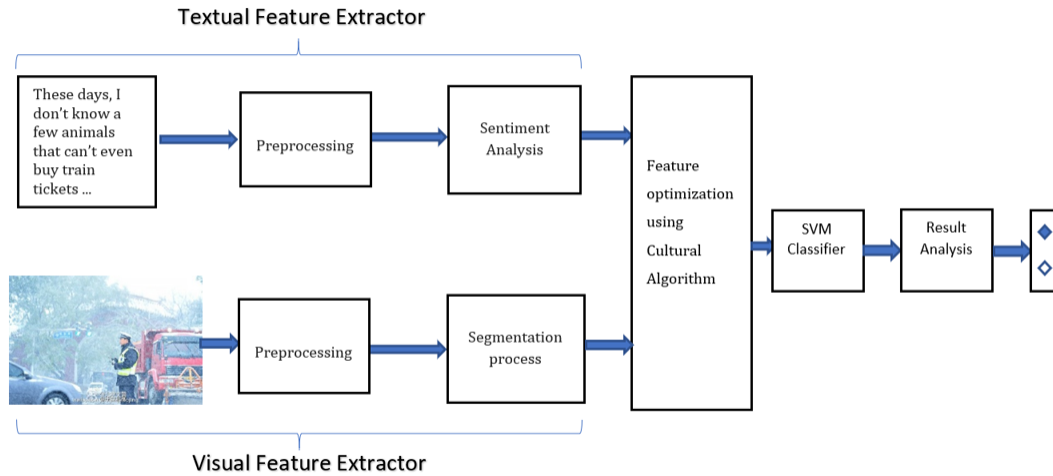
Figure 4.1: Architecture of proposed model

it with a Cultural algorithm using normative and situational knowledge. Then, the representations from both modalities are fused together, which is finally used for classification.The overall architecture of the proposed model is illustrated in figure 4.1 .

We divided our model into three main components.

- The first one is a textual feature extractor that uses Sentiment analysis to extract meaningful content from textual data.

- The second component is Visual feature extractor which extracts visual features from the post using preprocessing techniques and Segmentation process.

- Feature representation from both components were passed through a Cultural algorithm to extract optimum features. The last component is a fake news detector that uses a classifier to detect fake news.

4.3 Textual Feature Extractor

The input to the textual feature extractor are the text of the post and it output is the features learnt from textual modality. The textual feature extractor can be broke down into two steps: Preprocessing and Sentiment process.

- **Preprocessing step** :Text data for applying machine learning algorithms on them need special preprocessing. Numerous techniques are commonly used to transform text data into a prepared-modeling form. The preprocessing steps we discuss below relate to the headlines as well as the news articles.

(1) **Stop Word Removal** : We begin with removing stop words from the available text data. Stops Words (most common words in a language that does not have much context) can be processed and filtered out of the text, because they are more frequent and provide less useful information. Stop words behave more like a linking part of the sentences, for instance conjunctions like "and", "or" and "but", prepositions like "of","in", "from", "to" etc., and articles like "a","an","the".
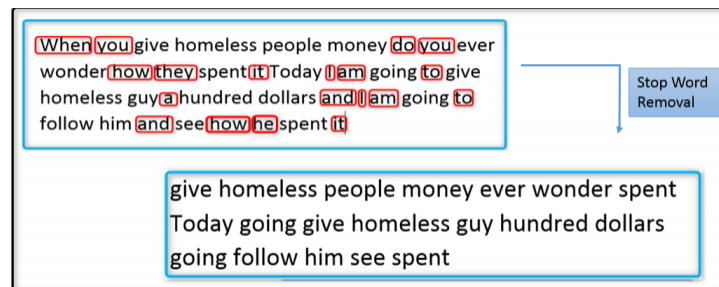
Figure 4.2: Example for Stop Word removal

These less important stop words may take important processing time and hence

eliminating stop words as part of pre-processing data is a key first step in the processing of natural language. Above figure gives an example of stop word removal.

(2) **Punctuation Removal** : Natural language punctuation provides the sentence's grammatical context. Punctuation like a comma may not add much value to the understanding of the sentence. Figure gives an example of the method of eliminating Punctuation.
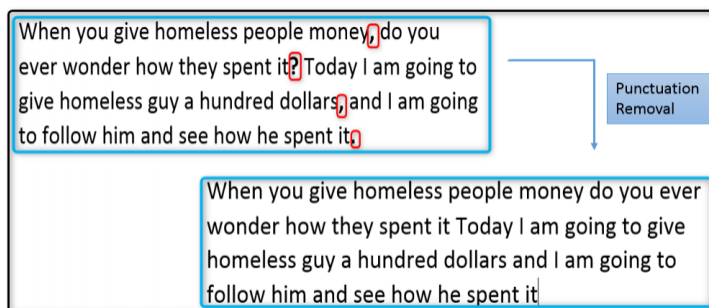
Figure 4.3: Example for punctuation removal

(3) **Stemming** : Stemming is a technique in which prefixes and suffixes are eliminated from a word, ending with the stem.

Figure 4.4: Example Stemming

We may reduce inflectional forms and often derivation related forms of a word to

a specific base form using stemming. Example of stemming technique is shown in figure above.

- **Sentiment process**

In order to extract informative and meaningful features from the textual content, we employ Sentiment analysis. Sentiment analysis is also known as opinion mining which is used to understand the effective meani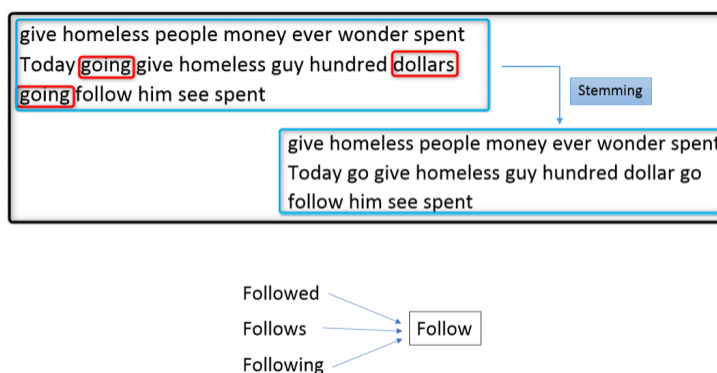ng of sentences and phrases. It assigns classification levels to statements made by the authors; also known as polarity. It could be as basic as binary distinction forms such as positive and negative or sometimes neutral. While sentiment analysis from the document goes beyond polarity, it could also involve evaluating the writer's emotional state as furious, nervous, distressed and excited. Some sentiment dictionaries exist to help in the achievement of this task such as [32] and [23].

In the context of our work, we use the text's sentiment to the negative, positive and neutral polarities of a text message.Sentiment classification can be performed using lexicon-based approach.Under lexicon-based approaches, one may use either a dictionary or a corpus-based approach. Dictionary-based approach would use an existing dictionary, which consists of gathering opinion terms along with their positive(+ve) or negative(-ve) power. The detailed steps of the sentiment analysis of text are summarized in algorithm 1. We believe that the sentiment of writing a news article may serve as a key element in characterizing the news as false or actual. It will also help to further increase accuracy and consider related feelings and sentiments.

Inputs : Text File τ , the sentiment lexicon x
Output: S $= P, N_g, N$ and Strength S, where P: Positive,
        N_g:Negative,N:Neutral.
Initialization: $SumPos$ and $SumNeg = 0$, where
  $SumPos$: collects the polarity of positive tokens $t_{i-s_{mt}}$ in τ
  $SumNeg$: collects the polarity of negative tokens $t_{i-s_{mt}}$ in τ
begin
    foreach $t_i \in \tau$ **do**
        Search for t_i in x
        if $t_i \in Pos - list$ **then**
            SumPos $\leftarrow SumPos + t_{i-smt}$
        end
        else if $t_i \in Neg - list$ **then**
            SumNeg $\leftarrow SumNeg + t_{i-smt}$
        end
    end
    if $SumPos > |SumNeg|$ **then**
        $S_{mt} = P$
        $S =$SumPos/(SumPos+SumNeg)
    end
    else if $SumPos < |SumNeg|$ **then**
        $S_{mt} = N_g$
        $S =$SumNeg/(SumPos+SumNeg)
    end
    else
        $S_{mt} = N$
        $S =$SumPos/(SumPos+SumNeg)
    end
end

Algorithm 1: Sentiment analysis for text

4.4 Visual Feature Extractor

The attached images of the news content are input to the visual feature extractor. In order to extract visual features efficiently, the proposed method first preprocess the image, apply image representation to segmentation process and lastly for optimal feature extraction and minimizing it's cost it using the concept of Cultural algorithm.

- **Preprocessing step**: In our preprocessing step,

  I. We first resize all images to 200*200.

  II. Convert image to gray scale.

  III. Use Otsu's binarization method for automatic image threshold. Thresholding creates binary images from gray-level ones by turning all pixels below some threshold to zero and all pixels above that threshold to one.

Figure 4.5: Example of image preprocess

- **Segmentation process**: The segmentation of images is the method of partitioning a digital image into multiple segments (pixel sets, often known as image objects). The segmentation purpose is to simplify and/or change an image's representation into something that is more relevant and simpler to evaluate.

Here, we use K-means for color based segment which give valuable clues of color, shape and texture of image. Discrete wavelet transform(DWT) analyses an image in different scales and orientations. And additional statistical features like Contrast, correlation, energy , homogeneity are used to extract meaningful content from the images.

Wavelet transform (WT) is an effective tool for extracting features from pictures because its multi-resolution analytical property enables image analysis at multiple scale rates.The DWT is a logical method for complex image disintegration.The discrete wavelet transformation (DWT) can provide special and discriminating representations that can measure crucial and interesting structures such as edges and details for a few coefficients, with good resolution. It is also efficient in computational terms due to the limited amount of data it deals with. The final wavelet coefficients can be used as characteristics directly and can be derived from the wavelet domain directly, defining the anomalies in the image data. Basically, in some wavelet coefficients, the discrete wavelet transformation reduces the association between wave coefficients and provides compaction of the energy.
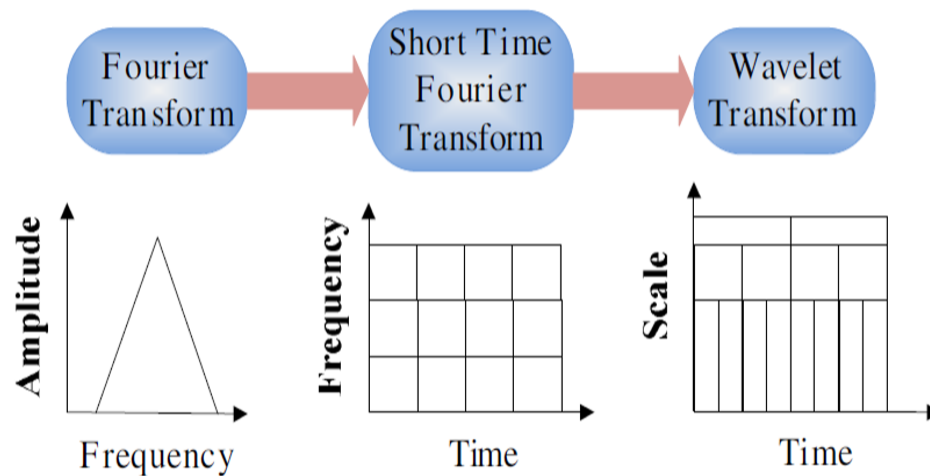
Figure 4.6: The development of DWT signal analysis
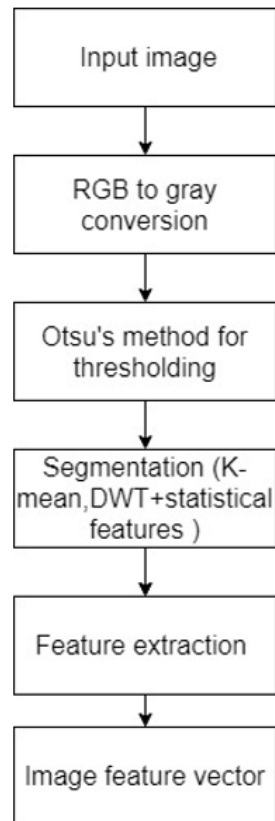
Figure 4.7: Process flow of Visual feature extractor

4.5 Optimization using a Cultural Algorithm

All the feature vector from text and images are passed through the Cultural algorithm for optimal feature selection. Cultural algorithms (CA) are a branch of evolutionary computation where there is a knowledge component that is called the belief space in addition to the population component. It was first introduced by Reynolds [38].Its use as an optimization algorithm to restrict optimization, combinatorial optimization, and continuous function optimization for a wide variety of non-limited domains.

Our motive is using this algorithm to find optimal features from both text and

images at the best cost. The goal of using CA here is find optimum feature subset
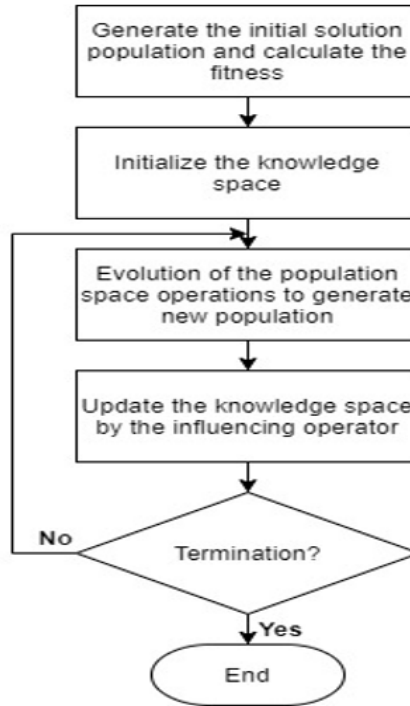
Figure 4.8: framework of Cultural Algorithm

that result in the best outcome for classification. Hence, a feature selection sub component is presented in which cultural algorithm is used as search strategy and SVM classifier is used as objective function. The pseudo code of the proposed model is presented in algorithm 2.

The process of proposed model is as below:

- At first, the initial population is generated based on acceptable points in problem space. Also initial knowledge base is generated. Before starting the evolution process, the initial population is evaluated by fitness function.

- In each iteration t, the next generation is generated by doing some operation on present population.

- The generated population is evaluated by fitness function.

- The belief space is updated based on fitness values of present population.

- The belief space influences the present population.

- Until the stopping criteria is not met, steps 2 to 6 is repeated.

begin
    t=0;
    Initialize Population POP(t);
    Initialize Belief Space BLF(t);
    Evaluate Population POP(t);
    while *termination condition achieved* **do**
        t=t+1;
        Variation(POP(t) $\leftarrow POP(t-1)))$
        $EvaluatePopulationPOP(t)$
        $Adjust(BLF(t), Accept(POP(t)))$;
        $Influence(POP(t), BLF(t));$ **end**
      End
  end

Algorithm 2: Cultural algorithm pseudo code

Chromosome Representation

In proposed model each chromosome is a subset of features. So, the chromosome size (number of genes) is equal to number of features. Each chromosome is represented in binary format in which 1 means corresponding feature is selected and vice versa. As shown in figure 4.8, the number of genes is equal to number of features that and they are equal to n.

Population

A population is a set of chromosome. Also a population is called a generation during the evolutionary process. In figure 4.8 the population size is m. In evolutionary process the initial population is generated randomly and the next generations are obtained through some operations.

Figure 4.9: Chromosome representation

Fitness function

For scoring the chromosomes and sort them by their score, all of them are evaluated by fitness function to be determined how fit are them. In feature selection, fitness function is the objective function. This function determines that which chromosome has the best outcome on classification task. In proposed model SVM classifier is performed as fitness function. For each chromosome and based on the selected features, SVM classifier trains on train set and tests on test set. The classification error of each chromosome is the scores of it.

Knowledge

In proposed model, situational knowledge and normative knowledge is used in cultural algorithm. The best chromosome of each generation influences the current population as knowledge. After reproduction of the next generation, best chromosome of that population is saved as knowledge. So, the belief space is updated.

It has five knowledge components among which here we have used normative knowledge and situational knowledge. Domain knowledge points to concepts, rules and principles that shape the domain of the problem , which can also be considerable

in fake news detection problem in future.

- **Normative knowledge** is knowledge represents a collection of desirable ranges of value which reside within the population space. This demonstrates acceptable behaviour for the individuals in the population.

- **Situational knowledge** refers to the beliefs pertaining to the vital individuals in the search space. It provides a set of important events(e.g. successful/unsuccessful) from an experience of specific individual. [12].

Operators

For reproducing new generation in order to maximize the fitness scores, some operators such as crossover and mutation are used. At first, some of elite chromosomes with the highest fitness score are selected to go to the next generation directly. Afterward, some of remained chromosomes are selected to be changed by crossover. Crossover exchanges substring from pairs of chromosome to generate two new chromosomes. In this study, two-point crossover is employed. In knowledge-based crossover, a copy of the best chromosome is used to produce two new offspring. Then mutation is applied on the remained chromosomes. In mutation, selected genes are inverted. Mutation prevents the search process from falling into local maxima . So, after applying cultural operation on a generation, the new generation is produced.

In order to explore the training process of the model,on selecting optimal value at each iteration training cost has been plotted in figure 4.10. The capability of CA in storing and reusing knowledge can be used to improve the exploitation ability of the search algorithm
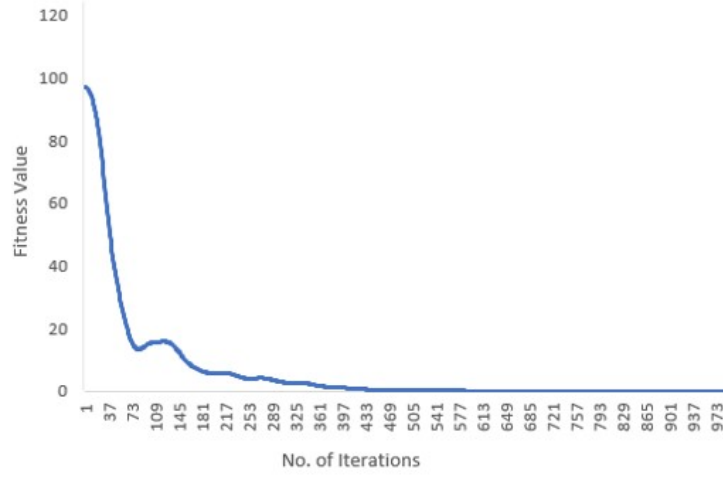
Figure 4.10: Evaluation of cost at each iterations

4.6   Model integration

Inputs : The multi-model input $\{m_i\}_{i=1}^{N}$,Label $\{n_i\}_{i=1}^{N}$ and the learning
rate η

begin

[1] Apply sentiment analysis (S_A) for text and segmentation process for
image (S_p);

[2] Create feature matrix of text (f_t) and image (f_i);

[3] Training set $\{f_t, f_i, i = 1, 2, \ldots n\}$ and weight matrix q_i,i=1,2,...n
passed to Cultural algorithm for feature optimization (γ);

[4] Calculate training set (S), error set (E) ,and remaining set (R);

[5] Apply kernel classification (K);

[6] Based on test data classify and update it with iteration.

end

Algorithm 3: Procedure of multimodal fake news detection

Here, a set of m news article containing the text and image information, we
can represent the data as a collection as a set of text-image tuples denoted as
$A = (A_i^T, A_i^I)_i{}^{\mathrm{m}}$. (S_A) is process of sentiment analysis for textual part of news ar-
ticle and features obtain from it is represented by (f_t).Similarly, process of visual
feature extraction is represented by (S_p) and all the features obtained from images
are represented by (f_i) . All features from text and image are fused together by
simple concatenate method.This combined representation is taken input to Cultural

algorithm for feature optimization denoted by (γ). Optimal features are further used for classification.

4.7   Fake News Detection

The two optimal feature vectors obtained after passing through Cultural algorithms are passed through kernel SVM. Recently, SVMs have grown rapidly, of which the most popular and effective one is kernel SVM. Here we have used radial based function kernel (RBF) for fake news classification. Since the classifier is trained by a given dataset, so it is possible to achieve high accuracy performance only for the particular training dataset and not any other independent datasets.

We need to incorporate cross-validation into our system to prevent over-fitting. Cross-validation will not improve the accuracy of the final classification but will allow the classifier reliable and generalize to other different datasets. Here, we implemented the K-fold cross-validation because of its property, as quick, easy and it uses all data for training and validation. In this study, by means of the trial-and-error process, we empirically calculated K as 10, which implies we assume parameter K varying from 3 to 10 with initial step rising, and then we train the SVM by each value. Finally, we pick the optimum K value corresponding to the valid classification.

Chapter 5

Experiments

In this section we explain first of all the two social media datasets used in the experiments. We then briefly discuss the state-of-the-art approaches to fake news detection. We are also exploring various training variants of our proposed model. Finally we are demonstrating the effectiveness of our proposed model. Comprehensive data-set statistics are provided in Table 5.1.

5.1 Dataset

Given the sparse availability of structured multimedia data, we make use of two standard datasets to evaluate our architecture for fake news detection.Our model training is carried out on two datasets, i.e. Weibo dataset and Twitter dataset, which are publicly available. These are the only databases accessible to the best of our knowledge that has combined picture and textual content.

- **Twitter Dataset**

  The Twitter dataset was published as part of MediaEval [7] for Verifying Multimedia Use task. The goal of the task was to identify fake multimedia material

on social media. The dataset is made up of tweets (short messages posted on Twitter), and each tweet has related textual material, image / video, and social background details. The dataset includes about 17,000 separate tweets spanning different events. The dataset is divided into two parts: the development set (9000 fake news tweets, 6000 real news tweets), and the test set (2000tweets). These are divided in such a way that there are no simultaneous occurrences on the tweets. We filtered out all tweets that have videos attached to them in the set, given our attention on the image and textual content. To retain the same data split scheme as the standard we use the development set for training and the test set for testing.

Total number of content	*Twitter*	*Weibo*
Number of real news	6026	4749
Number of fake news	7898	4779
Total number of images	514	9528

Table 5.1: The statistics of datasets

- **Weibo Dataset**

The Weibo dataset is used in [50] to track fake news. Real news from official Chinese news sources such as Xinhua News Agency was stored in this dataset. The fake news is crawled from May 2012 to January 2016 and checked by Weibo's official gossip debunking system. This system encourages specific users to report suspicious messages, and a committee of trustworthy users is investigating suspicious posts. We follow the same steps of pre-processing done by [50] for a fair comparison. To ensure the quality of the image , we first remove the duplicates and low-quality images. We split the dataset into the training, validation and testing sets in a ratio of 7:1:2 respectively.

5.2   Setting Hyper-parameters

We conduct random search iterations on possible combinations of hyper parameters to pick the right permutation of the hyper parameters. The number of possible permutations is reduced in each iteration, based on the previous iteration's results. Within our model, everything is configurable from the number of generations to the number of population size, acceptance ratio and decision variables. In Cultural algorithm , the population size for all runs was 100 and each run was 1000 generation, which is also a stopping criteria. Table III provides a complete list of hyper parameters.

Parameters	Values
#Decision variable	10
Decision variable range	[1,10]
# of iteration	1000
Population size	100
Acceptance ratio	0.35
Learning rate	0.02

Table 5.2: An overview of hyper parameters used in proposed model

5.3   Evaluation functions

Various assessment criteria were used to assess the efficiency of algorithms for fake news detection. In this section we will discuss the used metrics for calculating accuracy, F1-score and recall for identifying fake news.

- **True Positive (TP)**: when predicted fake news pieces are actually annotated as fake news;

- **True Negative (TN)**: when predicted true news pieces are actually annotated as true news;

- **False Negative (FN)**: when predicted true news pieces are actually annotated as fake news;

- **False Positive (FP)**: when predicted fake news pieces are actually annotated as true news.

We can define following metrics as follow:

$$\textbf{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\textbf{Recall} = \frac{TP}{TP + FN}$$

$$\textbf{Precision} = \frac{TP}{TP + FP}$$

$$\textbf{F1-score} = \frac{2 * Precision * Recall}{Precision + Recall}$$

In the fake news detection problem, the accuracy is the rate of correctly predicted news to all of the samples. Recall value shows the ratio of the fake news, which is correctly predicted over the total number of fake news. Precision metric measures the fake news, which is correctly predicted from the total predicted news in fake class. F-measure is the harmonic average value of the recall value and precision value obtained for fake news detection.

5.4 Baselines

We compared with two types of reference models to validate the efficiency of the proposed model: single modality models and multimodal models.

5.4.1   Single Modality Models

The proposed model leverages both visual and textual data to recognise potentially fake news. We also experimented with two unimodal models as mentioned below, against these multimodal approaches.

- **Textual** : This model uses only textual information present in posts to classify them as fake or not.We use sentiment process to extract textual features of text content from all of posts .

- **Visual** : Images from news article are the input of this model. We use segmentation process to extract visual features of image content from posts.

5.4.2   Multimodal models

For the purpose of fake news detection multimodal methods use data from different modalities.

- **VQA** [3]: Visual Question Answering is intended to answer questions about the given pictures. They adapted the binary classification task to the Visual QA model which was originally developed for a multi-class classification task. This is achieved by substituting a binary-class layer for the final multi-class layer. They use a single layer LSTM set to 32 with number of secret units.

- **NeuralTalk** [48]: In the domain of image captioning, Vinyals et al.[48]'s research proposes the generation of natural language sentences describing an picture using a deep recurrent system. Following a similar structure to theirs, they

provide latent representations by averaging the performance of the RNN as a joint representation of the image and text in tweets at each time stage. Such representations are closely associated with the layer followed by the predictions of an entropy loss layer to make.

- **att-RNN** [24]: Att-RNN uses processes of attention mechanisms to combine textual, visual and social context. The image features are integrated into the combined representation of text and social context in this end-to-end network, collected using an LSTM network. The neural input from the LSTM outputs is an important part of fusing the visual properties.

- **EANN** [50]: The Event Adversarial Neural Network (EANN) is composed of the key components: the extractor of multimodal content, the detector of false news and the discriminator of events. The extractor of multimodal features removes textual and visual features from posts. This works with the fake news detector to learn the discriminative representation of fake news for identification. The event discriminator is responsible for eliminating any unique features of an event and for preserving common features between events. Fake news can also be detected using only two components of the three, the multimodal extractor feature and the fake news detector.

- **MVAE**[28]: It uses a bimodal vector autoencoder combined with a conditional classifier for fake news identification tasks. The platform comprises of three main components, an encoder, a decoder, and a fake news detector element. The variable autoencoder is capable of learning probabilistic latent variable models by maximizing the marginal likelihood of observed data.

5.5 Variants of our model

Our proposed model is fused together the three different components. To investigate the impact of each modality , we simplified version of our model by removing certain components:

- **Textual** : We remove the visual feature extractor, and use only the textual feature extractor component of proposed model.

- **Visual** : We remove the textual feature extractor, and use only the visual feature extractor component of proposed model.
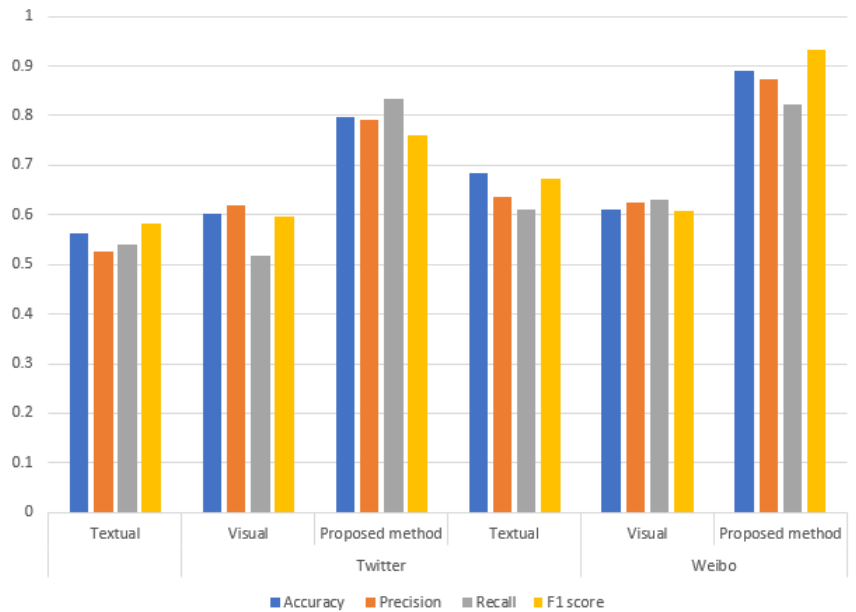
Figure 5.1: Performance of different variant of model

To justify the use of multimodal fake news detection using Cultural algorithm, we compared the performance of single modality models on twitter and weibo datasets. Textual model achieves an f1-score of 0.582 % and visual achieves an f1-score of 0.596 % for twitter dataset , whereas 0.672 % and 0.607 % respectively for weibo. Performance comparison of the single modality and the proposed architecture have been illustrated in figure 5.1.

Chapter 6

Results and Analysis

In this chapter, we first explain the toolkit and environments in which the experiments are performed in Section 6.1. Then, we compare and examine the results obtained by conducting various experiments with the baselines and our proposed method.

6.1 Environments and Tool-kits

In this section, the details of the environment and toolkit that were used for the implementation are described.

Software and Hardware configuration The implementation of proposed methodology was performed on Processor: Intel(R) Core(TM) i7-5820K CPU @ 3.30GHz, 3301 Mhz, 6 Core(s), 12 Logical Processor(s) and 16 GB RAM, running under Windows 10 operating system. The algorithm was in-house developed on Matlab 2018 software.

Item	Detail
OS	Windows 10 (64-bit)
Processor and storage	Core i7 , 16 GB
Language	MATLAB
IDE	MATLAB R2018a
Toolbox	Data Acquisition toolbox Image processing toolbox

Table 6.1: Implementation details

6.2 Performance Comparison

Table 6.2 displays both the baselines results and our proposed approach on both datasets. We report our fake news detector's accuracy as well as our proposed method's accuracy, recall and F1-score. We can easily see how much better our proposed approach does than the baseline methods.

Dataset	*Method*	*Accuracy*	*Precision*	*Recall*	*F1 score*
Twitter	Textual	0.563	0.527	0.539	0.582
	Visual	0.602	0.619	0.518	0.596
	VQA[3]	0.631	0.765	0.509	0.611
	Neural talk[48]	0.610	0.728	0.504	0.595
	att-RNN [24]	0.664	0.749	0.615	0.676
	EANN-[50]	0.648	0.810	0.498	0.617
	EANN [50]	0.718	**0.822**	0.638	0.719
	MVAE-[28]	0.656	N/A	N/A	N/A
	MVAE [28]	0.745	0.801	0.719	0.758
	Proposed Method	**0.798**	0.791	**0.833**	**0.760**
Weibo	Textual	0.683	0.635	0.612	0.672
	Visual	0.610	0.624	0.630	0.607
	VQA[3]	0.736	0.780	0.782	0.781
	Neural Talk[48]	0.726	0.683	**0.843**	0.754
	att-RNN[24]	0.772	0.778	0.799	0.789
	EANN-[50]	0.795	0.806	0.795	0.800
	EANN[50]	0.827	0.847	0.812	0.829
	MVAE-[28]	0.743	N/A	N/A	N/A
	MVAE[28]	0.824	0.854	0.769	0.809
	Proposed Method	**0.891**	**0.873**	0.822	**0.932**

Table 6.2: Performance of Proposed Model V/S other methods on 2 different datasets

6.3   Performance Analysis

Here , we compare and analyze our results with that of other methods. Table shows the performance comparison of the baselines as well as our proposed model on both the datasets. We can observe that the overall performance of the proposed method is better than the current methods in terms of accuracy, precision, recall, and F1 score.

On the twitter dataset, the visual model performs better than the textual model between the single modality models on the Twitter dataset. This can be due to the fact that image features learned with the help of the segmentation process have more sharing patterns than textual features to classify news.Although visual features perform better than textual features, single modality models perform worse than multimodal models. Among multimodal models, att-RNN beats EANN which tells us that the mechanism of attention can help improve model output by considering the aspects of the picture that are linked to the text.

We see similar patterns in the findings at the Weibo dataset. In single modality models the textual model beats the visual model. EANN and att-RNN that were proposed for this task perform better than NeuralTalk and VQA among the multimodal methods.

Both EANN and MVAE have versions of two types each. EANN-/MVAE is when fake news classifier is used independently. EANN / MVAE is when the fake news classifier is focused on a secondary task in the model. In the case of EANN, the secondary function is an event discriminator that excludes event-specific features and retains common features between activities. While in MVAE, the secondary sub-task is to discover the correlations across the modalities to improve shared representation.

Our proposed model on the Twitter dataset gains 24.33% and 18% improvement in accuracy over EANN-and EANN respectively. On Weibo dataset output benefit over EANN-and EANN is 9%, and 5.8% respectively. It outperform MVAE and MVAE by 23.5% and 14.6% respectively as compared to MVAE on twitter dataset.The brief summary of the findings is presented in Table 6.2.

6.4 Error Analysis

All of the above experiments were performed five times, and reported the best result. Although, the five outcomes did not vary by a huge margin. The difference in the observed error for each run of the algorithm was less than 1 . In the error the standard deviation was 0.09. This shows the results the algorithm gave in the experiment were consistent.

6.5 Case Study

To highlight the importance of considering multi-modal features for fake news detection, we compare the results reported by the proposed model and single modality feature models like a textual and visual model, and report the fake tweets that are correctly identified by the proposed model but missed by the single modality feature models.

Consider the tweets shown in Figure 6.1 .We will see that the text content does not provide any evidence to help classify the tweets as fake or not on initial review of the fake news tweets. All the details needed to determine a difference in the tweet's truth value is only obtained from the photos attached to it. Both of the pictures

in the tweets seem questionable and seem to have been altered/forged. Proposed model's ability to identify such tweets as fake restores trust in its ability to account accurately for the visual features derived from images associated with the tweets.

Text: **Five headed snake**   *Text:* **North Korea's Nuclear Missile Threat**

Figure 6.1: Fake News Tweets on the Twitter dataset found by proposed model but missed by Single Text Modality model.

Text: **A #Syrian girl who sells chewing gum in the cold streets of Amman. #Syrian #generation #lost**

Text: **This is scary! 'Samurai Ghost': Mysterious Boots Behind Girl.**

Figure 6.2: Fake News Tweets on the Twitter dataset found by proposed model but missed by Single Text Modality model.

Figure 6.2 shows two more cases where the model proposed outperforms others of the model. We compare the performance of a proposed model with that of a single model for the image modality. In both cases, pictures attached to posts provide little to no support in classifying the pictures as fake or not. The strong emotional content captured by the text for the left-hand tweet and the strange / mysterious content captured by the right-hand text is what plays a critical role in decision-making.Independently, the two images have entirely different meanings compared to the textual nature of their presence. As predicted, a simple model of the image modality classifies wrongly due to the lack of visual indications as discussed earlier. On the other hand, proposed method can account for the textual content associated with the tweets and easily identify fake news with a high degree of confidence.

Chapter 7

Conclusion and Future Work

7.0.1 Conclusion

With social media becoming popular, more and more people are receiving news from social media instead of traditional news media. Social media, however, has also been used to spread fake news, with significant negative impacts on individual users and broader society.The current multimodal state-of-the-art methods suffers from a problem of not being able to learn from fake news detection problems as a primary task. To address this issue, we proposed a model of multi-modal fake news detection using a Cultural Algorithm without considering any sub-task. It outperforms the current methods by an average of 9%. Previous literature has tackled the problem of detecting fake news from a variety of angles, such as natural language processing, knowledge graphics, computer vision, and user profiling.

There can still be a gain in performance for a larger data collection and more complex approaches that clarify how various modalities play an important role in the detection of fake news.

7.0.2 Future work

In future work, we would like to expand the research outlined in this thesis .Incorporating more user-independent features will be one important addition. These include demographics, sex, age, place, to learn more about the readers, and their related patterns of reading. We plan on investigating whether a social media post's dissemination pattern will help us better identify fake news. We would also like to see if the user features could help us in identifying users who are more likely to believe and spread fake news.

Bibliography

[1] Sadia Afroz, Michael Brennan, and Rachel Greenstadt. Detecting hoaxes, frauds, and deception in writing style online. In *2012 IEEE Symposium on Security and Privacy*, pages 461–475. IEEE, 2012.

[2] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2):211–36, 2017.

[3] Stanislaw Antol, Aishwarya Agrawal, Jiasen Lu, Margaret Mitchell, Dhruv Batra, C Lawrence Zitnick, and Devi Parikh. Vqa: Visual question answering. In *Proceedings of the IEEE international conference on computer vision*, pages 2425–2433, 2015.

[4] Esteban Alejandro Armas Vega, Ana Lucila Sandoval Orozco, Luis Javier García Villalba, and Julio Hernandez-Castro. Digital images authentication technique based on dwt, dct and local binary patterns. *Sensors*, 18(10):3372, 2018.

[5] Solomon E Asch. Effects of group pressure upon the modification and distortion of judgments. *Organizational influence processes*, pages 295–303, 1951.

[6] Alessandro Bessi and Emilio Ferrara. Social bots distort the 2016 us presidential election online discussion. *First Monday*, 21(11-7), 2016.

[7] Christina Boididou, Katerina Andreadou, Symeon Papadopoulos, Duc-Tien Dang-Nguyen, Giulia Boato, Michael Riegler, Yiannis Kompatsiaris, et al. Verifying multimedia use at mediaeval 2015. *MediaEval*, 3(3):7, 2015.

[8] Gary D Bond, Rebecka D Holman, Jamie-Ann L Eggert, Lassiter F Speller, Olivia N Garcia, Sasha C Mejia, Kohlby W Mcinnes, Eleny C Ceniceros, and Rebecca Rustige. 'lyin'ted','crooked hillary', and 'deceptive donald': Language of lies in the 2016 us presidential debates. *Applied Cognitive Psychology*, 31(6):668–677, 2017.

[9] Chloé Braud and Anders Søgaard. Is writing style predictive of scientific fraud? *arXiv preprint arXiv:1707.04095*, 2017.

[10] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. Information credibility on twitter. In *Proceedings of the 20th international conference on World wide web*, pages 675–684, 2011.

[11] Yimin Chen, Niall J Conroy, and Victoria L Rubin. Misleading online content: recognizing clickbait as" false news". In *Proceedings of the 2015 ACM on workshop on multimodal deception detection*, pages 15–19, 2015.

[12] Joshua Chittle, Ziad Kobti, et al. A new dimension division scheme for heterogeneous multi-population cultural algorithm. In *The Twenty-Seventh International Flairs Conference*, 2014.

[13] Lena Clever, Dennis Assenmacher, Kilian Müller, Moritz Vinzent Seiler, Dennis M Riehle, Mike Preuss, and Christian Grimme. Fakeyou!–a gamified approach for building and evaluating resilience against fake news. *arXiv preprint arXiv:2003.07595*, 2020.

[14] Song Feng, Ritwik Banerjee, and Yejin Choi. Syntactic stylometry for deception detection. In *Proceedings of the 50th Annual Meeting of the Association for*

Computational Linguistics: Short Papers-Volume 2, pages 171–175. Association for Computational Linguistics, 2012.

[15] Álvaro Figueira, Nuno Guimarães, and Luís Torgo. Current state of the art to detect fake news in social media: Global trendings and next challenges. In *WEBIST*, pages 332–339, 2018.

[16] Christie M Fuller, David P Biros, and Rick L Wilson. Decision support for determining veracity via linguistic-based cues. *Decision Support Systems*, 46(3):695–703, 2009.

[17] Matthew Gentzkow, Jesse M Shapiro, and Daniel F Stone. Media bias in the marketplace: Theory. In *Handbook of media economics*, volume 1, pages 623–645. Elsevier, 2015.

[18] Chuan Guo, Juan Cao, Xueyao Zhang, Kai Shu, and Huan Liu. Dean: Learning dual emotion for fake news detection on social media. *arXiv preprint arXiv:1903.01728*, 2019.

[19] Chuan Guo, Juan Cao, Xueyao Zhang, Kai Shu, and Miao Yu. Exploiting emotions for fake news detection on social media. *arXiv preprint arXiv:1903.01728*, 2019.

[20] Aditi Gupta, Hemank Lamba, Ponnurangam Kumaraguru, and Anupam Joshi. Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. In *Proceedings of the 22nd international conference on World Wide Web*, pages 729–736, 2013.

[21] Naeemul Hassan, Chengkai Li, and Mark Tremayne. Detecting check-worthy factual claims in presidential debates. In *Proceedings of the 24th acm international on conference on information and knowledge management*, pages 1835–1838, 2015.

[22] Valerie Hauch, Iris Blandón-Gitlin, Jaume Masip, and Siegfried L Sporer. Are computers effective lie detectors? a meta-analysis of linguistic cues to deception. *Personality and social psychology Review*, 19(4):307–342, 2015.

[23] Minqing Hu and Bing Liu. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177, 2004.

[24] Zhiwei Jin, Juan Cao, Han Guo, Yongdong Zhang, and Jiebo Luo. Multimodal fusion with recurrent neural networks for rumor detection on microblogs. In *Proceedings of the 25th ACM international conference on Multimedia*, pages 795–816, 2017.

[25] Zhiwei Jin, Juan Cao, Yongdong Zhang, Jianshe Zhou, and Qi Tian. Novel visual and statistical image features for microblogs news verification. *IEEE transactions on multimedia*, 19(3):598–608, 2016.

[26] Eric Kee and Hany Farid. A perceptual metric for photo retouching. *proceedings of the National Academy of Sciences*, 108(50):19907–19912, 2011.

[27] Nitin Khanna, Aravind K Mikkilineni, George TC Chiu, Jan P Allebach, and Edward J Delp. Forensic classification of imaging sensor types. In *Security, Steganography, and Watermarking of Multimedia Contents IX*, volume 6505, page 65050U. International Society for Optics and Photonics, 2007.

[28] Dhruv Khattar, Jaipal Singh Goud, Manish Gupta, and Vasudeva Varma. Mvae: Multimodal variational autoencoder for fake news detection. In *The World Wide Web Conference*, pages 2915–2921, 2019.

[29] Yoon Kim. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*, 2014.

[30] Tony Lesce. Scan: Deception detection by scientific content analysis. *Law and Order*, 38(8):3–6, 1990.

[31] Francesco Marra, Diego Gragnaniello, Davide Cozzolino, and Luisa Verdoliva. Detection of gan-generated fake images over social networks. In *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, pages 384–389. IEEE, 2018.

[32] George A Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.

[33] Raymond S Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2):175–220, 1998.

[34] Jinie Pak and Lina Zhou. A comparison of features for automatic deception detection in synchronous computer-mediated communication. In *2015 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pages 141–143. IEEE, 2015.

[35] Jeff Z Pan, Siyana Pavlova, Chenxi Li, Ningxi Li, Yangmei Li, and Jinshuo Liu. Content based fake news detection using knowledge graphs. In *International Semantic Web Conference*, pages 669–683. Springer, 2018.

[36] Martin Potthast, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff, and Benno Stein. A stylometric inquiry into hyperpartisan and fake news. *arXiv preprint arXiv:1702.05638*, 2017.

[37] Peng Qi, Juan Cao, Tianyun Yang, Junbo Guo, and Jintao Li. Exploiting multi-domain visual information for fake news detection. *arXiv preprint arXiv:1908.04472*, 2019.

[38] Robert G Reynolds. An introduction to cultural algorithms. In *Proceedings of the third annual conference on evolutionary programming*, pages 131–139. World Scientific, 1994.

[39] Hua Shen, Fenglong Ma, Xianchao Zhang, Linlin Zong, Xinyue Liu, and Wenxin Liang. Discovering social spammers from multiple views. *Neurocomputing*, 225:49–57, 2017.

[40] Somayeh Shojaee, Masrah Azrifah Azmi Murad, Azreen Bin Azman, Nurfadhlina Mohd Sharef, and Samaneh Nadali. Detecting deceptive reviews using lexical and syntactic features. In *2013 13th International Conference on Intellient Systems Design and Applications*, pages 53–58. IEEE, 2013.

[41] Kai Shu, Deepak Mahudeswaran, Suhang Wang, Dongwon Lee, and Huan Liu. Fakenewsnet: A data repository with news content, social context and spatial-temporal information for studying fake news on social media. *arXiv preprint arXiv:1809.01286*, 2018.

[42] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1):22–36, 2017.

[43] Michael Siering, Jascha-Alexander Koch, and Amit V Deokar. Detecting fraudulent behavior on crowdfunding platforms: The role of linguistic and content-based cues in static and dynamic contexts. *Journal of Management Information Systems*, 33(2):421–455, 2016.

[44] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[45] Shivangi Singhal, Rajiv Ratn Shah, Tanmoy Chakraborty, Ponnurangam Kumaraguru, and Shin'ichi Satoh. Spotfake: A multi-modal framework for fake

news detection. In *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)*, pages 39–47. IEEE, 2019.

[46] Henri Tajfel, John C Turner, William G Austin, and Stephen Worchel. An integrative theory of intergroup conflict. *Organizational identity: A reader*, 56:65, 1979.

[47] Sarah Vieweg, Amanda L Hughes, Kate Starbird, and Leysia Palen. Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1079–1088, 2010.

[48] Oriol Vinyals, Alexander Toshev, Samy Bengio, and Dumitru Erhan. Show and tell: A neural image caption generator. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3156–3164, 2015.

[49] Aldert Vrij. Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, 11(1):3, 2005.

[50] Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, Guangxu Xun, Kishlay Jha, Lu Su, and Jing Gao. Eann: Event adversarial neural networks for multi-modal fake news detection. In *Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining*, pages 849–857, 2018.

[51] Andrew Ward, L Ross, E Reed, E Turiel, and T Brown. Naive realism in everyday life: Implications for social conflict and misunderstanding. *Values and knowledge*, pages 103–135, 1997.

[52] Ian T Young, Jan J Gerbrands, and Lucas J Van Vliet. Fundamentals of image processing. 1998.

[53] Dongsong Zhang, Lina Zhou, Juan Luo Kehoe, and Isil Yakut Kilic. What online reviewer behaviors really matter? effects of verbal and nonverbal behaviors on detection of fake online reviews. *Journal of Management Information Systems*, 33(2):456–481, 2016.

[54] Xichen Zhang and Ali A Ghorbani. An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, 57(2):102025, 2020.

[55] Lina Zhou, Judee K Burgoon, Jay F Nunamaker, and Doug Twitchell. Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications. *Group decision and negotiation*, 13(1):81–106, 2004.

[56] Lina Zhou and Azene Zenebe. Representation and reasoning under uncertainty in deception detection: A neuro-fuzzy approach. *IEEE Transactions on Fuzzy Systems*, 16(2):442–454, 2008.

[57] Xinyi Zhou and Reza Zafarani. Fake news: A survey of research, detection methods, and opportunities. *arXiv preprint arXiv:1812.00315*, 2018.

Appendix A

All the experiments were performed using Matlab library. To develop the machine learning models for research purposes I strongly suggest using Matlab. For plotting the graphs, I have used Matlab plotting.

The Cultural Algorithm was created using the ideas from Yarpiz (available : `http://yarpiz.com/425/ypea125-cultural-algorithm`) who developed CA in Matlab. One may need to change the hyper parameters,proposed components or even the code used in above reference to better suit their dataset and experiments. Our source code is publicly available at GitHub `https://github.com/priyanshishah/Multimodal-fakenews-detection`

The code was executed on NVIDIA GeForceGTX with dedicated GPU memory of 16 GB. The code was executed on an Intel® Core™ i7-5820 CPU @ 3.30 GHz processor.Each run of the experiment took around 8-9 hours, the run-time of the algorithm can be improved by increasing the GPU memory.

Vita Auctoris

NAME: Priyanshi Shah

PLACE OF BIRTH: Gujarat , India

YEAR OF BIRTH: 1997

EDUCATION: B.E, Silver Oak College of Engineering
and Technology, Ahmedabad, India, 2018

Master of Science in Computer Science, University of
Windsor, Windsor, ON, Canada, 2020