Fake News Detection using Deep Learning

Submitted By

Yagnesh Bhadiyadra

21MCEC11

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2023

Fake News Detection using Deep Learning

Major Project - II

Submitted in partial fulfillment of the requirements

for the degree of

Master of Technology in Computer Science and Engineering
(Computer Science and Engineering)

Submitted By

Yagnesh Bhadiyadra
(21MCEC11)

Guided By

Dr. Ankit Thakkar

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
INSTITUTE OF TECHNOLOGY
NIRMA UNIVERSITY
AHMEDABAD-382481

May 2023

Certificate

This is to certify that the major project entitled **"Fake News Detection using Deep Learning"** submitted by **Yagnesh Bhadiyadra (Roll No: 21MCEC11)**, towards the partial fulfillment of the requirements for the award of the degree of Master of Technology in Computer Science and Engineering (Computer Science and Engineering) of Nirma University, Ahmedabad, is the record of work carried out by him under my supervision and guidance. In my opinion, the submitted work has reached the level required for being accepted for examination. The results embodied in this major project part-II, to the best of my knowledge, haven't been submitted to any other university or institution for the award of any degree or diploma.

Dr. Ankit Thakkar
Guide & Professor,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr. Sudeep Tanwar
Professor,
Coordinator M.Tech - CSE (CSE)
Institute of Technology,
Nirma University, Ahmedabad

Dr. Madhuri Bhavsar
Professor and Head,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Dr R. N. Patel
Director,
Institute of Technology,
Nirma University, Ahmedabad

Statement of Originality

I, **Yagnesh Bhadiyadra**, Roll. No. **21MCEC11**, give an undertaking that the Major Project entitled "**Fake News Detection using Deep Learning**" submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in **Computer Science and Engineering (Computer Science and Engineering)** of Institute of Technology, Nirma University, Ahmedabad, contains no material that has been awarded for any degree or diploma in any university or school in any territory to the best of my knowledge. It is the original work carried out by me and I give assurance that no attempt of plagiarism has been made. It contains no material that is previously published or written, except where reference has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

Signature of Student
Date: May 23, 2023
Place: Ahmedabad

Endorsed by
Dr. Ankit Thakkar

Acknowledgements

It gives me immense pleasure in expressing thanks and profound gratitude to **Dr. Ankit Thakkar**, Professor, Computer Engineering Department, Institute of Technology, Nirma University, Ahmedabad for his valuable guidance and continual encouragement throughout this work. The appreciation and continual support he has imparted have been a great motivation to me in reaching a higher goal. His guidance has triggered and nourished my intellectual maturity that I will benefit from, for a long time to come.

It gives me immense pleasure to thank **Dr. Madhuri Bhavsar**, Hon'ble Head of Computer Engineering/ Information Technology Department, Institute of Technology, Nirma University, Ahmedabad for her kind support and for providing basic infrastructure and a healthy research environment.

A special thank you is expressed wholeheartedly to **Dr. R. N. Patel**, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad for the unmentionable motivation he has extended throughout the course of this work.

I would also thank the Institution, and all faculty members of the Computer Engineering Department, Nirma University, Ahmedabad for their special attention and suggestions towards the project work.

I would also like to thank **Mr. Nisarg Vyas**[1], Executive Director, Infocusp Innovations Pvt. Ltd. for providing me the flexibility of work to conduct both the company work and research project parallelly, and also I would like to thank Infocusp Innovations for allowing me to use office machines to run simulations.

I would also like to thank **Dr. Jamal Nasir**[2], and other authors of base paper for providing me the code when I was stuck, and other guidance related to the base paper.

- Yagnesh Bhadiyadra

21MCEC11

[1] https://www.linkedin.com/in/nisargvyas/

[2] https://www.linkedin.com/in/jamal-nasir-44153ba/

Abstract

Social media plays a significant role in people's daily lives. More people read news online than in conventional newspapers. The risk of spreading false information is rising as online news outlets start to grow and social media applications gain more and more user popularity. Society is seriously harmed by fake news. Multimedia news is becoming more commonplace alongside text-based news [3]. To properly identify false news nowadays, several modalities, including pictures, audio, and video, must be taken into account. In this article, we present a thorough analysis of early, late, and hybrid fusion-based false news detection methods. We use two publicly available sources to show a hybrid CNN-RNN technique for false news identification. Along with that, we also try the Transformers-based approach to improving the results of the FAKES dataset. We describe the further improvement efforts which are data-specific and use more complex models than the baselines themselves.

[3] https://www.pewresearch.org/fact-tank/2021/01/12/more-than-eight-in-ten-americans-get-news-from-digital-devices/

Abbreviations

BERT	Bidirectional Encoder Representations from Transformers.
CNN	Convolutional Neural Networks
ANN	Artificial Neural Networks
SHAP	SHapley Additive exPlanations
RNN	Recurrent Neural Networks
LSTM	Long Short-Term Memory Networks
FFNN	Feed Forward Neural Network

–

Contents

List of Figures

Chapter 1

Introduction

1.1 General

Before the internet gained popularity, most news was read on paper by people all over the world. The situation eventually altered as technology advanced and televisions became popular. Televisions made it easier for news to reach more people. Everything changed when the internet era began in the late 2000s. Since its beginnings, it has gained so much popularity that everything has begun to be provided over the Internet. After the internet gained popularity and a large audience, several news outlets and publications eventually launched online versions. Smartphones allowed individuals to access news anywhere, which benefited the internet news sector. By the end of the year 2022, the way people consume news has undergone a dramatic transformation. Social Media platforms serve as a prime medium for millennials to consume news. According to a study, more than 80

News is a very important part of humanity. It can influence people to take action against injustice, it can provoke people to do inhumane activities, it can change political decisions, and most importantly, it can create hate among the people and change the stock market scenario. Just on receiving news of Russia Invading Ukraine, the Sensex suddenly tanked 2700 points [1]. We can see that as soon as the US IT sector reports a loss, the Indian IT sector and the entire industry suffers immediately, whether you are a public company or not [2]. After the Internet became very cheap in India around 2016, and widespread usage of smartphones, people are using online news (or rather social media

[1] https://economictimes.indiatimes.com/markets/stocks/news/sensex-crashes-over-1300-points-as-russia-orders-military-action-in-ukraine/articleshow/89789291.cms

[2] https://www.tomorrowmakers.com/financial-planning/how-will-recession-us-impact-indian-it-industry-article

based news) as the primary news source. Many Indian prime news organizations have started their online additions. The news provided by internet news outlets is simple to digest; some outlets, like Inshorts, offer 50-word news to encourage readers. With just one sharing, the news may be made available to a larger audience. The news is rich in media; the news piece includes audio, video, and photographs in addition to text. Compared to news that is solely text-based, photos and other multimedia information are simpler to grasp. The news is purposefully crafted to pique the user's curiosity about the specific subject. Here's when bogus news starts to matter. This has two sides to it. It's possible that those who forward bogus news are doing it on purpose or inadvertently. Whatever the cause, it is detrimental to society and we must halt the propagation of false information.

There are several definitions of false news. However, two features of this are well-liked: Factuality and harmfulness come first. It is considered misinformation by the EU. Disinformation is defined as "False Information" and "Items Spread to Harm Others." Kai Shu and others define [1] fake news as "Fake news is a news article that is intentionally and verifiably false". However, this definition disregards satire, rumors from sources other than news, etc.

1.2 The Need for Multimodality Fusion Techniques

As described earlier, media-rich news is becoming popular. There are various known reasons for that which include on-demand availability, direct sharing, etc. These particular news articles not only contain text, but also contain Images, Audio, or video. Users are drawn to multimedia news, and readers naturally pay greater attention to multimedia news than to text-based news. When multimedia news is used instead of text news, the likelihood of fake news spreading increases. The spread of fake news can destroy democracies, and financial markets, and cause catastrophes. The 2016 US election is a prime example of that. Some individuals damaged 5G towers during COVID because of a rumor that 5G towers disseminated COVID [3]. We don't require any further evidence to comprehend the significance of misleading news. Every day, hundreds of thousands of news articles are created globally, making it challenging for people to read them all. Consequently, automated false news detection is essential right now. False news must also be

[3] https://www.vox.com/recode/2020/4/24/21231085/coronavirus-5g-conspiracy-theory-covid-facebook-youtube

identified as soon as possible since it is quite difficult to get it out of any internet-based social media network once it has been posted there. The truth was wearing boots when Fake News started roaming the globe, as the saying goes.

1.3 Objective of Study

Through this paper, we try to provide a brief survey on Fake News Detection which does not only include text, but also other modalities such as images, video, and audio. We devise our survey into three types of fusion techniques: Early Fusion, Late Fusion, and Hybrid Fusion. We also try to implement fake news detection on a publicly available dataset FA-KES [2] and try to reproduce the results given in the state-of-the-art result [3]. We also apply transformers and see the results of the given dataset.

1.4 Scope of Work

The work provides a detailed survey on three types of fusion-based fake news detection, and along with that the survey also includes evolutionary algorithm-based fake news detection which is not explored in the available surveys to the best of the author's knowledge. The work also describes the reproduction of results of a paper that detects fake news on two publicly available datasets FA-KES [2] and ISOT [4].

[4] https://www.uvic.ca/ecs/ece/isot/datasets/fake-news/index.php

Chapter 2

Literature Survey

2.1 Search Strategy

The researchers searched Google Scholar for articles containing the terms "Multimodal Fake News Detection","Fusion based Fake News Detection", "Early Fusion based Fake News Detection", "Late Fusion based Fake News Detection", "Hybrid Fusion based Fake News Detection", and "Evolutionary Algorithm based Fake News Detection". The terms "Misinformation", "Propaganda", and "False News" have taken the place of "Fake News". The authors also searched for research publications on fake news identification using "Media-rich" false news detection. In order to find articles that discussed evolutionary algorithms, the authors employed "Evolutionary Algorithm-based Fake News Detection".

2.2 Related Surveys

Fake news detection methods using many modalities, including audio, video, picture, network, and temporal information, were surveyed by Firoj Alam et al. [4]. This study only examined the identification of false news based on multiple modalities individually and lacked a thorough examination of fusion-based approaches to do so. The authors of [8] investigated altering facial features in films and other forms of facial manipulation in an effort to spot phony videos. They also concentrated on identity switching, face synthesis, and skin color alteration. The major focus of this study was on face modifications. The authors [5] investigated artificial face fabrication, including face synthesis, identity switching, modifying the color of the skin and hair, and other techniques. They explained.

Table 2.1: Survey Comparison

Author - Details	**Images**	**Text**	**Audio**	**Video**	**Early Fusion**	**Late Fusion**	**Hybrid Fusion**	**Evolutionary Algorithm based Algorithm**
[4]	Y	Y	Y	Y	Y	Y	Y	
[5]	Y	Y		Y				
[6]	Y	Y	Y	Y	Y	Y		
Our Survey	Y	Y	Y	Y	Y	Y	Y	Y

The paper goes into great detail on all the approaches.

In a very thorough analysis, Chahat Raj and Priyanka Meel [6] looked at studies on early fusion and late fusion, as well as specifics on the datasets they utilized. Meme categorization is outlined in a framework by Tariq Habib Afridi et al. [7]. News is divided into hateful and non-hateful categories. However, we don't find any polls that use evolutionary techniques. In this survey, we also consider analyses based on evolutionary algorithms. The survey comparison is shown in Table 2.1. There was one other analysis done based on fake news detection which focused on different aspects of fake news detection. This survey [8] focused on three different parts of fake news detection: Multimodal data collection, multimodal feature selection, and multimodal model selection. The survey discussed here focuses more on the different fusion techniques to improve the quality of only text-based news detection.

2.3 Fake News Detection using Early Fusion

According to [1], early fusion is the process of merging several characteristics to create a common feature vector, which is subsequently used to further categorize the data using a model. The early fusion approach is depicted in Figure 2.1. [9] provides descriptions of many early fusion methods.

[1] https://medium.com/haileleol-tibebu/data-fusion-78e68e65b2d

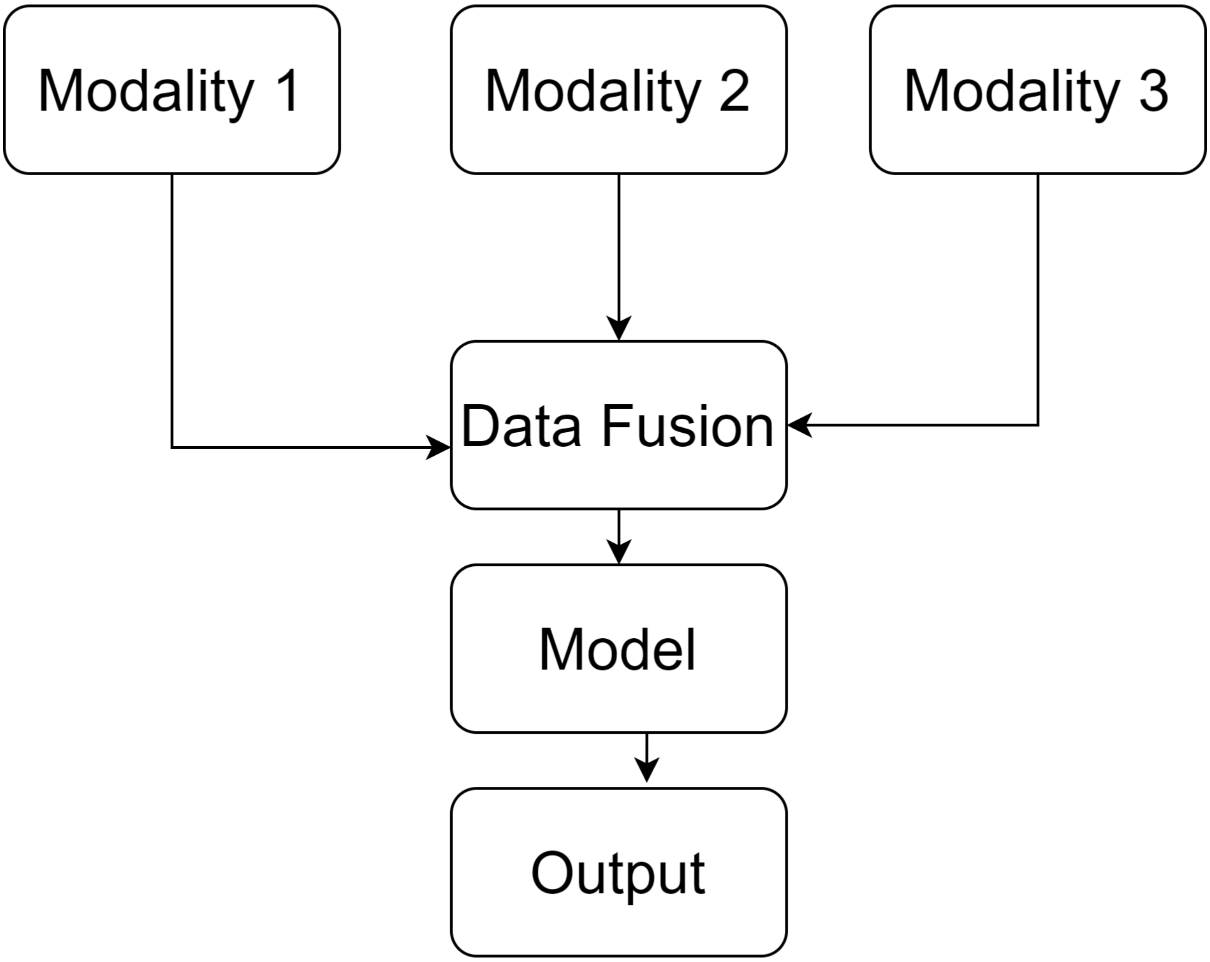

Figure 2.1: Early Fusion

Data loss at an early stage is one of the drawbacks of early fusion. Another situation where early fusion falls short is when we analyze video frames, where fusing several frames to create a single feature might be challenging. A very simplistic approach demonstrating early fusion-based fake news detection was proposed by Manoj Kumar Balwant where he passed the LIAR dataset[10] (News Content and Profile Information respectively) to CNN and BiLSTM (Bi-Long Short Term Memory) network to detect the fake news. They used a concatenation technique to fuse the results which came from different models. Similar model structure, although different news data was taken by Shahin Raza et al.[11] to detect fake news, this approach took different metadata information and not just the news content. The information they took were news content(news headline, news content, news source, publisher/author, publication timestamp) and social content(user ID, number of posts of that news, number of upvotes/likes, number of downvotes/dislikes, source of news, post title, replies to the post, user's credibility). This allowed them to focus on the wider breadth of the news and allowed them to detect fake news. For both the feature-set, they used Bidirectional and AutoRegressive Transformers(BART [12]). BART is based on transformers (BERT [13]), and Autoregressive models like GPT-2 [14]. Autoregressive models are models which depend on previous predictions for current prediction. They differ from neural networks such that AutoRegressive models only depend on the previous result to form the current result whereas neural networks take various other things into account like other things and can be useful for non-time-series data as well. [15] also did similar things using a different set of features with PHEME [16]. They also used SHAP [17] method to determine which features contributed the most to detecting fake news. One more approach tried to use the vectorized representation of news articles using TransE models [18]. The TransE model treats data as vectors and the difference between two entities is represented as the difference between their vectors. TransE models are trained to minimize the difference between the vectorized error and the actual error. Here they used two customized versions of TransE models to get two different feature representations and fused them to get the results. More approaches include semantic relationships-based approaches where authors try to explore relationships between different modalities in order to get the benefit of cross-model information. For example, if you can detect from a post that the image is saying the exact assertion but the text is saying the opposite then we can be pretty sure that the post must be fake news. One fusion-based approach used three things in a news article – Justification (News article), Claim (News Headline), and Metadata of the News(Subject, Speaker, Location, etc). They did sentence-level fusion with justification and claim after passing them through GloVe (Global Vectors for Word Representations) [19], ELMo [20], and ULMFit[21]. They did word level fusion where claim and metadata are used after passing them through ULMFit[21] and char-CNN[22]. Now one more approach where semantic information was sort of used was by Yingtong et al.[23]. They used users' previous social engagements, history, etc, along with the news content. They modeled users' endogenous context and news article content using BERT (Bidirectional Encoder Representation from Transformers). Now one more aspect of fake news is how you take it, for example, recently Sundar Pichai asked Google Employees to focus more on productivity as quarterly results came below expectations. Although one newspaper headline kept it like "Sundar Pichai says Google has too many employees but too few work, issues warning"[2]. So, this is misleading. Chuan Gao et

[2]https://www.indiatoday.in/technology/news/story/sundar-pichai-says-google-has-too-many-employees-but-too-few-work-issues-warning-1982992-2022-08-02

al [24] focused on this particular issue. In their early fusion-based false news detection technique, Another early fusion-related method was put out by [25], who employed text data to provide results on the LIAR [10] dataset that was almost 89% accurate. The dimensionality reduction methods MaxPooling and AveragePooling were combined with the BiLSTM-BiGRU model by the authors.

Now let's focus on Fake News Detection using Text and Image modalities combined. Text and Images are probably the most used modality combination to spread fake news. We see it across the internet –Facebook/Instagram/Twitter posts, WhatsApp status/messages, etc, and whatnot. Traditional newspapers also cherish the pool of information a single image can bring. That same pool can create havoc as well. For example, in the recent Gujarat Election, A flier was circulating on social media about AAP (Aam Aadmi Party) that displayed some promises that were false deliberately with Kejriwal's photo on it [3]. So fusion techniques are used to help people stay away from false information like this.

An approach like this was proposed in 2021 [26] by Mina Kumari and Asif Ekbal. Four distinct models were employed. They employed Attention Based Stacked Bidirectional Long Term Short Memory for textual feature extraction (ABS-BiLSTM). They employed Multimodal Factorized Bilinear Pooling (MFB) for fusing and Attention Based Multilevel Convolutional Neural Network-Recurrent Neural Network (ABM CNN-RNN) for multilevel convolutional neural networks. They used multilayer perceptrons to categorize.

Many methods for early fusion-based false news detection have been put forth, albeit some go beyond just fusing text and pictures and instead look for semantic connections to improve fake news prediction. H. Yang et al. presented one such method [27]. Using the COCO image captioning dataset and the text from the article, they attempted to extract features from the picture data. For pictures and text, they employed the ResNet and BERT models, respectively. Sara Sabour et al.[28] suggested a different strategy going in the same direction. To match visual entities with text data, they adopted a fine-grained technique in which they examined each word in a phrase and each pixel of the image. Cosine similarity was employed in another straightforward method to determine the semantic relationship between text and visuals. To determine the cosine similarity, they employed the article's title and image tag embedding. Dhruv Khattar et al. [29]

[3] https://www.vishvasnews.com/english/viral/fact-check-viral-post-with-purported-aap-gujarat-election-manifesto-is-fake/

suggested another method where dimensionality reduction and latent representation were taken into consideration. They made use of autoencoders. To eliminate noise and perform dimensionality reduction, autoencoders provide a latent representation of the data. In this manner, they paved the ground for a reliable false news detector. One more approach to fake news detection was about identifying minute details about text and minute details about images. [30] took a fine-grained and coarse-grained approach for both textual entities and image entities. The fine-grained method did break text word by word and image into pixel by pixel and uses these details as an advantage to detect fake news. They used capsule networks[28] to fuse the different modalities. Capsule networks are neural networks such that small neurons such that they are processed in a hierarchical manner and better preserve spatial relationships between objects in an image. Social media posts that are about news are not limited only by text and images, they also have images with text embedded, for example, memes. Now some news articles/posts which contain images (with text embedded) + text are also prone to false information and [31] focused on those sets of news articles as well. They took three considerations – One was the correspondence between the text in the news article and the text in the picture (embedded). OCR methods were utilized to extract the text from the picture, and a co-attention mechanism was employed to determine any relationships between the text characteristics of the news and the text features of the image that were acquired using BERT [13]. Entity consistency management, which focuses on instances where an entity described in both the text and the picture is different, was the other area they concentrated on. In order to create better relationships between visual and textual elements, the co-attention model aligns the visual features generated by CNN with the textual information. The final stage here is video-based fake news. Video-based fake news is also getting popular. Video-based fake news detection was worked upon by [32] in 2022 where the authors used the domain knowledge for better classification of fake news. They used YouTube video data in three different ways: they used video thumbnails and video frames and ran them through VGG19 [33], they used multilingual BERT [13] to encode comments after BiLSTM [34] and attention, and they used BERT [13] to encode title/description after Multilingual BERT, then Text-CNN. The three feature sets were combined using a linear combination, and the phony news was categorized using MLP after that. The next section describes another type of fusion, Late Fusion, and techniques used in literature to detect fake news.

2.4 Fake News Detection using Late Fusion

Ensemble classifiers serve as the basis for the Late Fusion approach [4]. When there is a significant discrepancy between data dimensions, units of measures, etc. across several modalities, it is employed. Through late fusion, several modalities can run various classifiers, obtain the results, and then combine them using methods like average fusion and max fusion. Figure 2.2 provides a description of Late fusion. The Late fusion approach is simpler than the Early fusion method [5].

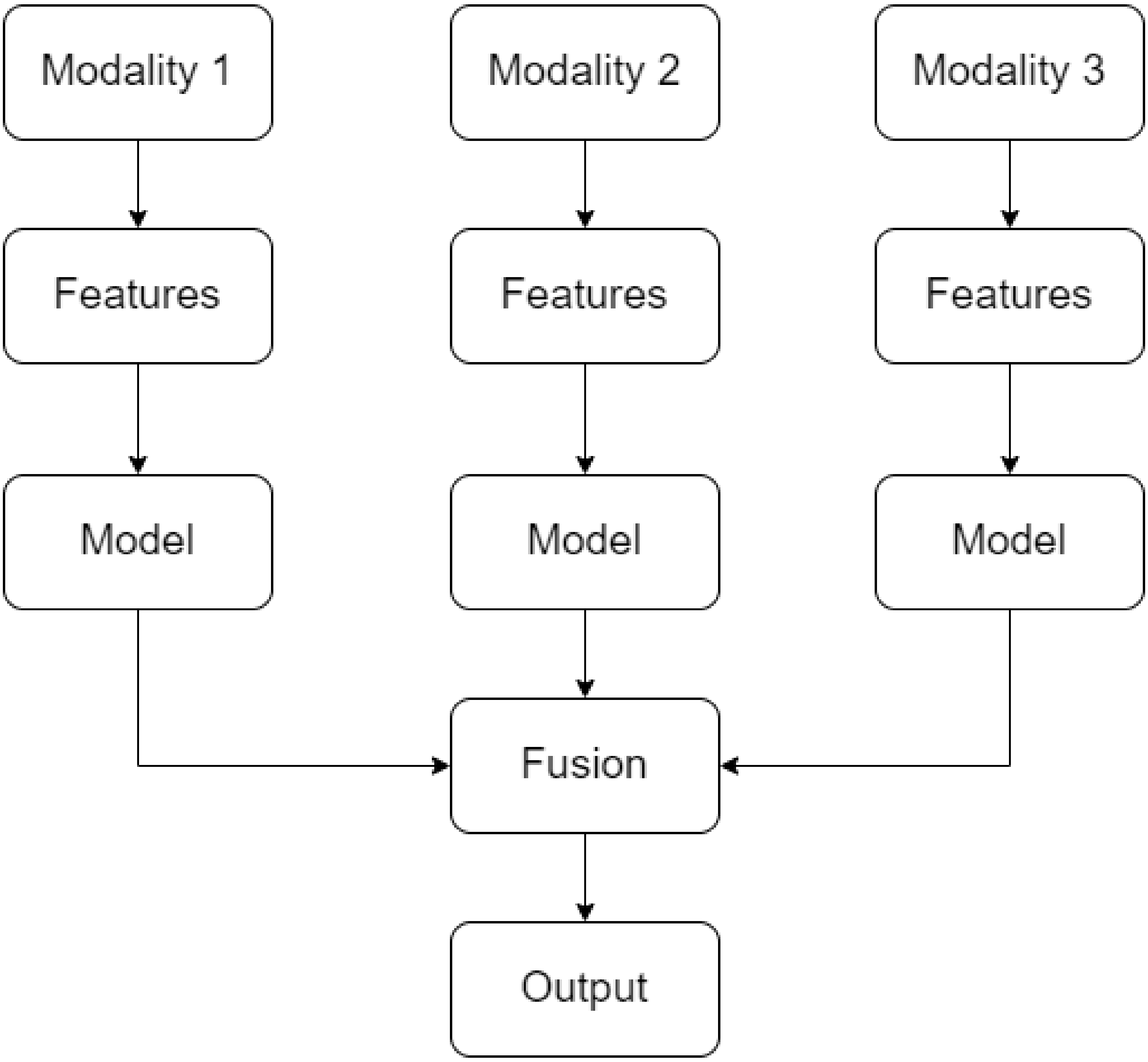

Figure 2.2: Late Fusion

[4]https://medium.com/haileleol-tibebu/data-fusion-78e68e65b2d

[5]https://medium.com/haileleol-tibebu/data-fusion-78e68e65b2d

Shengyi et al. in 2019 [35] suggested one method in which the authors attempted to identify the users who would disseminate fake news using a network-user-based network representation learning. They create a user-news network made up of users, their network of friends, and the news that has been spread. Two specific items, user and news representation, were obtained via the user-news network. To obtain the final findings, both representations were concatenated and late fusion was used with the news content. Results were obtained using datasets from Politifact and BuzzFeed that were given by FakeNewsNet [36]. One method based on late fusion with the Inception-ResNetv2 model, ALBERT [37], and BERT [13] was suggested. To improve outcomes on even smaller datasets, the methodology concentrated on transfer learning. For late fusion, the weighted average of probability was taken into consideration. After doing late fusion, the choice of classifiers can change the results of the fake news classifiers, this approach was also explored. A third strategy suggested by Pallabi Saikia et al. [38] in 2022 attempted to examine both news content and propagation graphs and fuse the characteristics gathered from each of them using two separate late fusion algorithms. The first was an aggregate, while the second was a meta classifier built on top of the underlying classifier. The propagation network was trained using GNN (Graph Neural Networks) [39] and the text features were learned using BERT [13] using the Politifact and Gossipicop datasets [36]. After that, either the mean/classifier or late fusion procedures were used. Mean worked well for Gossipicop, however, the meta-classifier was an efficient late fusion strategy for Politifact. With Late Fusion, we see fewer approaches that focus on only text-based approaches. Next are the Late Fusion-based approaches which use text and image both as modalities. The earlier-mentioned article [6] also suggested late fusion strategies to categorize false information. Six multi-modal datasets, including ReCOVery[40], CoAID [41], MediaEval 2020 [42], and their own proposed CovID dataset, was used to assess their approach. Text characteristics were implemented using Bi-LSTM [34]. The image was created using CNN and bottleneck feature extraction. Later, late fusion was carried out. Following early fusion, average fusion, sum fusion, and max fusion as the most accurate technique was weighted-average fusion. As a consequence, when the modalities were merged, the framework worked best when text and visuals were given the proper weights. In 2020, Abdullah Hamid et al. [42] offered six distinct approaches for binary and ternary

[6]https://medium.com/haileleol-tibebu/data-fusion-78e68e65b2d

classification in relation to the identification of false news. This was a component of the MediaEval 2020 competition. To handle the uneven dataset, they adopted late fusion. The greatest results came from doing N iterations over various samples within the dataset and merging them using a majority voting mechanism. They combined a logistic regression model with BERT [13]encoding. These are some good models/architectures which solved the problem of Fake News Detection with Late Fusion which were using Image and Text modalities combined. There were also some ideas where separately focusing on detecting forgery became a vital part of the fake news detection pipeline. In 2021, Junxiao et al. [43] put out a novel strategy in which they integrated a distinct module for visual manipulation. The strategies that we have so far seen combined text and picture data before doing detection. Here, the authors suggest a novel Visual Tamper detection module that essentially looks for visual picture tampering. The authors here first do Error Level Analysis followed by ResNet50[44] for detecting malicious features and recompression characteristics. The textual features were extracted by BERT-BiGRU (Gated Recurrent Unit) stacked model. For visual (image) features, the authors have used ResNet50 [44] along with BiGRU. The attention module at the end assigns weights to different modality features, and fake news detection is done based on the final features obtained using late fusion. The authors [45] beat results beating state-of-the-art in Twitter, Weibo, and MediaEval 2016 by using ALBERT [37]and Inception-ResNetv2 for text and image respectively. Yet another method, known as the inter- and intra-modality approach, sought to establish a connection between text and visuals. In 2022 [46], Shivangi Singhal et al. suggested a method in which visual characteristics are retrieved using bottom-up attention followed by a pooling layer, and textual features are extracted via BERT. To accomplish late fusion and obtain the result, multiplicative fusion with model loss was utilized. One of the very few models that employed multiplicative fusion was this one. They made use of Twitter, MediaEval, and publicly accessible datasets. There are some approaches that took Videos as a parameter and applied fake news detection to that. We have already seen quite a few approaches that work with images and text, the next articles focus on Video-Based Fake News Detection with Late Fusion. For deepfake detection, Facebook AI [47] put up yet another promising strategy in 2021. For deep fake detection, they essentially attempted to merge audio and visual information from audio. For instance, when a guy speaks, his mouth shape and words may both be indicative of something. Now that the video has been altered, the altered face might not exactly match the original footage. For instance, when uttering a certain word, a person's lips may expand fully or not at all. The fabricated video could miss capturing such information. The writers of this article are attempting to depict this sort of disparity through the merging of audio and video modalities. Chau et al [48] .'s deep fake and visual forgery detection method was another solution that was put out in 2020. In this method, the images/video frames were split into two representations: the original and the frequency domain representation. The authors took original and frequency domain representations while also extracting the face from the image. Following that, late fusion was used to concatenate the two representations after each was put into a separate EfficientNet [49]. A CNN version called EfficientNet [49] is employed to strike a compromise between efficacy and efficiency. Next is described Hybrid Fusion based Fake News Detection.

2.5 Fake News Detection using Hybrid Fusion

Late and early fusion were merged into hybrid fusion. We combine characteristics from many modalities through early fusion. We combine the result from it with the output from the remaining modalities via late fusion. Figure 2.3 describes hybrid fusion.

Let's first look at approaches that are using Hybrid Fusion to combine only text-based features and then detect fake news. Oluwafemi Oriola [50] proposed an approach in 2021 to investigate unigram-based n-gram, skip-gram-based word2vec, and Latent Dirichlet Allocation (LDA)-based topic modeling, in order to better focus on the word embeddings and context of the news topic, in order to increase the effectiveness of fake news detection. The LDA [7] is a topic modeling approach that is unsupervised and fundamentally categorizes material into different subjects. The topic is expressed as the likelihood that different terms will be contained in it. The words in the text that correspond to the terms in the given subject are categorized under that topic. The Vertical Stack Approach was used to stack the Term-Frequency Inverse Document Frequency (TFIDF) weighted n-gram model and TFIDF weighted topic model. By dividing our training data into n-folds and setting one fold aside, we may train an ensemble of multiple models into that training set and obtain the results for the set-aside block using the vertical stacking strategy, which is similar to the k-fold method. The final findings are loaded into a meta classifier [8] after

[7] https://towardsdatascience.com/latent-dirichlet-allocation-lda-9d1cd064ffa2
[8] https://supunsetunga.medium.com/stacking-in-machine-learning-357db1cfc3a

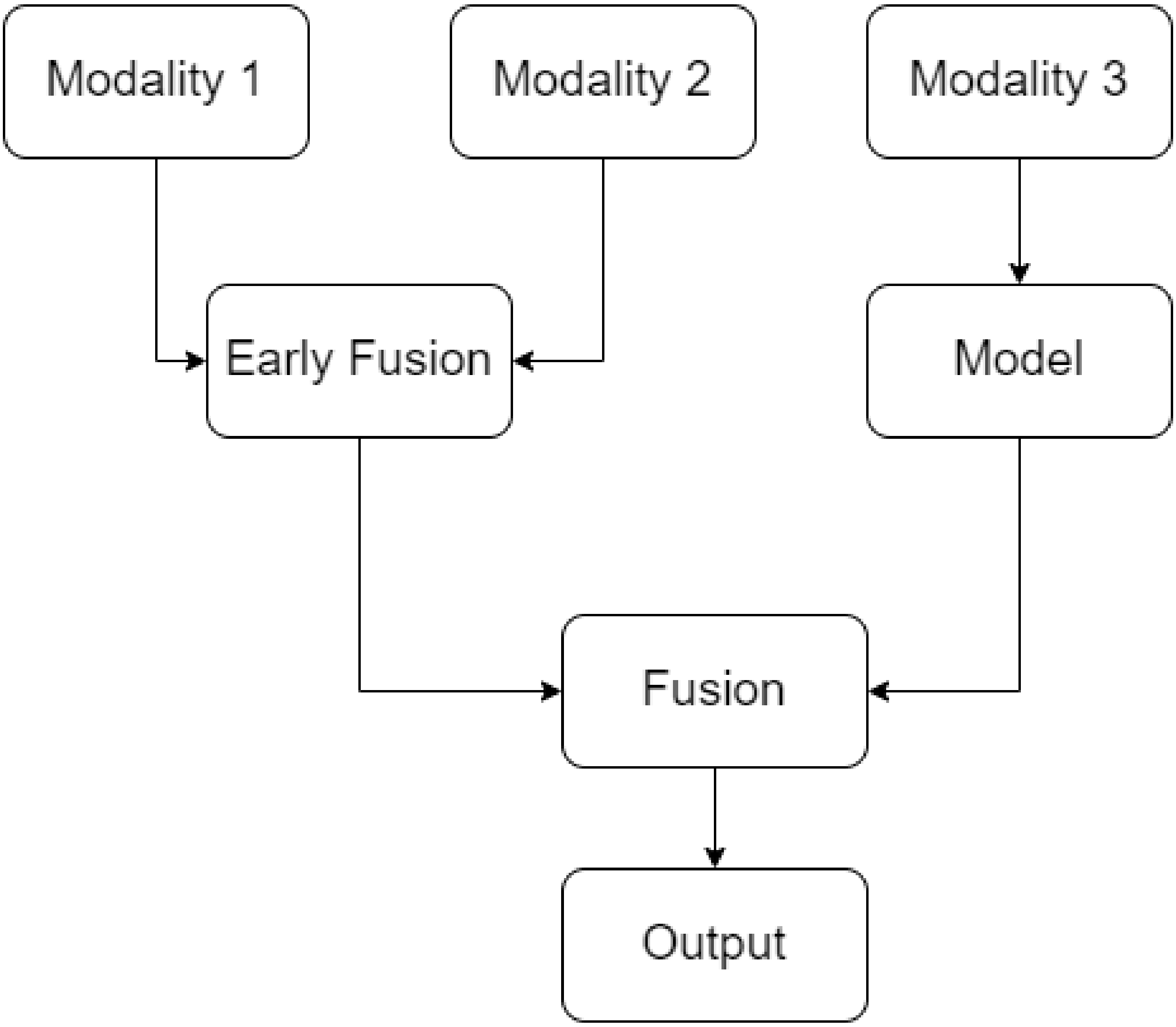

Figure 2.3: Hybrid Fusion

we repeat the process n times. Then that particular model was fused with the word2vec model with matrix multiplication, and finally, logistic regression was used to classify the news into fake or real. Another hybrid fusion approach was proposed by Deepak Kumar Jain et al.[51] in 2022 for taking into account two types of features. The features were context-based features and user-based features. The context-based features were POS (Part of Speech) tags, Word vectors, Question marks, Capital ratio, Word count, etc., and the user-based features included Tweet count, Follow ratio, Listed count, Age, and verified status. The experiment was done on PHEME [52] and a publically available twitter dataset. The article used a Hierarchical attention module for context-based features to get a feature vector by feature fusion. One more approach is based on a couple of new things which included removing noise from input modalities and looking at the interaction between different modalities at an early stage to get better information from both the text and image modalities. Lainwei Wu and Yuan Rao [53] suggested a different method for identifying false news in 2020. They attempted to combine Twitter posts and comments and create a cross-semantic fusion network that would attempt to learn the features that take into account two distinct aspects of a post-comment combination. The first is an emotional correlation, while the second is a semantic correlation. For instance, if the post claims that Biden disclosed some secret names that could endanger US security, then words like name, secret, etc. in the comments will be semantically correlated with the semantic correlation and negative words like bad, unacceptable, etc. will be emotionally correlated with the semantic correlation. Before forecasting the false news, they fused comments, correlations, posts, and correlations. The fused data were then sent via Gated Adaptive Interaction Networks (GAIN), followed by self-attention networks (separate for post-correlation, and comment-correlation). Now let's look at some approaches which took text and images and combined them to get the result of fake news detection. In 2017, Zhewei Jin et al.[54] suggested a hybrid fusion-based technique for detecting false news. They made advantage of the Weibo and Twitter dataset that was available to the public. A tweet's content and social context are combined using an LSTM [34] The visual information obtained from a deep CNN that has previously been trained is then integrated with the joint representation. The output of the LSTM is utilized in coordination with the picture data. Haizou Wang et al [55] in 2022 proposed a hybrid fusion-based fake news detection in which five kinds of features were extracted from text, image, and user information. From text embedded within the image, from article text, from user information, image-text correlation features extracted using SiameseNet [56]. They also took into account the embedded text unlike many other approaches, and VGG-19[33] models were used for various features from images and text. A hybrid model for using the information between text and visual elements was put out by Yi Lang et al. [57] in 2022. To reduce noise and enhance the fusion quality of many modalities, they adopted a deep auto-encoder. Early fusion is used to combine text and picture features, while late fusion is used to identify false news by passing all three (text, image, and fused features) through the MLP-softmax function.

2.6 Evolutionary Computation based Fake News Detection using Fusion

To the best of the author's knowledge, no survey work discussed the application of evolutionary algorithms to maximize feature extraction from various modalities. However, in 2020 [58], Priyanshi Shah and Ziad Kobti suggested a technique that entailed collecting elements from text and photos before running them through a cultural algorithm module to improve them and identify the false news. To extract text characteristics, they employed sentiment analysis. They employed Discrete Wavelet Transform segmentation to extract features from the picture. For the cultural algorithm, they have employed two Belief spaces: Situational Knowledge and Normative Knowledge. One approach described in 2021 did use only textual features, fused them, and used a heuristic algorithm to optimize the features. They used various text features from tweets and FakeNewsNet data-set such as username, URL, source, etc, and some other meta-data of the tweet post. They achieved 87.49% test accuracy which they claimed to be 9.56 % better than the state-of-the-art.

Chapter 3

Fake News Detection using Hybrid CNN-RNN Approach

3.1 Introduction to the Base Paper

The work here consists of a reproduction of the base paper [3] and tries to improve the results using two variations of a well-known approach of the model Transformers [59]. The paper also discussed some other data-specific approaches later. The Paper has used two publically available datasets, namely ISOT, and FAKES [2]. The core of the paper is that with a very simple model which consists of stacking of CNN and RNN, then can achieve good accuracy for both the dataset for the fake news detection task.

3.1.1 Dataset Description

Table 3.1 shows the data description of the FAKES dataset. Table 3.2 shows the description of the ISOT dataset. From both datasets, the article content feature was used to detect fake news. Google Colab with GPU runtime was used for carrying out all the experiments. Table 3.3 shows real and fake news instances of the datasets.

Table 3.1: FA-KES Dataset Description

Feature Name	Detail
ID	Article ID
Title	Article Title
News Content	Content of the Article
Article Source	Source of the article
Article Date	Date of publishing the article
Location	Location of the event
Label	Fake or not (0 or 1)

Table 3.2: ISOT Dataset Description

Feature Name	Details
Article Title	Title of the Article
Article Text	Text of the Article
Source	Source of the Article (News/Social Media Post)
Date	Date on which it was published

Table 3.3: Dataset Category Details

Dataset → Category	Fake	Real
ISOT	21417	23481
FAKES	376	428

3.2 Methodology Details

3.2.1 PreProcessing

First things first, everything was converted to lowercase so that "king" and "KING" do not stay different. Then we removed punctuations like ', ", ', etc. and we also removed URLs. We also removed the newline character that is present in the string when we read any particular line in any programming language. We also then removed stopwords from the article text, for example, articles, words like "not", "is", "was", etc. all the stopwords. We also did lemmatization of the words. For example, if "Going" was coming in a sentence, then we changed it to "Go". This helped us focus more on the meaning of the word and less on the actual meaning that it was focusing upon.

3.2.2 Tokenization and Word Embeddings

After Pre-processing, we did Tokenization of the dataset. The NLTK tokenizer was used. Tokenizer did give a unique integer to each unique word present in the corpus. After tokenization Word Embeddings were used to generate meaningful representations of the sentences. GloVe [1] was used to generate word embeddings, GloVe works by constructing co-occurrence matrix of the given corpus. Word Embeddings are generated such that the words retain their meaning and each word is represented into a vector of fixed length, in our case, it's 100. For that, a publically available file, "glove.6b.100d.txt" was used. This file contains data of 6 billion words, each having a 100-dimensional vector representation.

3.2.3 Neural Network Model

The entire architecture is shown in Figure 3.1. After we got the word embeddings, we passed them into one-dimensional CNNs. CNN filter size was 5 and a total of 128 filters were used. A highly well-liked neural network for pictures is a convolutional neural network. It has a filter with a defined size that iterates around the input and multiplies the filter with the input data stride by stride to produce a feature map. When processing text, one-dimensional convolution is employed, where a fixed-size input array (filter) repeatedly iterates through the training set of data to produce a feature map. During the training phase, the fixed filters' values are discovered. We use Rectified Linear Unit Activation Function here. The feature vectors are down-sampled using the MaxPooling layer.

[1] https://nlp.stanford.edu/projects/glove/

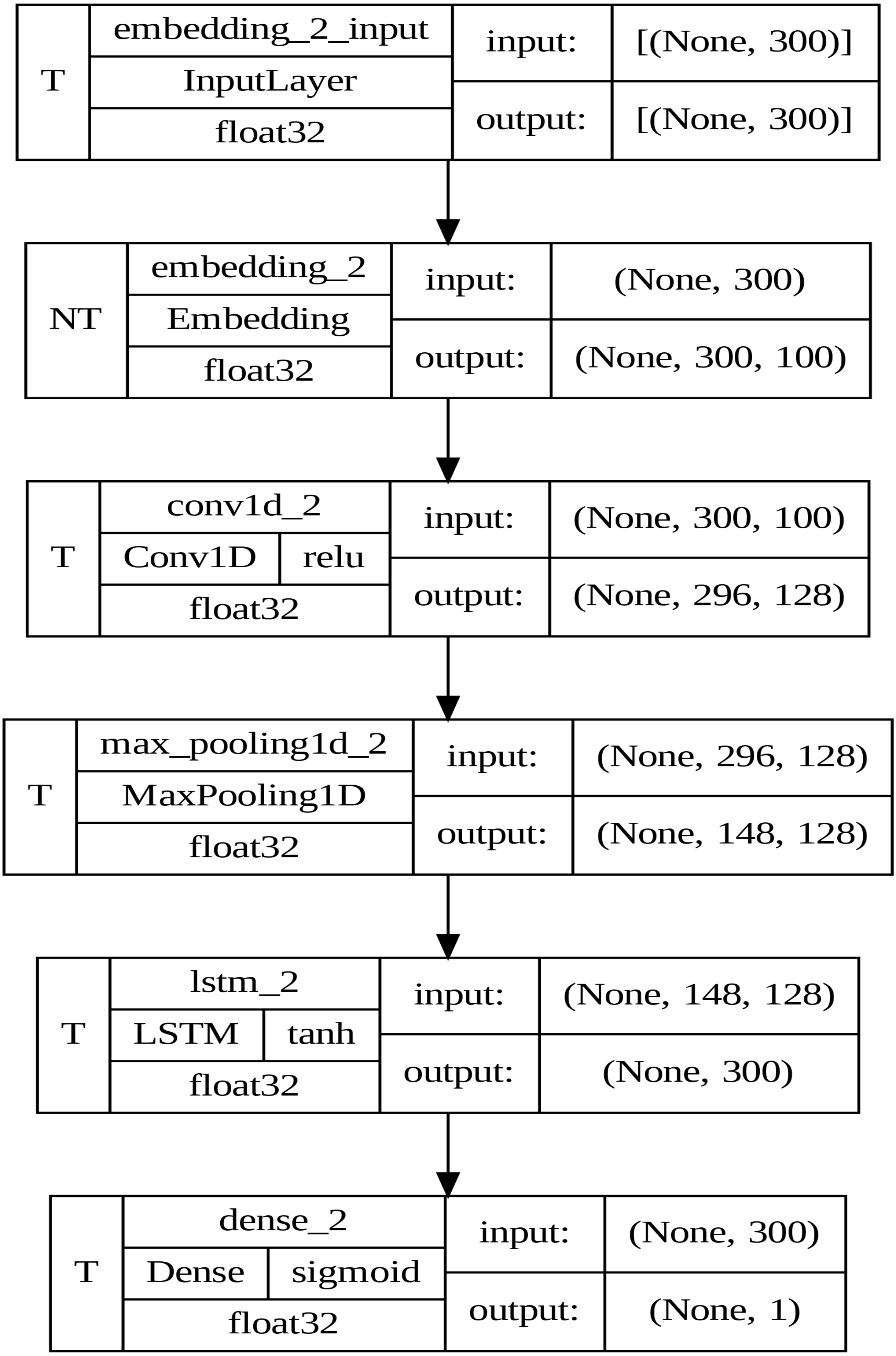

Figure 3.1: Model Architecture

It facilitates dimensionality reduction and lowers the number of parameters that must be learned for a particular model. For that, we've chosen a window size of 2. Then, for jobs involving sequence analysis, we input our model into Long Short-Term Memory (LSTM) Networks. A specific kind of recurrent neural network-based model is the LSTM. The weights and three different learning gates make up an LSTM cell. At each time step, there is an input gate for the current input, an output gate for forecasting values, and a forget gate for erasing irrelevant data.

Because they aid in the mitigation of vanishing gradients, LSTM is superior to RNN. In LSTM, we used 32 units (output dimension). The output is then passed into a thick layer of size 1, where it is determined whether or not it is fake news. This uses the sigmoid activation function.

A total of 1000 training epochs were employed, using a batch size of 64, the binary cross-entropy loss function, and the Adam optimizer. The base paper mentioned that the 10 epochs are enough, although the authors here observed that we get optimal output with 1000 epochs only.

3.3 Results and Analysis

Training data were evaluated using a validation split of 0.2, meaning that 20% of the data were used for testing and 80% for training. A total of five different trials of 5-fold cross-validation were done and results were averaged. The findings for both datasets are displayed in the table 3.4 after 1000 epochs. As we can see, we get a decent result in the ISOT dataset whereas a fairly good accuracy of 60% in the FA-KES dataset. For the ISOT dataset, the results are quite promising with more than 99% accuracy, and 99% F1 Score. For the FA-KES dataset, we are getting a fairly good F1 Score and reasonable accuracy. The same model was used for both datasets. Figure 3.2 shows one instance of the training for the FAKES dataset[2]. The validation accuracy is higher than 60 % because one of the folds might get higher accuracy, although when averaged out, it reduced to around 60 %. The same for ISOT is shown in Figure 3.3[3]. Table 3.5 and 3.4 show results original and reproduced respectively.

[2]The image is blurry because it is taken as a screenshot.
[3]The image is blurry because it is taken as a screenshot.

Table 3.4: Reproduced Results

Dataset ↓ **Performance Metric →**	**Validation** **Accuracy**	**Validation** **F1Score**
FA-KES	59.27 %	57.61 %
ISOT	99.69 %	99.69 %

Table 3.5: Original Results

Dataset ↓ **Performance Metric →**	**Validation** **Accuracy**	**Validation** **F1Score**
FA-KES	60.00 %	59.00 %
ISOT	99.00 %	99.00 %

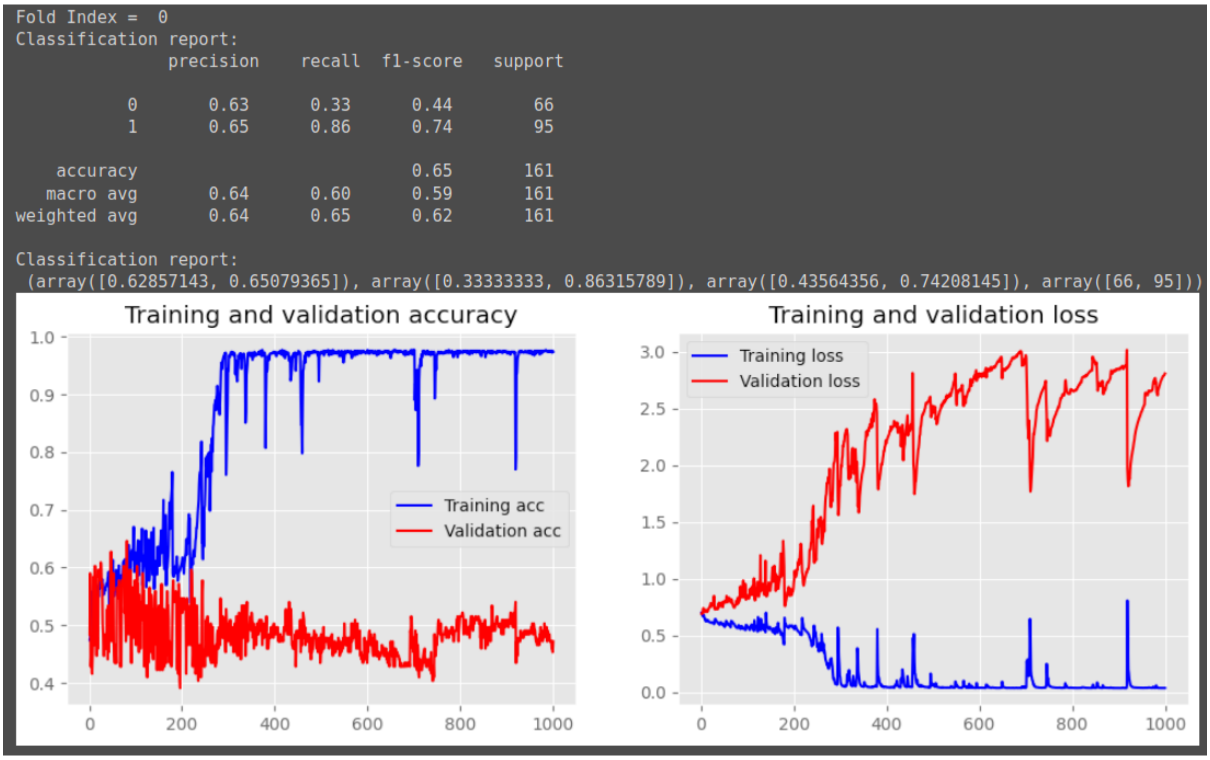

Figure 3.2: Training Loss, Training Accuracy, Validation Loss, Validation Accuracy for FAKES dataset

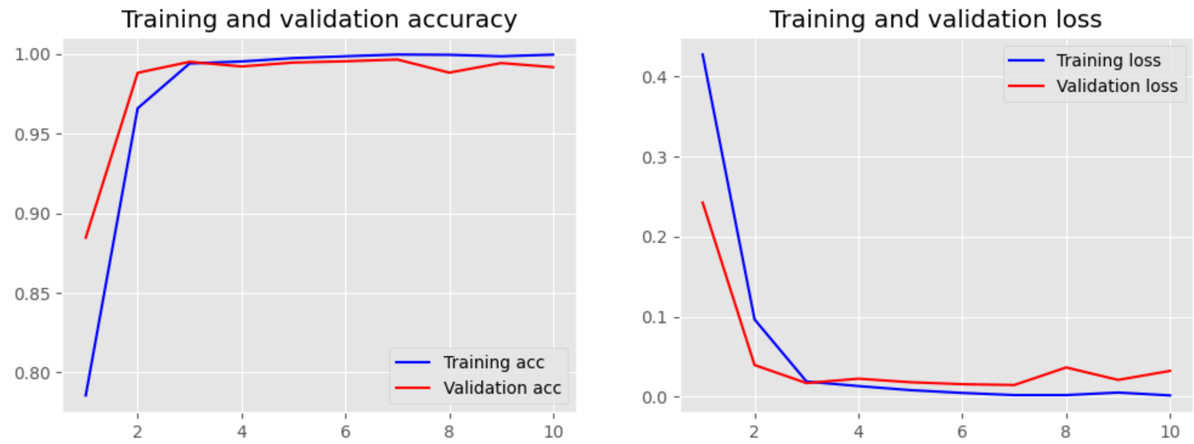

Figure 3.3: Training Loss, Training Accuracy, Validation Loss, Validation Accuracy for ISOT dataset

Chapter 4

Improvement Efforts

The ISOT dataset had an accuracy of more than 99% so that's not required to be improved. Here we discuss ideas that were used to improve the FAKES dataset which had an accuracy of 60 %.

4.1 Transformers

We tried to use the Transformers [59] to better our accuracy, although we failed. Before going into details, here we present some introduction to transformers. Recurrent Neural Networks are one of the early-age models that people used to deal with time series data. Although traditional RNN suffered from vanishing gradient problems when dealing with texts, or sequences which were too long[1]. To solve this, LSTM was introduced. LSTM had a memory that it maintained to keep track of important words in a long sequence which was important. The LSTM, however, had an issue, it took longer time to train and was prone to overfitting. The Encoder-Decoder networks that use these networks are suffering from the bottleneck, in that only one particular context vector, which is derived by the Encoder network is fixed in length, and thus is not able to get better results when passed to the decoder network. This issue was solved by the attention mechanism. The attention mechanism tried to focus on different parts of the input sequence, and the decoder paid "attention" to various parts of the input sequence, and not only on one fixed-length context vector. The transformer architecture [59] demonstrated that RNN/LSTM type of networks are not needed at all, and using only the "attention" mechanism, better

[1] https://www.analyticsvidhya.com/blog/2021/07/lets-understand-the-problems-with-recurrent-neural-networks/

results can be achieved. Transformers use “self-attention” which allowed each position in the sequence to attend to all other positions. Transformers can process various input sequences parallel, and they are better at scalability as well. The Transformers however failed on the fakes dataset, and we didn’t get any visible improvement. This could be because to get better performance on transformers, we need a large amount of data, and the FAKES dataset had only 804 rows.

4.2 Sentence Transformers

BERT encodes one word in natural language processing tasks. Although the authors tried to experiment with sentence transformers, which encode entire sentences with the help of transformer architecture. Sentence transformer was first introduced by Nils Reimers and Iryna Gurevych in 2019[2]. They work by taking pair of sentences and training them on a single BERT one after another. Now, the difference of vectors of both sentences is taken, word by word. The two sentences are concatenated and the difference vector of the two sentences is also concatenated. Then this vector is trained with FFNN to get the similarity score and is matched with the original scores. The similarity score here is of three types. The first sentence here is Premise and the second Hypothesis. The Sentence Transformers are depicted in Figure 4.1[3] and Figure 4.2.[4]

1. Entailment: The Premise Suggests the Hypothesis.
2. Neutral: The Premise and Hypothesis could both be true, but they are not necessarily related.
3. Contradiction: The Premise and Hypothesis contradict each other.

[2] https://www.pinecone.io/learn/sentence-embeddings/
[3] Figure taken from https://www.pinecone.io/learn/sentence-embeddings/.
[4] Figure taken from https://www.pinecone.io/learn/sentence-embeddings/.

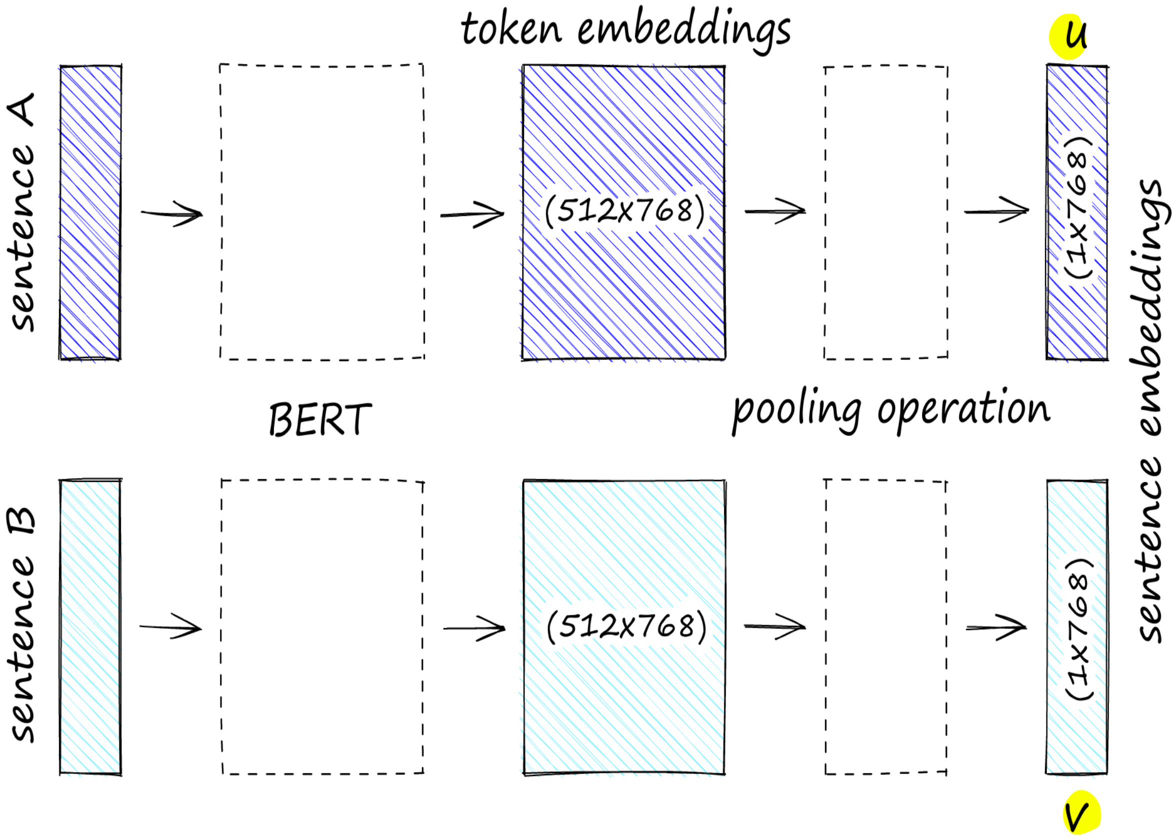

Figure 4.1: Sentence Transformers - Generating tokens for sentences

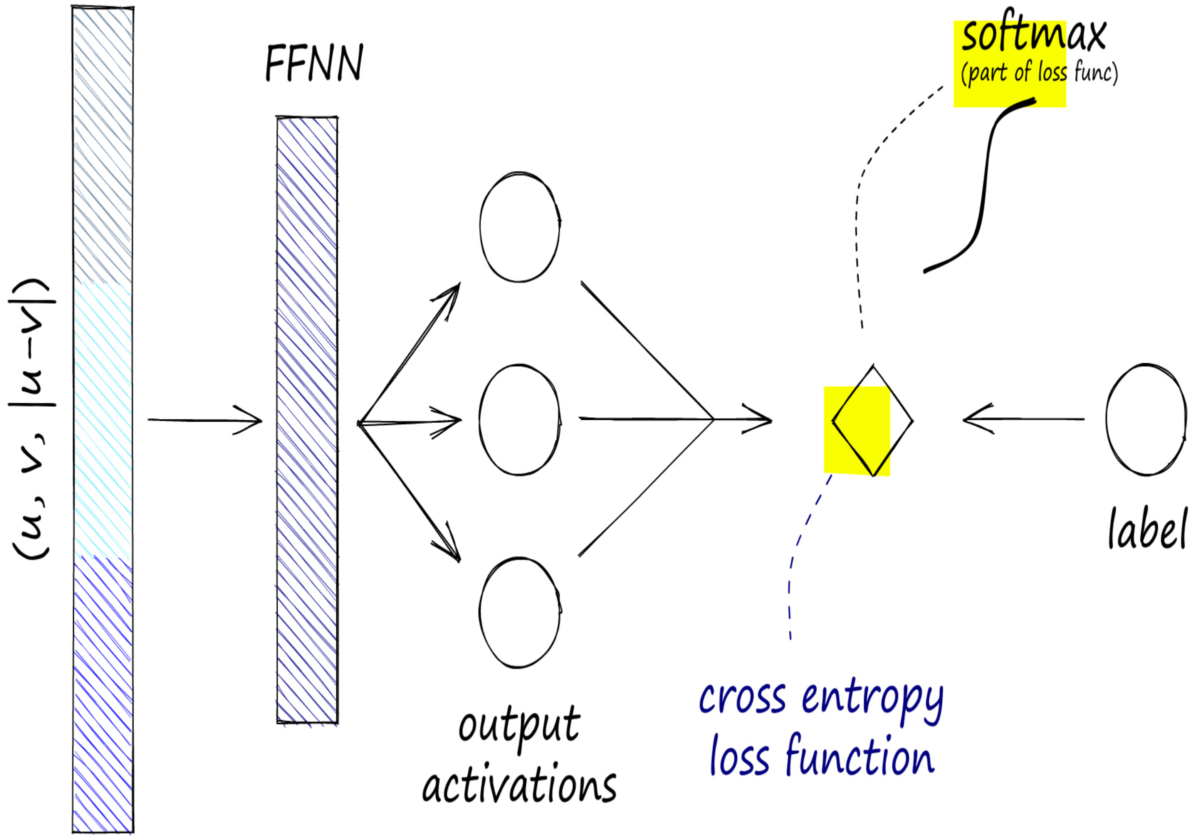

Figure 4.2: Sentence Transformers - Comparison between sentences and Training

4.3 Sentiment Analysis-based Training

This approach involved transforming data into time series data. In this approach, `textblob` library was used for sentiment analysis. At the start, the "article content" of the FA-KES dataset is used. After getting all the data that was in the content, we tried to split each training example into a set of sentences. For example, if there was a news article that involved 10 sentences in total, then that article was split into a set of 9 to 10 sentences, and then for each news article, we applied `textblob` library-based sentiment analysis over each sentence of the news article and generated a sentiment score of the same. For example, the following is one article from the FA-KES dataset.

> "Wed 05 Apr 2017 Syria attack symptoms consistent with nerve agent use WHO. Victims of a suspected chemical attack in Syria appeared to show symptoms consistent with reaction to a nerve agent the World Health Organization said on Wednesday. "Some cases appear to show additional signs consistent with exposure to organophosphorus chemicals a category

> of chemicals that includes nerve agents" WHO said in a statement putting the death toll at at least 70. The United States has said the deaths were caused by sarin nerve gas dropped by Syrian aircraft. Russia has said it believes poison gas had leaked from a rebel chemical weapons depot struck by Syrian bombs. Sarin is an organophosporus compound and a nerve agent. Chlorine and mustard gas which are also believed to have been used in the past in Syria are not. A Russian Defence Ministry spokesman did not say what agent was used in the attack but said the rebels had used the same chemical weapons in Aleppo last year. The WHO said it was likely that some kind of chemical was used in the attack because sufferers had no apparent external injuries and died from a rapid onset of similar symptoms including acute respiratory distress. It said its experts in Turkey were giving guidance to overwhelmed health workers in Idlib on the diagnosis and treatment of patients and medicines such as Atropine an antidote for some types of chemical exposure and steroids for symptomatic treatment had been sent. A U.N. Commission of Inquiry into human rights in Syria has previously said forces loyal to Syrian President Bashar al-Assad have used lethal chlorine gas on multiple occasions. Hundreds of civilians died in a sarin gas attack in Ghouta on the outskirts of Damascus in August 2013. Assads government has always denied responsibility for that attack. Syria agreed to destroy its chemical weapons in 2013 under a deal brokered by Moscow and Washington. But Russia a Syrian ally and China have repeatedly vetoed any United Nations move to sanction Assad or refer the situation in Syria to the International Criminal Court. "These types of weapons are banned by international law because they represent an intolerable barbarism" Peter Salama Executive Director of the WHO Health Emergencies Programme said in the WHO statement. - REUTERS 0"

Now this one news article is converted into a few sentences, as displayed below in Table 4.1. The third column titled "Sentiment Score by `textblob`" is the sentiment scores generated by a library named `textblob`[5]. TextBlob works by applying Naive Bayes

[5] https://textblob.readthedocs.io/en/dev/

Table 4.1: Example of Statements out of a news article

Serial No.	Sentence	Sentiment Score by `textblob`
1	Wed 05 Apr 2017 Syria attack symptoms consistent with nerve agent use WHO.	0.25
2	Victims of a suspected chemical attack in Syria appeared to show symptoms consistent with reaction to a nerve agent the World Health Organization said on Wednesday."Some cases appear to show additional signs consistent with exposure to organophosphorus chemicals a category of chemicals that includes nerve agents" WHO said in a statement putting the death toll at at least 70.	0.067
3	The United States has said the deaths were caused by sarin nerve gas dropped by Syrian aircraft.	0.0
4	Russia has said it believes poison gas had leaked from a rebel chemical weapons depot struck by Syrian bombs.	0.0
5	Sarin is an organophosporus compound and a nerve agent.	0.0
6	Chlorine and mustard gas which are also believed to have been used in the past in Syria are not.	-0.25
7	A Russian Defence Ministry spokesman did not say what agent was used in the attack but said the rebels had used the same chemical weapons in Aleppo last year.	0.0
8	The WHO said it was likely that some kind of chemical was used in the attack because sufferers had no apparent external injuries and died from a rapid onset of similar symptoms including acute respiratory distress.	0.196
9	It said its experts in Turkey were giving guidance to overwhelmed health workers in Idlib on the diagnosis and treatment of patients and medicines such as Atropine an antidote for some types of chemical exposure and steroids for symptomatic treatment had been sent.	0.0
10	A U.N. Commission of Inquiry into human rights in Syria has previously said forces loyal to Syrian President Bashar al-Assad have used lethal chlorine gas on multiple occasions.	0.042
11	Hundreds of civilians died in a sarin gas attack in Ghouta on the outskirts of Damascus in August 2013.	0.0
12	Assads government has always denied responsibility for that attack.	0.0
13	Syria agreed to destroy its chemical weapons in 2013 under a deal brokered by Moscow and Washington.	-0.2
14	But Russia a Syrian ally and China have repeatedly vetoed any United Nations move to sanction Assad or refer the situation in Syria to the International Criminal Court. "These types of weapons are banned by international law because they represent an intolerable barbarism" Peter Salama Executive Director of the WHO Health Emergencies Programme said in the WHO statement. - REUTERS.	-0.133

classifier on a large corpus of a labeled dataset. It also takes into account already labeled words for detecting sentiment in informal posts such as social media posts.

After converting the dataset into sentiment-based data, we feed each list of sentiment scores as one training data into the model. Before feeding into the model, we make the length of all the training data equal. Some news articles have 15 sentences, some have 14, etc. We average the length of all the articles and chop off some articles if needed and pad some articles with zero. For the FA-KES dataset, the average length came out to be 11 sentences per article. The model here is very simple. We just feed our data into a BiLSTM [60] with 64 units and then a Dense layer to classify the news. The total trainable parameters of the model are 8769.

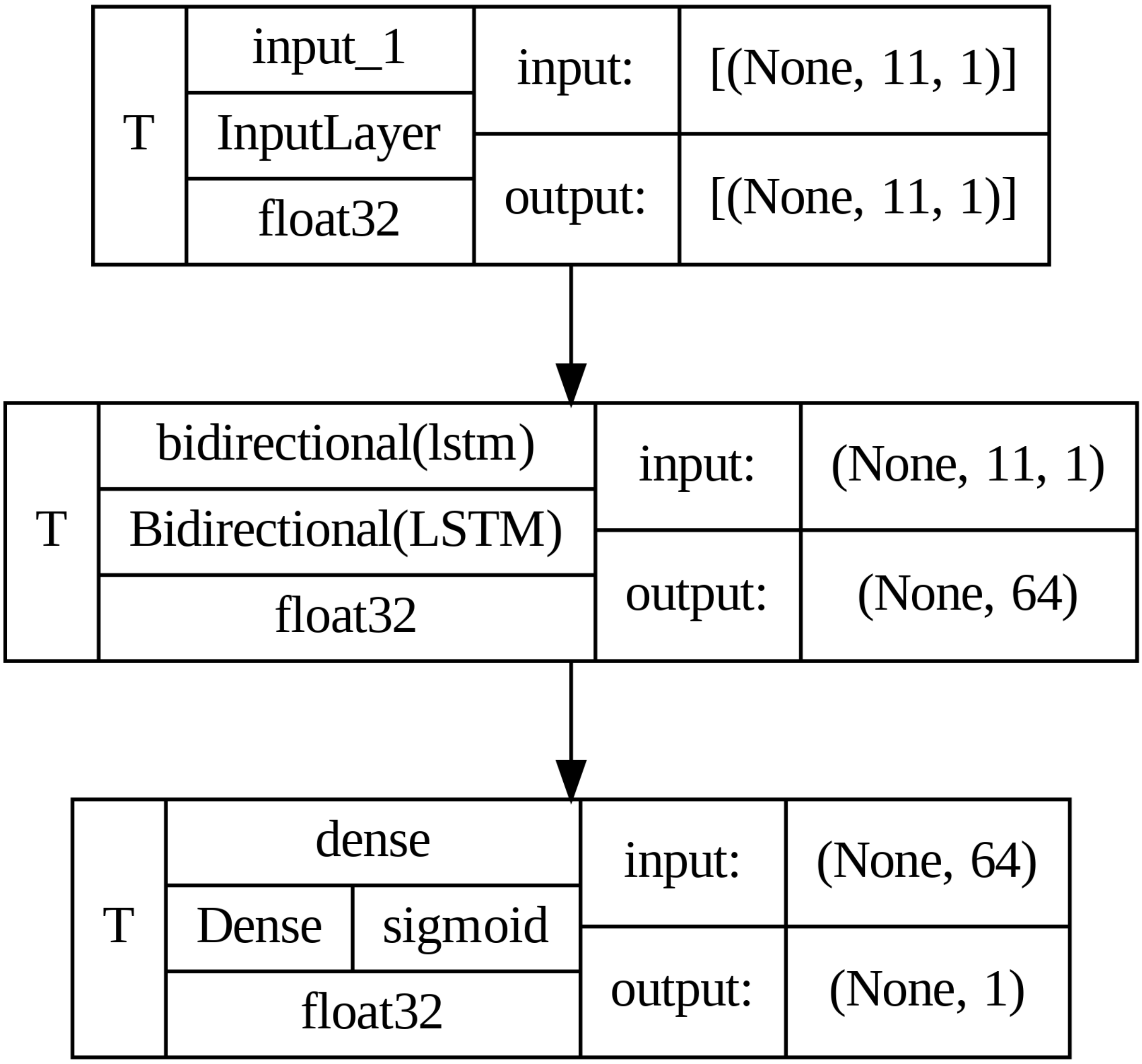

Figure 4.3: Sentiment Model Architecture

In addition to this, two additional parameters were added to the training process.

1. Sentiment score of the title of the article.
2. Source credibility score (score between 0 and 1 for a news publisher where 1 indicates that the news publisher is most reliable, and 0 indicates least reliable).

In addition to splitting the paragraph into sentences and using text blob-based sentiment scores, the authors also used the sentiment of the title and the credibility score of the news publisher as the feature. The credibility score of the news publishers was generated using chatGPT[6] Generative AI model. The reason behind using chatGPT was the unavailability of a single news-rating agency that had all the news publishers which were there in the FAKES dataset. The news source and title sentiment were used as another feature and hence they increased the length of one data point from 11 to 13. However, this approach backfired, and got accuracy decreased from baseline rather than increased. The authors observed that this particular model could not learn one particular type of dataset. Sometimes it learned only real news, and sometimes fake. The reason could be that the LSTM is not the right model for concatenated data - sentiment scores + title sentiment score + source credibility score.

4.4 3-4-3 Neurons based Model

One more approach that the authors have tried here, in order to improve the accuracy, and F1 score on the FAKES dataset is to replace the classifier in the base paper with a little more complex structure. The authors have replaced the single dense layer-based classification with a model that is having a "3 - 4 - 3" structure. That is after the CNN-RNN stack, a 3-neuron hidden layer follows, and 4-neuron and 3-neuron hidden Dense layers follow the initial 3-neuron layer. The complete model is shown in Figure 4.4. The activation functions are kept linear for the newly added layers. The reason behind not keeping the ReLU activation function was such that the experiments that were conducted with "ReLU" were showing the effect of not learning over one of the types of news, either fake or real. The results of this approach were matching with the base paper with a little gap between accuracy and F1 score when averaged over 5 runs with the same dataset, and the same fivefold cross-validation split. It was observed that the model got to the baseline accuracy, although it never reached more than the baseline. This indicates that the model is not able to learn more about the veracity of the dataset. The reasons could

[6] https://chat.openai.com/

be insufficient data (as it has only 804 rows), or the baseline model itself is not adequate enough to learn.

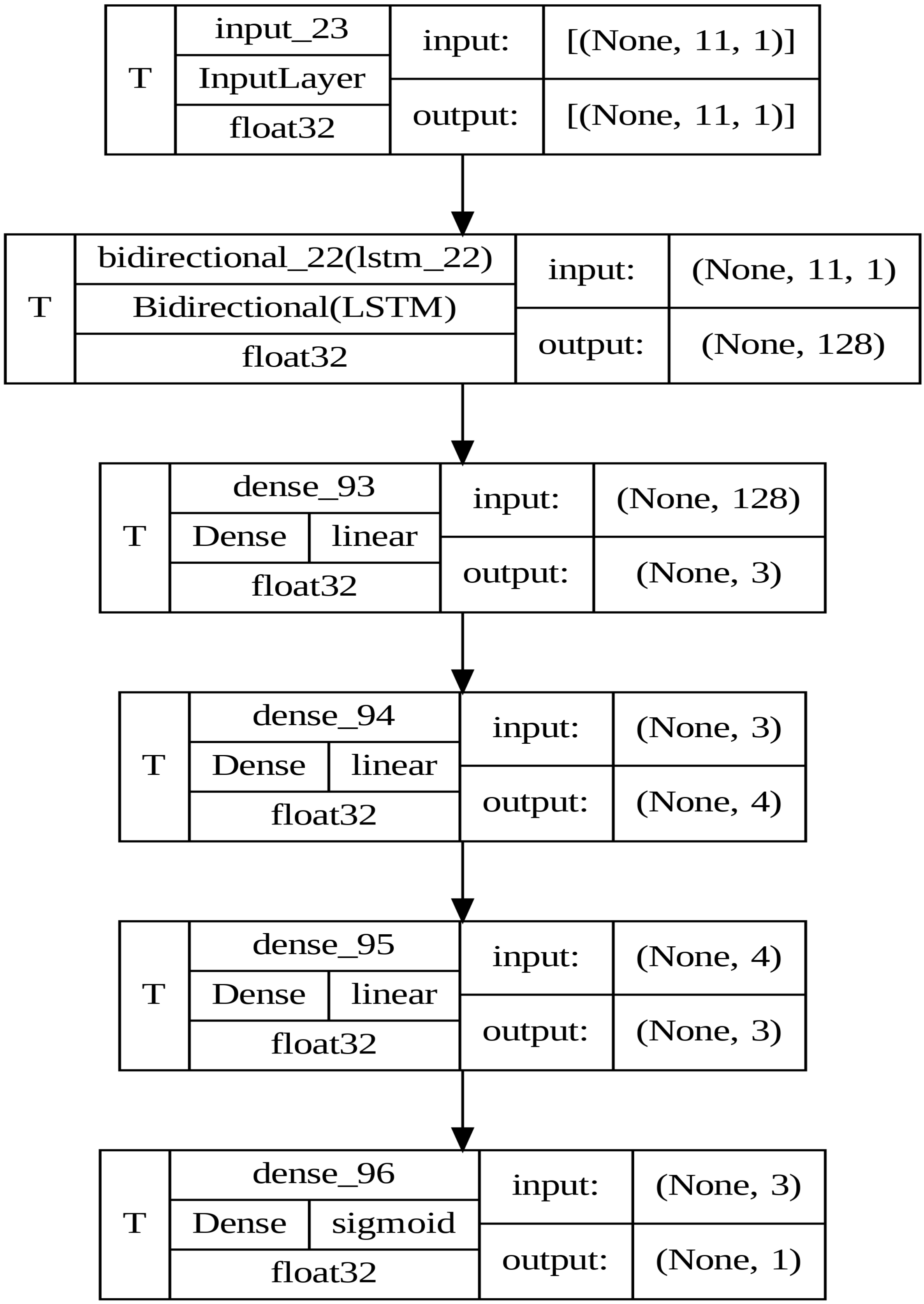

Figure 4.4: 3-4-3 Model Architecture with Sentiment Analysis-based Approach

Along with testing the 3-4-3 approach on the baseline model, the authors also tested the same on the sentiment analysis-based model described in section 4.3. The last dense layer in the sentiment analysis-based model was replaced 3-4-3 model as it was done in the baseline as well. There too authors observed that it could not improve the results of the paper.

4.5 Summary of Improvement Efforts

The improvement efforts first focused on using different modern-age transformer-based models and the authors observed that it could not improve the accuracy. The dataset being too small could be one of the reasons why LLM-based training didn't help here. The next set of approaches was focused on looking at datasets and making dataset-specific changes that could lead to better accuracy and F1 score. However, those approaches broadly could only match the baseline, and could not improve it. The authors believe that the dataset itself could be the reason behind the failure to improve accuracy. As the baseline authors also claim that when the model was trained on ISOT and got 99 % accuracy, and then when that same model was applied to FAKES dataset, the accuracy was reduced to 50 %. That indicates that the patterns are generally observed and hence the DL models can learn those patterns, although this dataset is clustered around only one event, which is the Syrian war 2015[7] might be the reason why we cannot improve the accuracy. The summary of Improvement efforts is described in Table 4.2

[7] https://en.wikipedia.org/wiki/Syrian_civil_war

Table 4.2: Comparison between Improvement Efforts

Performance Metric → **Details of Methods ↓**	**Validation Accuracy**	**Validation F1Score**	**Trainable Parameters of the Model**
Base Paper	60.00 %	**59.00 %**	5,79,229
Base Paper Reproduction	59.28 %	57.61 %	5,79,229
Base Paper with 3-4-3 Model	**60.34 %**	56.35 %	5,79,866
Sentiment Analysis based Training (Only text)	57.16 %	52.02 %	33,921
Sentiment Analysis based Training (Only text, 3-4-3 model)	56.62 %	52.78 %	34,214
Sentiment Analysis based Training (Text + Title sentiment expanded)	56.92 %	53.03 %	33,921
Sentiment Analysis based Training (Text + Title sentiment + Source score, 3-4-3 model)	47.02 %	30.21 %	34,214

Chapter 5

Conclusion and Future Work

Here, we conducted a systematic survey of Fake News Detection with Deep Learning Techniques that involved the Fusion of multiple modalities, such as text, audio, video, etc. We also included Evolutionary algorithm-based studies. We also looked at numerous efforts of improving the validation accuracy of the FAKES dataset, which we couldn't do, and for which we have mentioned a detailed analysis in Section 4.5. Further research on the survey part could focus on using AI to detect AI-generated fake news. If people are writing fake news articles using chatGPT, or BARD[1]. People could also use AI-generated images along with AI-generated texts, and a combination of that with manual text, and images.

The improvement efforts on the FAKES dataset could use Deep Learning models that can better feature engineer any particular dataset then which might improve the results, although that model could only work on FAKES then, and it would fail to generalize on other fake news datasets.

[1] https://bard.google.com/

Bibliography

[1] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake news detection on social media: A data mining perspective,” *SIGKDD Explor. Newsl.*, vol. 19, p. 22–36, sep 2017.

[2] F. K. Abu Salem, R. Al Feel, S. Elbassuoni, M. Jaber, and M. Farah, “Fa-kes: A fake news dataset around the syrian war,” *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 13, pp. 573–582, Jul. 2019.

[3] J. A. Nasir, O. S. Khan, and I. Varlamis, “Fake news detection: A hybrid cnn-rnn based deep learning approach,” *International Journal of Information Management Data Insights*, vol. 1, no. 1, p. 100007, 2021.

[4] F. Alam, S. Cresci, T. Chakraborty, F. Silvestri, D. Dimitrov, G. D. S. Martino, S. Shaar, H. Firooz, and P. Nakov, “A survey on multimodal disinformation detection,” *arXiv preprint arXiv:2103.12541*, 2021.

[5] R. Tolosana, R. Vera-Rodríguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, “Deepfakes and beyond: A survey of face manipulation and fake detection,” *Inf. Fusion*, vol. 64, pp. 131–148, 2020.

[6] C. Raj and P. Meel, “A review of web infodemic analysis and detection trends across multi-modalities using deep neural networks,” *arXiv preprint arXiv:2112.00803*, 2021.

[7] T. H. Afridi, A. Alam, M. N. Khan, J. Khan, and Y.-K. Lee, “A multimodal memes classification: A survey and open research issues,” in *The Proceedings of the Third International Conference on Smart City Applications*, pp. 1451–1466, Springer, 2021.

[8] S. Abdali, "Multi-modal misinformation detection: Approaches, challenges and opportunities," *arXiv preprint arXiv:2203.13883*, 2022.

[9] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Information fusion*, vol. 14, no. 1, pp. 28–44, 2013.

[10] W. Y. Wang, ""liar, liar pants on fire": A new benchmark dataset for fake news detection," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, (Vancouver, Canada), pp. 422–426, Association for Computational Linguistics, July 2017.

[11] S. Raza and C. Ding, "Fake news detection based on news content and social contexts: a transformer-based approach," *International Journal of Data Science and Analytics*, vol. 13, no. 4, pp. 335–362, 2022.

[12] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," *arXiv preprint arXiv:1910.13461*, 2019.

[13] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[14] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, *et al.*, "Language models are unsupervised multitask learners," *OpenAI blog*, vol. 1, no. 8, p. 9, 2019.

[15] S. Rezayi, S. Soleymani, H. R. Arabnia, and S. Li, "Socially aware multimodal deep neural networks for fake news classification," in *2021 IEEE 4th International Conference on Multimedia Information Processing and Retrieval (MIPR)*, pp. 253–259, IEEE, 2021.

[16] A. Zubiaga, M. Liakata, R. Procter, G. Wong Sak Hoi, and P. Tolmie, "Analysing how people orient to and spread rumours in social media by looking at conversational threads," *PloS one*, vol. 11, no. 3, p. e0150989, 2016.

[17] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

[18] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," *Advances in neural information processing systems*, vol. 26, 2013.

[19] J. Pennington, R. Socher, and C. D. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543, 2014.

[20] J. Sarzynska-Wawer, A. Wawer, A. Pawlak, J. Szymanowska, I. Stefaniak, M. Jarkiewicz, and L. Okruszek, "Detecting formal thought disorder by deep contextualized word representations," *Psychiatry Research*, vol. 304, p. 114135, 2021.

[21] J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," *arXiv preprint arXiv:1801.06146*, 2018.

[22] Y. Zhang and B. Wallace, "A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification," *arXiv preprint arXiv:1510.03820*, 2015.

[23] Y. Dou, K. Shu, C. Xia, P. S. Yu, and L. Sun, "User preference-aware fake news detection," in *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 2051–2055, 2021.

[24] C. Guo, J. Cao, X. Zhang, K. Shu, and M. Yu, "Exploiting emotions for fake news detection on social media," *arXiv preprint arXiv:1903.01728*, 2019.

[25] N. Aslam, I. Ullah Khan, F. S. Alotaibi, L. A. Aldaej, and A. K. Aldubaikil, "Fake detect: A deep learning ensemble model for fake news detection," *complexity*, vol. 2021, 2021.

[26] R. Kumari and A. Ekbal, "Amfb: attention based multimodal factorized bilinear pooling for multimodal fake news detection," *Expert Systems with Applications*, vol. 184, p. 115412, 2021.

[27] H. Yang, X. Zhao, D. Sun, Y. Wang, H. Zhu, C. Ma, and W. Huang, "Multi-modal fake news detection on social media with dual attention fusion networks," in *2021 IEEE Symposium on Computers and Communications (ISCC)*, pp. 1–6, IEEE, 2021.

[28] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," *Advances in neural information processing systems*, vol. 30, 2017.

[29] D. Khattar, J. S. Goud, M. Gupta, and V. Varma, "Mvae: Multimodal variational autoencoder for fake news detection," in *The world wide web conference*, pp. 2915–2921, 2019.

[30] P. Li, X. Sun, H. Yu, Y. Tian, F. Yao, and G. Xu, "Entity-oriented multi-modal alignment and fusion network for fake news detection," *IEEE Transactions on Multimedia*, 2021.

[31] P. Qi, J. Cao, X. Li, H. Liu, Q. Sheng, X. Mi, Q. He, Y. Lv, C. Guo, and Y. Yu, "Improving fake news detection by using an entity-enhanced framework to fuse diverse multimodal clues," in *Proceedings of the 29th ACM International Conference on Multimedia*, pp. 1212–1220, 2021.

[32] H. Choi and Y. Ko, "Effective fake news video detection using domain knowledge and multimodal data fusion on youtube," *Pattern Recognition Letters*, vol. 154, pp. 44–52, 2022.

[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[34] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[35] S. Jiang, X. Chen, L. Zhang, S. Chen, and H. Liu, "User-characteristic enhanced model for fake news detection in social media," in *CCF International conference on natural language processing and chinese computing*, pp. 634–646, Springer, 2019.

[36] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "Fakenewsnet: A data repository with news content, social context and spatialtemporal information for studying fake news on social media," *arXiv preprint arXiv:1809.01286*, 2018.

[37] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, "Albert: A lite bert for self-supervised learning of language representations," *arXiv preprint arXiv:1909.11942*, 2019.

[38] P. Saikia, K. Gundale, A. Jain, D. Jadeja, H. Patel, and M. Roy, "Modelling social context for fake news detection: A graph neural network based approach," in *2022 International Joint Conference on Neural Networks (IJCNN)*, pp. 01–08, IEEE, 2022.

[39] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE transactions on neural networks*, vol. 20, no. 1, pp. 61–80, 2008.

[40] X. Zhou, A. Mulay, E. Ferrara, and R. Zafarani, "Recovery: A multimodal repository for covid-19 news credibility research," in *Proceedings of the 29th ACM International Conference on Information amp; Knowledge Management*, CIKM '20, (New York, NY, USA), p. 3205–3212, Association for Computing Machinery, 2020.

[41] L. Cui and D. Lee, "Coaid: Covid-19 healthcare misinformation dataset," *ArXiv*, vol. abs/2006.00885, 2020.

[42] A. F. Hamid, N. Shiekh, N. Said, K. Ahmad, A. Gul, L. Hassan, and A. Al-Fuqaha, "Fake news detection in social media using graph neural networks and nlp techniques: A covid-19 use-case," *ArXiv*, vol. abs/2012.07517, 2020.

[43] J. Xue, Y. Wang, Y. Tian, Y. Li, L. Shi, and L. Wei, "Detecting fake news by exploring the consistency of multimodal data," *Information Processing Management*, vol. 58, no. 5, p. 102610, 2021.

[44] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer Vision – ECCV 2014* (D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, eds.), (Cham), pp. 740–755, Springer International Publishing, 2014.

[45] P. Meel and D. K. Vishwakarma, "Multi-modal fusion using fine-tuned self-attention and transfer learning for veracity analysis of web information," *ArXiv*, vol. abs/2109.12547, 2021.

[46] S. Singhal, T. Pandey, S. Mrig, R. R. Shah, and P. Kumaraguru, "Leveraging intra and inter modality relationship for multimodal fake news detection," in *Companion Proceedings of the Web Conference 2022*, pp. 726–734, 2022.

[47] Y. Zhou and S.-N. Lim, "Joint audio-visual deepfake detection," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 14800–14809, October 2021.

[48] C. X. T. Du, L. H. Duong, H. T. Trung, P. M. Tam, N. Q. V. Hung, and J. Jo, "Efficient-frequency: a hybrid visual forensic framework for facial forgery detection," in *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 707–712, 2020.

[49] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proceedings of the 36th International Conference on Machine Learning* (K. Chaudhuri and R. Salakhutdinov, eds.), vol. 97 of *Proceedings of Machine Learning Research*, pp. 6105–6114, PMLR, 09–15 Jun 2019.

[50] O. Oriola, "Exploring n-gram, word embedding and topic models for content-based fake news detection in fakenewsnet evaluation," *International Journal of Computer Applications*, vol. 176, pp. 24–29, 2020.

[51] D. Jain, A. Kumar, and A. Shrivastava, "Canardeep: a hybrid deep neural model with mixed fusion for rumour detection in social data streams," *Neural Computing and Applications*, vol. 34, pp. 1–12, 09 2022.

[52] E. Kochkina, M. Liakata, and A. Zubiaga, "PHEME dataset for Rumour Detection and Veracity Classification," 6 2018.

[53] L. Wu and Y. Rao, "Adaptive interaction fusion networks for fake news detection," *ArXiv*, vol. abs/2004.10009, 2020.

[54] Z. Jin, J. Cao, H. Guo, Y. Zhang, and J. Luo, "Multimodal fusion with recurrent neural networks for rumor detection on microblogs," in *Proceedings of the 25th ACM international conference on Multimedia*, pp. 795–816, 2017.

[55] H. Wang, S. Wang, and Y. Han, "Detecting fake news on chinese social media based on hybrid feature fusion method," *Expert Systems with Applications*, vol. 208, p. 118111, 2022.

[56] H. Ren, M. Pan, Y. Li, X. Zhou, and J. Luo, "St-siamesenet: Spatio-temporal siamese networks for human mobility signature identification," in *Proceedings of the*

26th ACM SIGKDD International Conference on Knowledge Discovery amp; Data Mining, KDD '20, (New York, NY, USA), p. 1306–1315, Association for Computing Machinery, 2020.

[57] Y. Liang, T. Tohti, and A. Hamdulla, "False information detection via multimodal feature fusion and multi-classifier hybrid prediction," *Algorithms*, vol. 15, p. 119, 03 2022.

[58] P. Shah and Z. Kobti, "Multimodal fake news detection using a cultural algorithm with situational and normative knowledge," in *2020 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1–7, 2020.

[59] A. Vaswani, N. M. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *ArXiv*, vol. abs/1706.03762, 2017.

[60] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

21mcec11_thesis_from_chapter_1

ORIGINALITY REPORT

4%	4%	0%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.todayonline.com Internet Source	4%

Exclude quotes	On	Exclude matches	< 2%
Exclude bibliography	On		

Head Office:
301-310, 4th Floor, Gala Hub, Near Basant Bahar,
Bopal, Ahmedabad, Gujarat - 380058
GSTIN: 24AAECI7718L1ZA

infocusp
Innovations

Date: May 05, 2023

Internship Certificate

Dear Sir/Madam,

This is to certify that **Mr. Yagnesh Bhadiyadra,** a student of **Nirma University** undergoing **M.Tech.,** is doing an internship with Infocusp Innovations and the internship period is from **13 June 2022** to **31 May 2023.**

During this internship programme with us, he was found to be punctual, hardworking and inquisitive.
We wish him every success in life.

Thank you,
Yours sincerely,

NPatel

Nirav Patel
Chief Executive Officer,
Infocusp Innovations

Infocusp Innovations Pvt. Ltd. AHMEDABAD

Branch Office:
601 - Westport, Pan Card Club Road,
Baner, Pune, Maharashtra - 411045
GSTIN: 27AAECI7718L1Z4

Email: inquiries@infocusp.com
Website: www.infocusp.com
CIN: U74994GJ2018PTC100914