JIMMA UNIVERSITY

JIMMA INSTITUTE OF TECHNOLOGY

SCHOOL OF GRADUATE STUDIES

FACULTY OF ELECTRICAL AND COMPUTER ENGINEERING

COMPUTER ENGINEERING

Ethio-Fact Check

Stance Based Fake News Detection for Amharic news Using Dense Neural Network

A Thesis Submitted to School of Graduate Studies, Jimma University, Jimma Institute of Technology, Faculty of Electrical and Computer Engineering in Partial Fulfillment of the Requirements of the Degree Masters of Science in Computer Engineering

By

Tigist Wondiye Tefera

Declaration

I declare that this thesis entitled as **Stance Based Fake News Detection for Amharic news Using Dense Neural Network** is my original work and has not been presented for a master's degree in this or any other universities, and all materials and resources used have been fully cited.

Student Name: ________________________ Signature: _______________

This research has been presented for review with my approval as university advisors.

Signature

Main Advisor: Dr. Eng. Getachew Alemu____________________________

Co-Advisor: ___________________________ **Signature** ___________________

This thesis has been approved by

Board of Examiners **Signature** **Date**

External Examiner: ________________________ ____________________ __________

Internal Examiner: ________________________ ____________________ __________

Chair Person: ____________________________ ____________________ ___________

Abstract

Fake news is characterized as a story made up with the deliberate of misdirecting or deluding. In this article we display the arrangement to the fake news location action utilizing Profound Learning structures for the Amharic dialect. Gartner's investigate [21] predicts that "By 2022, most individuals in developed economies will devour more untrue data than genuine data." The exponential increment within the generation and dispersion of wrong news in Ethiopia and within the world presents the quick have to be consequently tag and identify such bent news articles. In any case, programmed discovery of fake news may be a troublesome assignment to achieve because it requires the demonstrate to get it the subtleties of common dialect. In expansion, most existing fake news discovery models treat the issue in address as a twofold classification movement, which limits the model's capacity to get it how related or irrelevant detailed news is compared to genuine news. To address these gaps, we present neural network architecture to accurately predict the position between a given pair of titles and the body of the article in Amharic language. Our model able to achieve 95.21% accuracy on test data.

List of Abbreviations

Abbreviation	*Meaning*
API	Application Programming Interface
AUC	Area Under Curve
AUPRC	Area Under Precision Recall Curve
GAN	Generative Adversarial Network
GNMT	Google Neural Machine Translation
LCWA	Local Closed-World Assumption
LSTM	Long-short Term Memory Network
NER	Named Entity Recognition
NLP	Natural Language Processing
PCA	Principal Components Analysis
POS	Part of Speech
RDF	Resource Description Framework
TF-IDF	Term Frequency-Inverse Document Frequency
URL	Uniform Resource Locator

Table of Contents

CHAPTER ONE

INTRODUCTION

1.1. BACKGROUND

A proverb in one of Ethiopian language (Amharic) (official working language of the Ethiopian government) describes fake news or fabricated news as "**በሬ ወለደ**!" This means: "**An Ox delivers a baby calf!**" which is totally fabricated.

Within the later a long time, a significant sum of inquire about has been conducted in this region with palatable comes about. With the victory and development of Counterfeit Insights and Machine Learning, innovation has soothed human from unessential endeavors. Fake news discovery utilizing these innovations can spare the society from superfluous chaos and social distress.

This research "Fake News Discovery System" employments machine learning calculations and normal dialect handling methods. Machine learning may be a subset of counterfeit insights within the field of computer engineering that frequently employments measurable procedures to donate computers the capacity to memorize with information, without being unequivocally modified [23]. Natural –language processing is an zone of two computer sciences and counterfeit insights concerned with intuitive between computers and human (characteristic) dialects, in specific how to program computers to handle and analyze expansive sums of characteristic dialect information [4].

Sentiment analysis refers to the management of sentiments, opinions, and subjective text [21]. Opinion examination gives the comprehension data related to open sees, because it analyzes diverse tweets and surveys. It could be a confirmed apparatus for the expectation of numerous critical occasions such as box office execution of motion pictures and general elections [22]. Open surveys are utilized to assess a certain substance, i.e., individual, item or area and may be found on diverse websites like Amazon. The suppositions can be categorized into negative, positive or unbiased. The reason of estimation investigation is to consequently decide the expressive heading of client surveys [23]. The request of estimation examination is raised due to extend prerequisite of analyzing and organizing covered up data which comes from the social media within the shape unstructured information [4]. As of late social media platforms have ended up a fundamental medium for individuals to precise their day by day exercises, responses and feelings. Blogs and small scale blogs are the foremost common shape of social media [5]. Blogs are casual locales on the around the world web where clients are utilized to post thoughts, dialogs, and contemplations on a specific issue [6]. Whereas microblogs are littler blogs with brief posts up to a restricted number of signs [7]. One thing they have in common is that they comprise of sections that are recorded in a chronologically slipping arrange (i.e. the most recent news is on beat). They are instruments that empower talks and comments on data shared with other clients. They are characterized by its energetic and up to datedness. Around 16,037,811 Web clients on June/2017, 15.4% of the Ethiopian populace per ITU 4,500,000 Facebook clients on June/2017, 4.3% entrance rate [8]. Among this cyber clients citizen most of them taking an interest on distinctive issues and deliver input towards the issue raised by the his/her companions, blogs, Medias, daily papers by mother tongue dialect. Be that as it may still presently as my concern there's no work done on assumption mining for socio-political space based on opinioned Amharic content utilizing profound learning approach. Past supposition mining procedures and opinion mining models are frequently created for Amharic dialect in motion picture survey and item benefit regions, the past analysts Abebe [9], Mengistu [10] and Abraham[11] taken after a administered machine learning approach which utilizes Naïve Bayes Choice Tree, Bolster Vector machine, Multinomial Naïve Bayes and Greatest Entropy calculations utilizing Sack of words, Data pick up, n-grams nearness, n-grams recurrence & n-grams-TF-IDF highlights for report level Amharic assumption classification. In show disdain toward of hence analysts our investigate center on sociopolitical space and we collect the information set from the official Facebook page of Fan broadcasting by utilizing Facebook information extraction instrument known as Facepager utilizing chart API. In expansion to this we take after a profound learning approach to prepare and test the show which is novel strategy for the region in our dialect. One of the current patterns in social media individuals express their felling towards freely posted issues utilizing feeling symbol (Emojis). Hence, feeling symbol related with Amharic content express diverse opinion of the individuals approximately issues raised by the online community. The analyst come up with novel thought which is consolidation of feeling symbols with Amharic comments in opinion investigation. In terms of information required to prepare and test the demonstrate the profound learning endeavors to memorize tall level reflection by misusing the progressive models. It may be a promising approach and has been broadly connected learning, semantic parsing, common handling and numerous more. We select profound learning approach since of progressed capacity of chip handling (GPU unit), broad lower consumption of equipment and noteworthy improvement in machine learning calculation.

1.1.1. But what is fake news?

Concurring to the writing, fake news implies untrue news, lie, misdirect, cheating, figment, deceiving, sham, reenactment, creation, control and purposeful publicity. In other words, fake news incorporates too "defamation, through spreading of unfaithful actualities of another individual intentionally; actuation of contempt or the making of dangers in a way which will aggravate open peace. [2]

1.1.2. Why social media?

With the headway of innovation, data is unreservedly open to everybody. Web gives a tremendous sum of data but the validity of data depends upon numerous components. Colossal sum of data is distributed every day through online and print media, but it isn't simple to tell whether the information could be a genuine or untrue. It requires a profound consider and examination of the story, which incorporates checking the actualities by evaluating the supporting sources, by finding unique source of the data or by checking the validity of creators etc. The manufactured data may be a think endeavor with the aim in arrange to damage/favor an organization, substance or individual's notoriety or it can be essentially with the thought process to pick up monetarily or politically. "Fake News" is the term coined for this kind of created data, which deludes individuals.

The reason we are attending to check realities on social media is, since the private press has been battling to outlive and the street ahead is full of challenges. One of the strategies the government employments to debilitate the press is to starve them of pivotal salary by debilitating government organizations from publicizing. A few businesses are too perplexed to publicize in media seen as basic of the government, and have been undermined by insights agents for doing so. [1]

The other challenge comes from the open. A few individuals are not fascinated by perusing the surviving distributions since they consider them cowed by the government. They are seen as only attempting to remain above water, cautiously observing which way the wind blows, and dodging being basic on critical political issues.

Since of the over and other unmentioned issues people groups believe and utilize social media news as trusted and nongovernmental predisposition news in Ethiopia, of course social medias are a source of news all over the world, as news are back bone of our world at this time so a few administrative medias particularly in creating nations like Ethiopia may not transmit the data on time and not in one-sided way, so that people groups within the social media will claim the report as quick as conceivable in that way social medias are exceptionally portable and fair media within the world, within the other hand a few people groups purposely use this media for a few awful reason, for misinform the client, to make a few undesirable sentiments between the people groups, to form mutilated or totally fake news to the people groups, so that they get their fulfillment of the terrible feelings between people groups.

There are social media rules that can be executed to avoid such kind of terrible assignments within the social media, such as, Ethiopia received its to begin with national arrangement on Data and Communication Advances (ICTs) in 2002, which has since been overhauled in 2009 and 2016. Other related approaches have moreover been as of late received by the government as of late such as the Data Security Approach of 2011 and the Broadband Technique of 2016. A few of these approach records are deciphered into a number of laws counting the a few draft pieces of enactment on e-commerce, e-signature and the more as of late ordered Cybercrime and telecom extortion decrees. Actualizing offices have too been introduced to roll the arrangements and laws into activity. But the record so distant in creating a coherent and legitimately faultless approach in managing with rising Cyber dangers is dreary. This can be especially the case in managing with the rising rebel web culture in Ethiopia. [2]

But when we see the rules on the social media it negates with the concept of majority rule government, anyone have a full right to claim anything. In Ethiopian structure Article 29 Right of Thought, Conclusion and Expression say's "Everyone has the correct to hold opinions without interference.", "Everyone has the correct to opportunity of expression without any impedances. This right shall incorporate opportunity to look for, get and confer data and concepts of all sorts, in any case of wildernesses, either orally, in composing or in print, within the frame of craftsmanship, or through any media of his choice." "Freedom of the press and other mass media and flexibility of artistic inventiveness is guaranteed." [3] So that the government cannot take absent people groups majority rule government which is composed within the structure but people groups utilize those rights and manhandle with fake and disconnected news on the media, when we watch from literary works what are the greatest fake claims(news) in Ethiopia?

1.1.3. Impact of Fake News in Ethiopia

** October 23, 2018 was sad day for two longtime companions who went to their hometown to conduct a therapeutic inquire about. The investigate comes about and its point would have been a awesome bargain for the nearby community. In any case, their exertion and goal was cut brief with a catastrophe that claims their lives. They were beaten to passing by an irate horde in West Gojjam Zone, Amhara Territorial State. A analyst and a PhD candidate at the Ethiopian Organized of Water Assets and his companion, who is additionally an Ethiopian analyst, were murdered whereas conducting a investigate at a school in West Gojiam. Agreeing to reports from the region, things took a turn to most noticeably awful on that game changing day and it was all since of unwarranted rumors or fake news supposedly circulated within the social media on the day of the slaughtering. The rumors were to begin with started within the social media and after that found their way into the town, agreeing to those near to the case. [4] Clearly, the episodes had distant coming to results for the wellbeing segment. Since at that point, wellbeing laborers in different locales are confronting issues to practice their regular schedule, and it is all since of fake news within the virtual world, especially Facebook.

** Another fake news which said to relate the rally in Mekelle was circulated by activists within the social media. The news was said to be compile by the Washington Post and detailed the rally as in case it was primarily organized by dissidents against Prime Serve Abiy Ahmed (PhD). Afterward, it was demonstrated that Washington Post has never done any detailing approximately the rally and the claim was completely a creation, and so numerous fake claims. [4]

Nowadays, noxious parties utilize untrue accounts to impact conclusions around the world. Brought to light amid in 2016, the term ”fake news” is presently a all inclusive significant issue. Tragically, the pace of fake news advancement is quick drawing closer a point at which the normal human will be incapable to recognize truth from camouflaged fiction. A few progressed generative models such as GPT-2 discharged by OpenAI have as of now come to a point at which the makers considered the demonstrate substance to be as well ”human-like” for open discharge, inciting fear and caution approximately the increasing speed of about imperceptible fake content [22].

There's moreover presently a solid money related motivating force for content fashion that's continually advancing and get away discovery by state of the craftsmanship fake news discovery calculations. In little towns as inaccessible as Veles in Macedonia, numerous presently see fake news era as a profitable career way that indeed has an official educational programs and college to lure a developing number of youthful individuals to create created articles [10]. These worldwide patterns require the investigation of overhauling, transient models able of taking care of gushing information and complex designs [30]. It is anticipated that by 022, the created world will see more fake news than genuine data on the web [21]. Modern methods in manufactured insights are driving the charge within the generation of such fakes, but similarly offer us the opportunity to analyze tremendous sums of information and confirm substance to combat the deluge of misinformation[16]. Ethiopia as, of late, too seen a prime illustration of the pervasiveness of fake news. Within the later 2018 decision, an Oxford think about found that 1 in 3 articles shared on social media during the race period were undoubtedly false[14]. It is evident that there's a require for littler dialect communities to be able to evaluate the veracity of their news sources and their related claims.

1.1.4. Impact of Fake News in the World

Fake news location has as of late pulled in a developing intrigued from the common open and analysts as the circulation of deception online increments, especially in media outlets such as social media nourishes, news blogs, and online daily papers. A later report by the Jumpshot Tech Web journal appeared that Facebook referrals accounted for 50% of the entire activity to fake news destinations and 20% add up to activity to legitimate websites. Since as numerous as 62% of U.S. grown-ups devour news on social media (Jeffrey and Elisa, 2016), being able to identify fake substance in online sources could be a squeezing need [5].

Rumor Cascade:

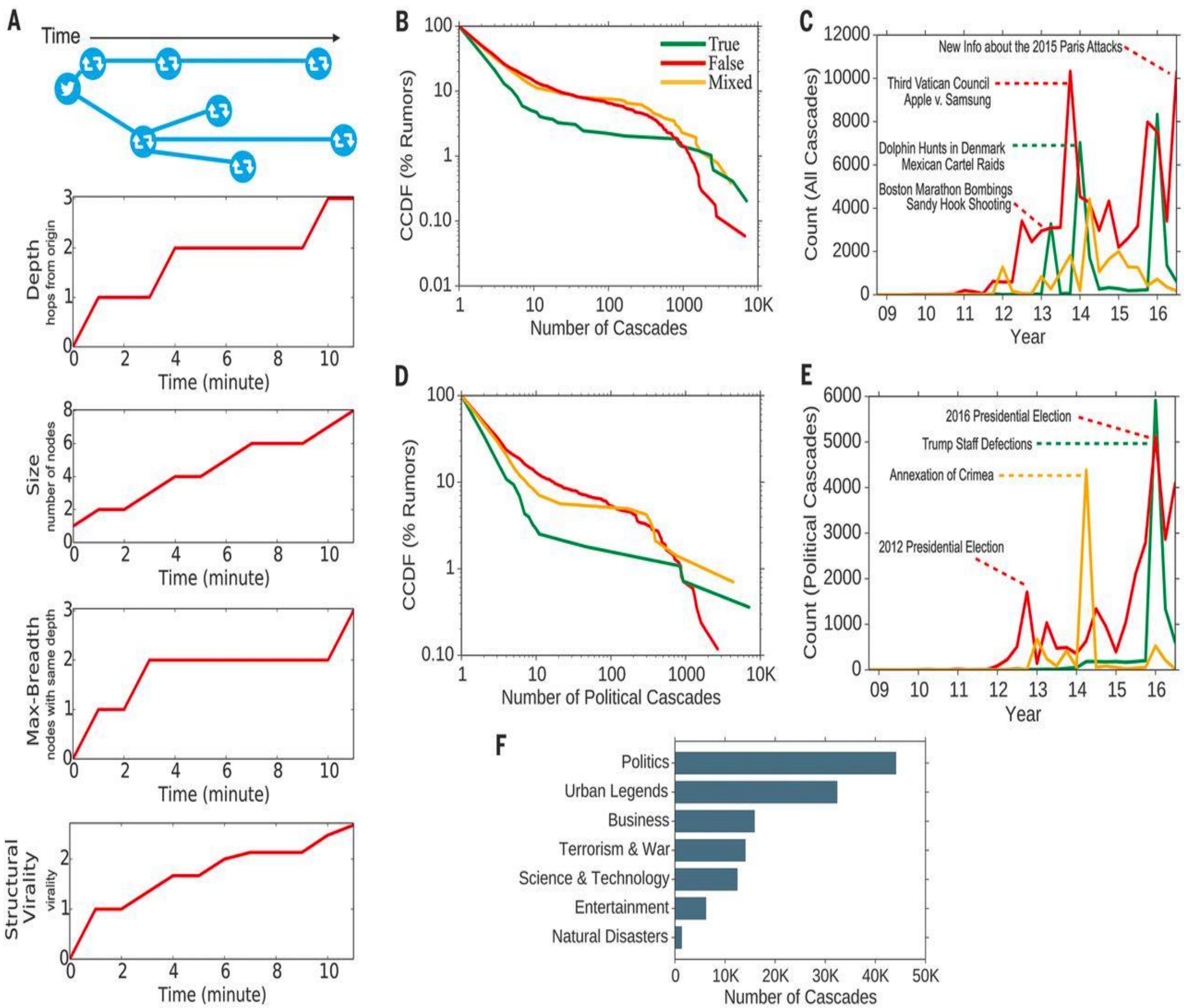


Figure 1: Rumor cascades. [5]

An case rumor cascade collected by the creator strategy as well as its profundity, measure, most extreme breadth, and basic essentialness over time. "Nodes" are clients. (B) The complementary total dispersion capacities (CCDFs) of genuine, untrue, and blended (mostly genuine and in part untrue) cascades, measuring the division of rumors that display a given number of cascades. (C) Quarterly tallies of all genuine, untrue, and blended rumor cascades that diffused on Twitter between 2006 and 2017, commented on with case rumors in each category. (D) The CCDFs of genuine, untrue, and blended political cascades. (E) Quarterly tallies of all genuine, wrong, and blended political rumor cascades that diffused on Twitter between 2006 and 2017, explained with case rumors in each category. (F) A histogram of the whole number of rumor cascades in our information over the seven most visit topical categories.[5]

1.1.5. Why Amharic Language?

We choose Amharic language because "Amharic shall be the working language of the Federal Government." as the constitution says, and also we can talk, listen and write the language.

1.1.6. State Of The Art

At present, false news can be created and easily diffused through many social networking platforms, resulting in a wide-ranging impact in the real world. Modeling and distinguishing how misinformation spreads on social platforms and why they succeed in fooling readers is critical to developing effective algorithms and tools for early detection. Modern boom in research in this area aims to address key issues using ways rely on automated learning and deep learning and engineering features and mining on the graph and analyze images and video, in addition to data that have been newly created Web services groups to identify phishing content. Most of the research targeted fake reviews, biased messages and information against facts (false news and deception).

And nowadays the researchers are using combination of feature engineering and machine learning in order to tackle the problem, but in Ethiopia we are still trying to tackle the problem traditionally which means by searching or referring news sources to check whether the claim is true or fake. Prime minister office announce that there will be rule on hate speech and fake news dissemination on social media but still the problem will remain because of the characteristics of the media which anybody can have any number of fake accounts, and so on.

So our model tackle fake news problem in local language by checking all the claims and says "**እዉነት**, **ሀሰት**".

1.1.7. Statement of the Problem

Misinformation is also known as a "gimmick" that occupies a large amount of cyberspace around the world. The rapid and widespread spread of electronic technology threatens it. Advertising through this false news in cyberspace has been adopted today by states, institutions and individuals for various reasons and forms. Erotic news is often created and disseminated across social networks to achieve the desired goal. On the other hand, it can also involve a real fact, but deliberately exaggerated. This may also include creating web pages with misleading headlines or logos to attract readers' attention. This misrepresentation can lead to fences, social unrest, or financial fraud because of this distortion, political gain, increased readership or earnings associated with click, etc. This can also affect the importance of serious media.

Another danger lies in other electronic media that use this as a source of their news, and thus continue to publish news. The problem is to determine the validity of news and content online. Equally important is the problem of identifying robots involved in spreading false news.

Listed Problems of the Statement

Fake-News may lead to:

- Fencing
- Social unrest
- Financial fraud
- Bad political Interest gain
- Affect the importance of serious media News

1.1.8. Objectives

General Objective:

The general objective of this thesis research is to design and implement an algorithm for Automatic system which can check the truthiness of claims posted in social media using local language Amharic.

Specific Objectives:

- ✓ To create fake news dataset in Amharic
- ✓ To prepare dataset by preprocessing
- ✓ To select and adopt the best algorithm for our research
- ✓ To detect and classify the truthiness of a given claim
- ✓ To reduce the social conflict caused by fake news

1.1.9. Scope of the research

This investigate venture centers on as it were a restricted set of news articles over a given occasion skyline inside a given time period. It is in this way not planned to speak to a expansive body of information, but or maybe a centered set of articles that speak to ”fake news” inside a specific setting over a specific time period.

1.3. Fake News

Fake news can be come in numerous shapes, counting: inadvertent mistakes committed by news aggregators, by and large wrong stories, or the stories which are created to delude and impact reader's conclusion. Whereas fake news may have numerous shapes, the impact that it can have on individuals, government and organizations may by and large be negative since it varies from the facts.

1.3.1. The Problem of Defining Fake news

The task of defining fake news is a challenge by itself and is open to interpretations. Table 1 has the prevalent definitions of fake news:

Definition statement
A made-up story with an intention to deceive [14]
News articles that are intentionally and verifiably false, and could mislead readers [22]
Fake news is a type of yellow journalism or propaganda that consists of deliberate misinformation or hoaxes spread via traditional print and broadcast news media or online social media [15].

Misinformation is a common theme when fake news is mentioned. Misinformation can itself be classified as shown in Table. 2.

Table 1: Definitions for fake news

	False connection	Misleading content	False context	Imposter content	Manipulated content	Fabricated content	Satire or Parody
Poor Journalism	✓	✓	✓				
To parody				✓		✓	✓
To Provoke or to 'punk'				✓	✓	✓	
Passion			✓				
Partisanship		✓	✓				
Profit	✓			✓			
Political Influence		✓	✓		✓	✓	
Propaganda		✓	✓	✓	✓	✓	

Table 2: Misinformation Matrix [7]

Considering these challenges included in discovery of fake news, a great to begin with step is to identify the position between the body of content and the substance it's portraying. The assignment of position discovery can be portrayed as the method of consequently anticipating in case the news article or social media substance is concurring, opposing this idea or disconnected to the substance it's portraying. The underneath case in Table 3, clarifies position between news features and news article.

Article body	Headline	Stance
በዚህም ምክንያት የኢንፎርሜሽን መረብ ደህንነት ኤጀንሲ አግኝቸዋለሁ ባለው ምክንያት ኢንተርኔት ሙሉ ለሙሉ በሐገር ዓቀፍ ደረጃ ለ30 ደቂቃዎች ተቋርጦ እንደነበር ኢትዮ ቴሌኮም ተናግሯል።	በኢትዮጵያ የፋይናንስ ተቋማት የሳይበር ጥቃት ሊደርስባቸው ነበር ተባለ።	Agree
ከአቅም በላይ በተፈጠረ ችግርም ኢንተርኔት መቋረጡን የኢትዮ ቴሌኮም ዋና ስራ አስፈፃሚ ፍሬህይወት ታምሩ ለሸገር ነግረዋል፤ ደንበኞችንም ይቅርታ ጠይቀዋል።	በኢትዮጵያ የፋይናንስ ተቋማት የሳይበር ጥቃት ደረሰባቸው ተባለ።	Disagree
ኢትዮ ቴሌኮም ኢንተርኔት የተቋረጠው አቅጀው ሳይሆን አጣዳፊ ጉዳይ ስለገጠመኝ ነው ብሏል፤ለደንበኛም ለማሳወቅ ግዜ እንዳልነበረ ፍሬህይወት ታምሩ አስረድተውናል።	ዩኒቨርስቲዎች በፌደራል ፖሊስ ሊጠበቁ ነው- የሳይንስና ከፍተኛ ትምህርት ሚኒስቴር	Unrelated
ግብጽ በሕዳሴ ግድብ ላይ እያሰማች ያለውን የተዛባ አመለካከት ለማስተካከል ጠንካራ የዲፕሎማሲ ስራ እንደሚሰራ ምክትል ጠቅላይ ሚኒስትር ደመቀ መኮንን ተናገሩ ። የታላቁ የኢትዮጵያ ሕዳሴ ግድብ ሕዝባዊ ማስተባበሪያ ምክር ቤት ፅህፈት ቤት የስራ አስፈጻሚ አባላት ትናንት ባካሄዱት መደበኛ ስብሰባ ግድቡ ከኃይል ምንጭነት በተጨማሪ የወንዙን ጂኦ ፖለቲካ የቀየረ መሆኑ ተገልጿል። የስራ አስፈጻሚ አባላቱ ግብጽ በግድቡ ዙሪያ እያከናወነች ያለውን የተዛባ ትርክት ማስተካከል እንደሚገባ አንስተዋል።	ግብጽ የምታሰማውን የተዛባ ትርክት ለመቀየር ቀጣይነት ያለው ጠንካራ የዲፕሎማሲ ስራ ይከናወናል – አቶ ደመቀ መኮንን	Discussed

Table 3: Example of News Article and Headlines and their corresponding stance in Amharic

The core task of detecting fake news involves identifying the language (set of words or sentences) which is used to deceive the readers. The idea of classifying fake news by learning word-level meaning is a very challenging task under the skin. For instance, consider Table 4.

Type of Fake news	Example
100 % Fake	ተወዳጁ ድምፃዊ በራሱ ጊቢ በር ላይ መኪናዉ ዉስጥ ሙቶ ተገኘ። የቴዲ ሚስት ለፖሊስ እንደተናገረች ዛሬ ጧት የጣናን እምቦጭ ለመታደግ ከባለሀብቶች ጋር ስብሰባ አለኝ ብሎ ስልክ ተደዉሎለት ነው ከቤት የወጣዉ ብላለች። በአሁን ሳዓት እስክሬኑ ለምርመራ ተብሎ ጳውሎስ ሆስፒታል ይገኛል። ኢትዮጵያ ከባድ ሰዉ አጣች የአገር ፍቅር የሚያገበግበዉ ኢትዮጵያ በጎደላት ቦታ አለዉ ብለዉ የምሞላና ለጋሽ የሚስክኖች እምባ ጠራጊ ነበር ለቴድ አፍሮ ቤተሰቦች እና ለሞላዉ ኢትዮጵያ ህዝብ መፅናናትን እንመኛለን ነብስህ በአፀደ ገነት ያኑርህ
Slanted and Biased	ሀገሪቷን የውጭ ጠላት የወረራት ይመስል ያለውን ጦር በሙላ ወደ ኦሮሚያ እየላከ ነው። ዛሬ ወደ ሻሸመኔ ከተማ የተላከ የጦር መሳሪያ ነው። ከዚህ በኋላ አይደልም ታንክ ንኩለር ብታፈናዳ ወደኋላ የሚመለስ ትግል የለም።
Misusing the Data	ትግራይ በምርጫ ጣቢያ ምክንያቱ ባልታወቀ ጉዳይ ግጭት ተቀስቅሷል ከፍተኛ ድብድብ ነበር ፤ በዚህም በርካታ ድምፅ የተሰጠባቸው ኮሮጆዎች ጠፍተዋል። ከስፍራው የደረሰኝ ቪዲዮ ይህንን ይመስላል።

Table 4: Examples of fake news [13]

1.3.2. Challenges

The detection of fake news presents a slew of challenges, some of which are discussed further in this report, including:

1. Fake news is troublesome to characterize concisely and reliably, as its nature changes altogether over time. This implies that whereas within the past, simply complex approaches were very effective, the joining of fake news to the composing fashion of genuine news will likely lead to corrupting execution. Understanding how fake news is created seem, in this manner, lead to insights that are pro-active instead of review in fake news detection.

2. Later ponders have appeared that fake news stories spread more rapidly than they can be recognized, so the sources of fake news too got to be identified instead of centering as it were on person articles. [16]

3. Ground-truth confirmation of article claims in total isn't continuously conceivable since ponders have appeared that people are ”average” at identifying duplicity, with an precision within the extend 55-58% [24].

4. Characteristic dialect handling approaches are vulnerable to ill-disposed assaults (e.g. a fake news article created by a GAN calculation) that mimicks the see and feel of a trusted news source [8].

5. ’Fake news’ could be a intensely context-dependent and time-dependent classification, as news is as it were current for a certain period of time, and rectifications or withdrawals are common.

6. The subject of fake news discovery in dialects other than English has been underrepresented in inquire about and hence administered approaches that work well in English don't perform as well in non-English domains. One of the most reasons for this can be the need of labeled preparing data.

The over cases capture the complex nature of discovery and classification of fake news. To properly classify the over sorts of fake news, our dialect model needs to get it the nuances included in passing on messages through text. Detecting fake news is difficult for numerous reasons. To begin with, manual errand of identifying fake news is exceptionally subjective. Evaluating the veracity of a news story could be a complex and cumbersome task, indeed for prepared specialists. News isn't as it were spread through traditional media outlets any longer but moreover through different social media channels. Computerized arrangement requires understanding the normal dialect handling which is difficult and complex. These complexities make it a overwhelming errand to classify content as fake news.

The combination of over issues of definition and location makes the assignment of stance location to unravel the programmed fake news classification challenging.

1.4. Contributions

The primary contribution of this research is the design of Amharic claim fact-checking model, which is going to check the stance of a claim with the trusted news agencies news articles. This proposal investigate too centers on demonstrate clarify capacity, as the framework proposed within the venture aims for a human within the circle plan, meaning that the model should expand the capacity of the end-user instead of be a black-box robotization arrangement. Since typically the primary endeavor to utilize position discovery for fake news discovery in Amharic so, it is the trust that this investigate brought a standard dataset for proceeded inquire about into fake news location in Amharic in our nation Ethiopia.

1.5. Thesis Structure

The proposal is organized within the taking after way. In Chapter 2, Related Work, the hypothetical establishments of position discovery as well as the issue of ’Fake News’ are examined. In Chapter 3, the strategy of the chosen models is investigated. In chapter 4 the Amharic fake news dataset planning handle has been clarified in detail. The comes about from these strategies are collected in Chapter 5, counting an assessment of results and a discussion of their impediments. The ultimate chapter, chapter 6 contains an reply to the objectives set in Chapter 1 as a conclusion as well as a talk on future work in this range.

CHAPTER TWO

2. LITERATURE REVIEW

Broad writing surveys from distinctive books; diaries articles, proposal, and the Web are conducted, so as to have a strong and concrete understanding on standards, methods, and apparatuses of a Stance-based fake news checking. Moreover, inquiries about that are conducted on distinctive fake news location framework in social media and other related work are surveyed.

The impacts of fake news on peace and development in the world: the case study of Ethiopia. [24]

The creator of this Paper has looked into and analyzed purposively chosen reports of mass media (“mainstream media”), i.e. electronic and print media, particularly within the Western world, for occasion in Germany concerning the turmoil in a few towns of Oromia and Amhara Territorial States of Ethiopia in harvest time 2016.

Afaan Oromo Text Content-Based Fake News Detection using Multinomial Naive Bayes. [25]

In this paper the creators investigate the application of characteristic dialect handling strategies with multinomial naïve Bayes for the location of "fake news" on 752 news datasets in Afaan Oromo dialect. The information set was tried on multinomial naïve Bayes classification calculation. They apply Term recurrence, and term frequency-inverse record recurrence (TF-IDF) of unigram and bi-grams and discover that term recurrence of unigram of this demonstrate recognizes fake news sources with an precision of 96%, but we are able see the demonstrate is having slight impacts on review.

Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques by **Hadeer Ahmed, Issa Traore, and Sherif Saad [1]**

in college of Victoria 2017. In this paper the analysts propose, fake news discovery show that utilize n-gram investigation and machine learning procedures as an objective. They explore and compare two distinctive highlights extraction methods (TF and TF-IDF) and six distinctive machine classification methods (SVM, LSVM, KNN, DT, SGD, and LR). From the comes about gotten in their tests, Linear-based classifiers (Straight SVM, SDG, and Calculated Relapse) accomplished superior comes about than nonlinear ones. In any case, nonlinear classifiers accomplished great comes about as well; DT accomplished 89% precision. The most noteworthy precision was accomplished utilizing Direct SVM as 92%. This classifier performs well no matter the number of include values utilized. Too with the increment of n-gram (Tri-gram, Four-gram), the exactness of the calculation diminishes. Moreover, TF-IDF outflanked TF. The least exactness of 47.2% was accomplished utilizing KNN and SVM with four-gram words and 50,000 and 10,000 include values. We will watch that

Fette et al. utilized machine learning to classify an mail as phishing or not by utilizing highlights such as age of URL, number of dabs in URL and HTML substance of e-mail whereas getting a tall exactness of 99.5%.[7]

Bajaj et al. utilized Two-Layer Bolster forward neural arrange, Repetitive neural arrange, Convolution neural arrange, Long-Short term recollections, Gated repetitive units and watched that the RNN design with GRUs beated one with LSTM cells for the location of fake news. The bolster forward arrange performed well as a classifier but the convolution arrange didn't perform well. [10]

Yang et al. displayed a trailblazing vision on how criminal accounts tend to be socially associated shaping a little world organize. Criminal accounts arbitrarily take after accounts, anticipating them to take after back. There are three sorts of criminal supporters found specifically, social butterflies, social promoters and sham.[8]

Conroy et al. utilized phonetic approach (Both words and sentence structure) as well as organize approach which says that message metadata or organized information organize questions can be tackled to supply total double dealing measures. They utilized Centering Reverberation Investigation (CRA), a mode of network-based content examination.[9]

A work by **Antoniadis et al. (Antoniadis et al., 2015**) attempted to recognize deception on Twitter. The creators explained a expansive dataset of tweets and created a show utilizing the highlights from the Twitter content, the clients, and the social criticism it got (number of re-tweets, number of favorites, and number of answers). At long last, they evaluated the capability of the show in recognizing deception in genuine time, i.e. in a need (when the social input isn't however accessible). Assessments on real-time as it were rot in 3% when compared with the demonstrate that employments all accessible highlights.

Shu et al. said that pernicious accounts can be effortlessly and rapidly made to boost the spread of fake news, such as social bots, cyborg clients, or trolls. The existing calculations for location of fake news are either (i). News Substance Based or (ii). Social Setting Based. News substance based approaches center on extricating different highlights in fake news substance, counting knowledge-based and style-based. Social setting based approaches point to utilize client social engagements as assistant data to assist identify fake news. [12]

Galuba et al. focussed on characterizing and modeling the information cascades formed by individual URL mentions in the Twitter follower graph. They constructed two models of information propagation in social networks namely At-least-one (ALO) model and Linear Threshold model. The linear threshold model was able to correctly predict almost half of the URL mentions.[14]

Pisarevs kayaet al. aimed to reveal fake and truthful news using markers from different linguistic levels. They used POS tags, length of words, sentiment terms and punctuation on the lexics level. They also used Rhetorical Structures Theory (RST) relations as markers on the discourse level.[17]

Summarize the literatures that we have seen previously, now a day's peoples are in the vicinity of social media news both true and false and we cannot prevent that fake news because of peoples and media speech and write rights that has been given in country law as media freedom.

So that fake news has become a leading issue in the world right at this time. As the literature says most of the social unrests in the world are occurring because of fake or untrue news especial in developing countries fake news have a big chance to spread out fast and which is uncontrollable but we can detect those rumors using machine learning.

So to tackle this problem many researchers throughout the world are spending their time and exertion in order to prevent social unrest that will be occurred because of those fake news in different medias by detecting and telling which one is true and which one is not using machine learning different methodologies and techniques. We may be able to prevent but we can detect fake news as we already seen in the related works review with different techniques.

Our main objective here is to implement a model that can detect fake news from social media in Ethiopian local language (Amharic). Since the topic (fake news) is new research area there is no or very few papers are there to review especial in our country with local languages we can say there are no papers to review also no fake news dataset in Amharic language, so with the intention of implementing and testing the model we need to construct a dataset for the training and testing of the proposed model. We collect mixed type valued (fake and true) Amharic news from different news broadcasts and then need to do the data preprocessing step, since we are implementing fake news detection model for multilingual this process is the main process, before representing those data we need to do some refinements like stop-word removal, tokenization, a lower casing, sentence segmentation, and punctuation removal then we can represent the data in vector-based model.

As the literature says if the learning and testing datasets are well prepared the accuracy of the model will be much higher than the model which is trained by with the data which is not well organized.

CHAPTER THREE

3. METHODOLOGY STRUCTURE

This chapter starts by defining the data framework for our model in machine learning and goes deep in to fake news detection methods and picks the stance detection method in order to detect how the claim is related to the trusted news agencies like Fana, EBC, Walta, Voa Amharic … news articles.

3.1. Data framework

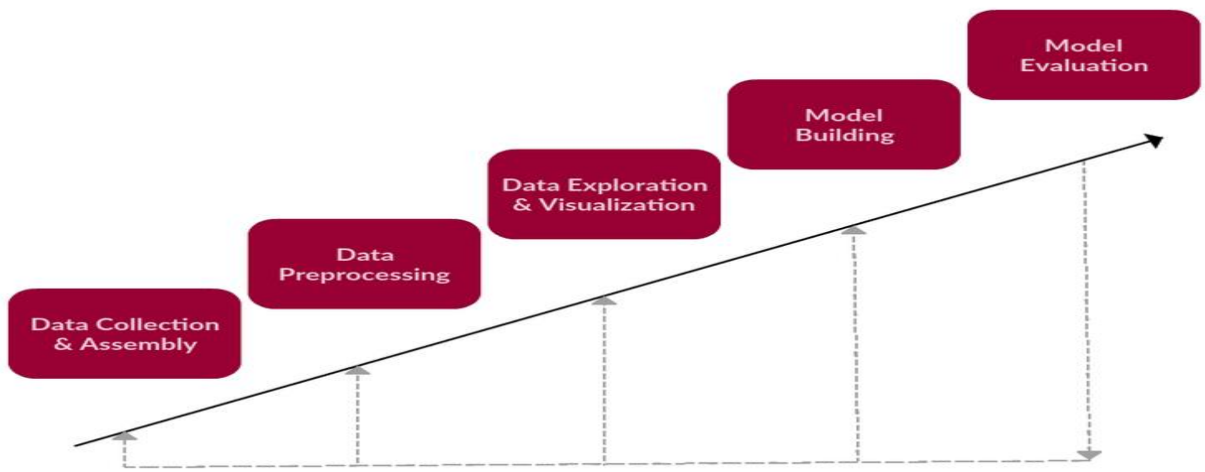


Figure 2: Data framework [1]

The entire research has been defined by following steps and illustrated in Fig. 2:

- Data Retrieval: collecting the data.
- Data Preprocessing: make the data ready to feed to the model
- Data Visualization:
- Tokenization
- Feature Extraction
- Machine Leaning Algorithms
- Training & Testing Model
- Evaluation Metrics

3.1.1. Data Retrieval and Preprocessing

Instead of succeeding this thesis research the first job to do is collecting news from different sources which are true and also false in different categories which means politics, health, social, entertainment, religion, business, technology, sport and religion. Amharic fake news dataset is collected from different news agencies which are Amharic news reporters, since we are working in Amharic language we need Amharic news agencies like, ኤፍ ቢ ሲ, ቪኦኤ ዜና, ቢቢሲ ዜና, ኢቢሲ, Ethiopian Reporter, walta, voaAmharic, BBCamharic, wolotube, axumitetube, diretube, ethiopianDj, etegetube, and also our dataset is based on social media (Facebook) so we need to have some news from individuals who has much followers on Facebook like, ጌጡ ተመስገን, Elias Meseret. Totally our dataset have 5,000 numbers of rows collected from 15 Amharic news sources with 9 news categories.

3.1.2. Data Preprocessing

Content information requires uncommon preprocessing to actualize machine learning or deep learning calculations on them. There are different procedures broadly utilized to change over text data into a frame that's prepared for modeling. The information preprocessing steps that we outline underneath are connected to both the features and the news articles. We moreover provide insights into diverse word vectors representations we utilized as portion of our analysis.

Stemming

Stemming may be a procedure to evacuate prefixes and additions from a word, ending up with the stem. Utilizing stemming we are able decrease inflectional shapes and sometimes derivationally related shapes of a word to a common base shape. Figure 5 appears the example of stemming technique. We have done the stemming with the help of language experts.

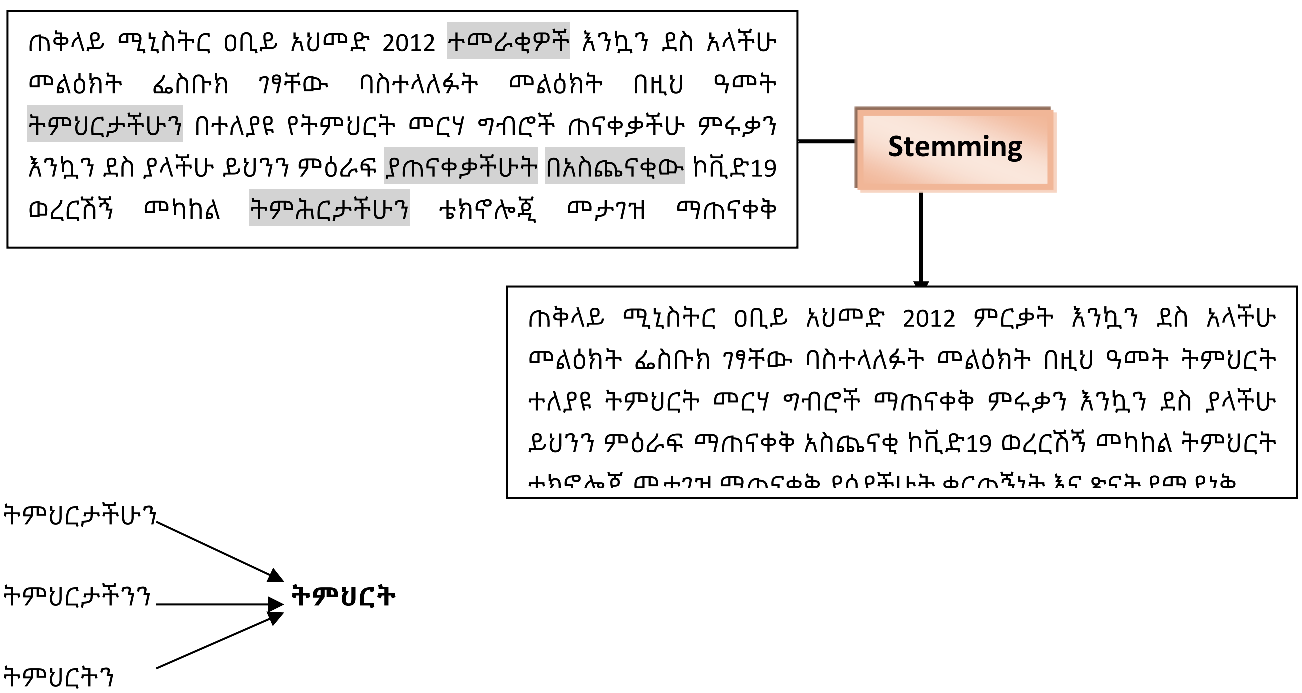


Figure 3: shows an example of stemming process.

Stop Word Removal

Stops Words (most common words in a dialect which don't give much setting) can be handled and sifted from the content as they are more common and hold less valuable data. Stop words acts more like a connecting part of the sentences, for example, conjunctions in Amharic like “**እና**”, “**ወይም**” and “**ግን**”, prepositions like “**ከ**”, “**ለ**”, “**የ**”, “**በ**” etc. and the articles “**ነዉ**”, “**አዎ**”, “**ናቸዉ**”, “**እንደ**”, “**ስለ**”, “**ላይ**”, “**ነበር**”, “**ምን**”, “**መቼ**”, and “**የት**”, Such stop words which are of less importance may take up valuable processing time, and hence removing stop words as a part of data preprocessing is a key first step in natural language processing. We used Natural Language Toolkit – (NLTK) library to remove stop word.

This task by itself is an independent research area for Amharic language so that we try to refer an Ethiopian researcher who did this topic as a research area and work together with them because this is beyond the scope of our research. An Ethiopian researcher at Addis Abeba University has already done the stop word removal work, so that we just adopt it. [19]

Figure 3 illustrates an example of stop word removal.

ጠቅላይ ሚኒስትር ዐቢይ አህመድ ለ2012 ዓ.ም ለተመራቂዎች የእንኳን ደስ አላችሁ መልዕክት አስተላለፉ። በፌስቡክ ገፃቸው ባስተላለፉት መልዕክት በዚህ ዓመት ትምህርታችሁን በተለያዩ የትምህርት መርሃ ግብሮች ጠናቀቃችሁ ምሩቃን እንኳን ደስ ያላችሁ ብለዋል። ይህንን ምዕራፍ ያጠናቀቃችሁት በአስጨናቂው የኮቪድ19 ወረርሽኝ መካከል ቢሆንም፤ ትምሕርታችሁን በቴክኖሎጂ በመታገዝ ለማጠናቀቅ ያሳያችሁት ቁርጠኝነት እና ጽናት የሚደነቅ ነው በማለት አስፍረዋል።

Stop Word Removal

ጠቅላይ ሚኒስትር ዐቢይ አህመድ ለ2012 ተመራቂዎች እንኳን ደስ አላችሁ መልዕክት። ፌስቡክ ገፃቸው ባስተላለፉት መልዕክት በዚህ ዓመት ትምህርታችሁን ተለያዩ ትምህርት መርሃ ግብሮች ጠናቀቃችሁ ምሩቃን እንኳን ደስ ያላችሁ። ይህንን ምዕራፍ ያጠናቀቃችሁት አስጨናቂው ኮቪድ19 ወረርሽኝ መካከል ቢሆንም፤ ትምሕርታችሁን ቴክኖሎጂ መታገዝ ማጠናቀቅ ያሳያችሁት ቁርጠኝነት እና ጽናት የሚደነቅ።

Figure 4: Example for Stop Word removal

Punctuation Removal

Accentuation in characteristic dialect gives the linguistic setting to the sentence. Punctuations such as a comma, might not include much esteem in understanding the meaning of the sentence.

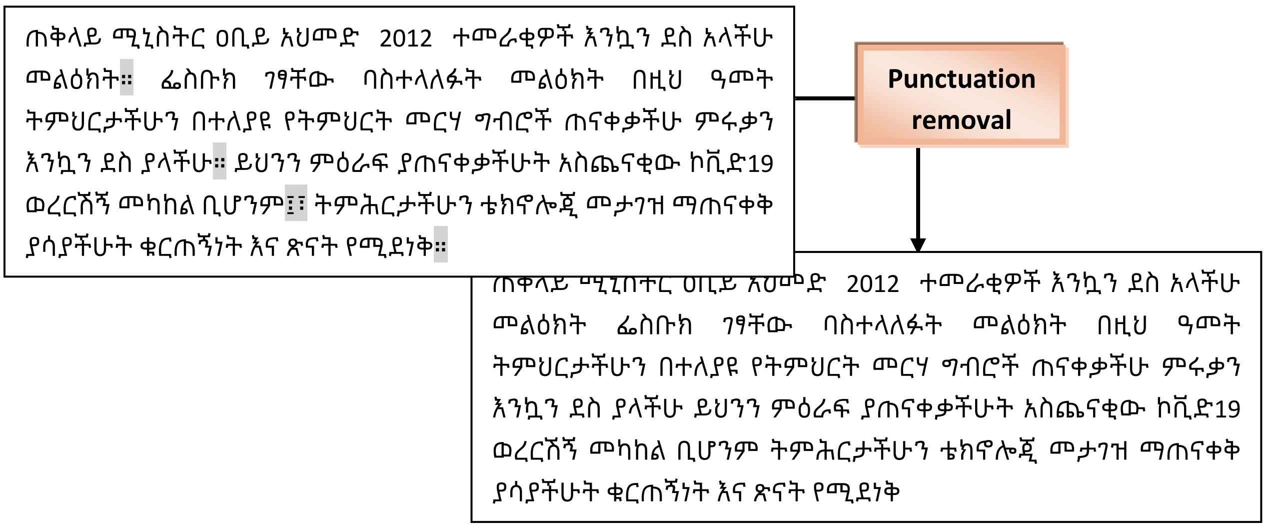


Figure 5: shows an example of Punctuation removal process.

3.1.3. Word Vector Representation

Planning the content from the body and feature of the news article for modeling is quite challenging. To perform content analytics, we have to be change over raw text into numerical features. We tested with two strategies to convert the crude content and feature extraction: Pack of Words and TF-IDF.

Bag of Word

The Pack of Words (BoW) method forms each news article as a report and calculates the recurrence number of each word in that archive, which is assist utilized to create numerical representation of the information, moreover called as vector highlights of fixed length. Sack of Words changes over crude content to word tally vector with Check Vectorizer function for highlight extraction. Check Vectorizer parts the content frame substance, builds the vocabulary and encodes the content into a vector. This encoded vector will have a count for events of each word that shows up more like a recurrence check as a key value pair. This strategy has disadvantages in terms of data misfortune. The relative position of the words isn't considered, and the data approximately the setting is lost. This misfortune can be costly now and then, compared to the pick up in computing simplicity with the ease the pack of words for real.

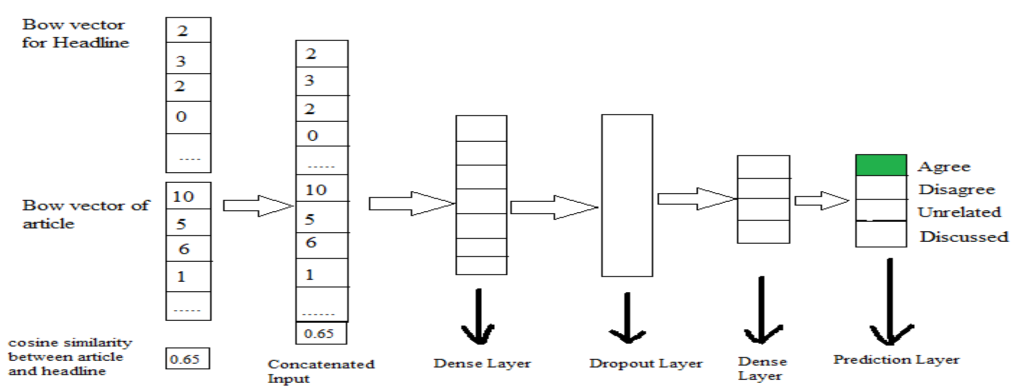


Figure 6: Bag of words vector with dense neural network architecture.

TF-IDF vectorizer

We have moreover utilized the procedure "Term Frequency-Inverse Record Frequency" (TF-IDF) for highlight extraction. Term Recurrence and Converse Record Frequency are two components of TF-IDF. Term Recurrence recognizes nearby significance of a word by its event in a record. Reverse Report Recurrence distinguishes the signature words, which are not showed up more frequently over the documents Word with a tall TF-IDF could be a signature word which is critical for the document in thought, has high recurrence within the report but isn't a common word across other texts.

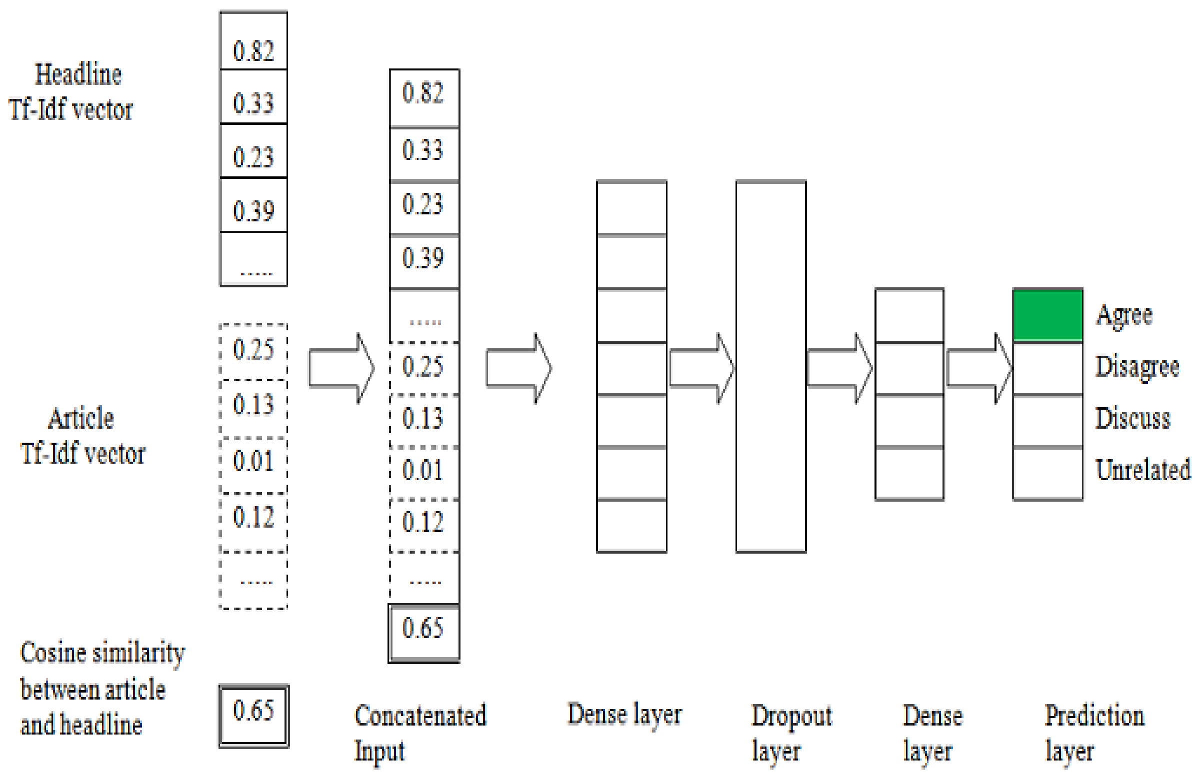


Figure 7: TF-IDF vector with dense neural network architecture.

3.1.4. Sampling Techniques

We part our information into Prepare, Approval and Test information sets. Since our dependent variable (Position) has four diverse classes (concur, oppose this idea, talk about and disconnected) and the information is uneven, we utilized stratified corresponding assignment and irregular rearranging to perform our information parts. We apportioned 67% of our Amharic news information to the Prepare set and the remaining 33% to the Test set. The preparing information is advance separated into approval sets (80/20 part). All of our experiments are conducted on preparing and approval sets in a 3-fold cross-validation setup.

3.1.5. Classification Process

It begins with preprocessing the information set, by evacuating pointless characters and words from the information. N-gram highlights are extricated, and a highlights lattice is shaped speaking to the records included. The final step within the classification handle is to prepare the classifier. We examined distinctive classifiers to anticipate the lesson of the records. We examined particularly six distinctive machine learning calculations, to be specific, Stochastic Slope Plunge (SGD), Bolster Vector Machines (SVM), Straight Back Vector Machines (LSVM), K-Nearest Neighbor (KNN) and Choice Trees (DT). We utilized executions of these classifiers from the Python Normal Dialect Toolkit (NLTK). We part the dataset into preparing and testing sets. For occurrence, within the tests displayed hence, we utilize 3-fold cross approval, so in each approval around 80% of the dataset is utilized for preparing and 20% for testing.

N-gram Model

N-gram modeling could be a prevalent highlight distinguishing proof and investigation approach utilized in dialect modeling and Natural dialect handling areas. N-gram could be a coterminous arrangement of things with length n. It may be a grouping of words, bytes, syllables, or characters. The foremost utilized n-gram models in content categorization are word-based and character-based n-grams. In this work, we utilize word-based n-gram to speak to the setting of the archive and produce highlights to classify the archive. We create a basic n-gram based classifier to distinguish between fake and genuine news articles. The thought is to produce different sets of n-gram recurrence profiles from the preparing information to speak to fake and honest news articles. We utilized a few standard n-gram highlights based on words and inspected the impact of the n-gram length on the exactness of distinctive classification calculations.

Output Sentence: $W = \{w_1 w_2 w_3 \dots w_{m_1} \mid m_1 \in (1, \infty)\}$
Reference Sentence: $W^r = \{w_1^r w_2^r w_3^r \dots w_{m_2}^r \mid m_2 \in (1, \infty)\}$
$\forall x \in (1, \infty)$, The Alignment of word w_x:

if $\forall y \in (1, \infty): w_x \neq w_y^r$ *// ∀ means for each, ∃ means there is/are*
  $(w_x \to \emptyset)$; *// → shows the alignment*
elseif $\exists!\, y \in (1, \infty): w_x = w_y^r$ *// ∃! means there exists exactly one*
  $(w_x \to w_y^r)$;
elseif $\exists y_1, y_2 \in (1, \infty): (w_x = w_{y_1}^r) \wedge (w_x = w_{y_2}^r)$ *// ∧ is logical conjunction, and*
  foreach $k \in (-n, -1) \cup (1, n)$
    foreach $j \in (-n, -1) \cup (1, n)$
      if $\exists k_1, k_2, j_1, j_2: (w_{x+k_1} = w_{y_1+j_1}^r) \wedge (w_{x+k_2} = w_{y_2+j_2}^r)$
        if $Distance(w_x, w_{y_1}^r) \leq Distance(w_x, w_{y_2}^r)$
          $(w_x \to w_{y_1}^r)$;
        else
          $(w_x \to w_{y_2}^r)$;
      elseif $\exists k_1, j_1: (w_{x+k_1} = w_{y_1+j_1}^r) \wedge (\forall k_2, j_2: (w_{x+k_2} \neq w_{y_2+j_2}^r))$
        $(w_x \to w_{y_1}^r)$;
      else *// i.e.* $\forall k_1, k_2, j_1, j_2: (w_{x+k_1} \neq w_{y_1+j_1}^r) \wedge (w_{x+k_2} \neq w_{y_2+j_2}^r)$
        if $Distance(w_x, w_{y_1}^r) \leq Distance(w_x, w_{y_2}^r)$
          $(w_x \to w_{y_1}^r)$;
        else
          $(w_x \to w_{y_2}^r)$;
else *// when more than two candidates, the selection steps are similar as above*

Figure 8: N-gram algorithm [2]

Classification Process

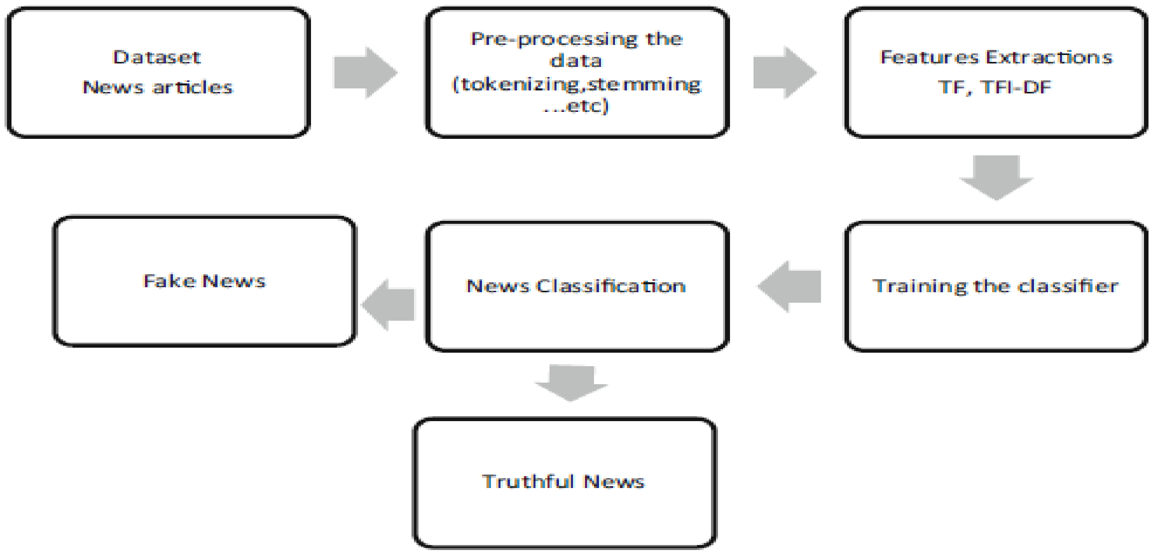


Figure 9: Classification Process [20]

Assume that… …

$$\Delta = [d_i]_{1 \le i \le m}$$

is our training set consisting of m documents

$$d_i$$

Utilizing one of the include extraction strategies (i.e., TF or TF_IDF), we calculate the highlight values comparing to all the terms/words included in all the records within the preparing corpus and select the p terms t_j (1≤j≤p) with the most elevated include values. Following, we construct the highlights network.

$$x = [x_{ij}]_{1 \le i \le m,\ 1 \le j \le p'}$$

Where: $$\begin{cases} x_{ij} = feature(t_j)\,if\ t_j \in d_i \\ x_{ij} = 0\ otherwise \end{cases}$$

In other words, x_ij compares to the highlight extricated (utilizing TF or TF-IDF) for term t_i for document d_i, such esteem is invalid (0) in case the term isn't within the record.

The multi classes of the machine learning algorithm have the following formula:-

$$Tpi = C_{ii} \; \forall i \; \in \; [1, N]$$

$$Fpi = \sum_{l=1}^{N} c_{li} - tp_i$$

$$Fni = \sum_{l=1}^{N} c_{il} - tp_i$$

$$Tni = \sum_{l=1}^{N} \; c_{lk} \; - \; tp_i - \; fp_i - \; fn_i$$

Finally what the algorithm will tell is, if a given claim is written in a similar way to a real news article or not it conclude and provide the following four arguments:

Agree **(አውነትነው):** The given claim is True

Unrelated **(ሀሰትነው):** The given claim is Fake or False

Discussed (**ሀሳቡ ተነስቷል):** The given claim has been raised

Disagree **(ሀሳቡተጋንኗል):** The truthiness of the given claim is bellow threshold

3.1.6. Evaluation metrics

In order to evaluate our multi-class classification model and compare between models, a confusion matrix is used.

	Predicted agree	Predicted disagree	Predicted discuss	Predicted unrelated
Labeled agree	True agree (TTA): Articles correctly classified as agree.	False disagree (FFD): Articles incorrectly classified as disagree	false discuss (FTD): Articles incorrectly classified as discuss	False unrelated (FFU): Articles incorrectly classified as unrelated
Labeled disagree	False agree (FTA): Articles incorrectly classified as agree	True disagree (TFD): Articles correctly classified as disagree	False discuss (FTD): Articles incorrectly classified as discuss	False unrelated (FFU): Articles incorrectly classified as unrelated
Labeled Discuss	False agree (FTA): Articles incorrectly classified as agree	False disagree (FFD): Articles incorrectly classified as disagree	True discuss (TTD): Articles correctly classified as discuss	False unrelated (FFU): Articles incorrectly classified as unrelated
Labeled unrelated	False agree (FTA): Articles incorrectly classified as agree	False disagree (FFD): Articles incorrectly classified as disagree	False discuss (FTD): Articles incorrectly classified as discuss	True unrelated (TFU): Articles correctly classified as unrelated

Table 5: Confusion matrix with explanation of outcomes from the confusion matrix in Table (3.4) collection of performance measurements can be calculated.

Performance measurements for each class

True Positive (TP): It refers to the number of predictions where the classifier correctly predicts the positive class as positive.

True Negative (TN): It refers to the number of predictions where the classifier correctly predicts the negative class as negative.

False Positive (FP): It refers to the number of predictions where the classifier incorrectly predicts the negative class as positive.

False Negative (FN): It refers to the number of predictions where the classifier incorrectly predicts the positive class as negative.

It's continuously superior to utilize disarray lattice as your assessment criteria for your machine learning demonstrate. It gives you a really straightforward, however proficient execution measures for our model. Here are a few of the foremost common execution measures able to utilize from the confusion matrix.

Accuracy: It gives you the overall accuracy of the model, meaning the fraction of the total samples that were correctly classified by the classifier. To calculate accuracy, use the following formula:

(TP+TN) / (TP+TN+FP+FN).

Misclassification Rate: It tells you what fraction of predictions was incorrect. It is also known as Classification Error. We calculate it using

(FP+FN) / (TP+TN+ FP+FN) or ***(1-Accuracy)***.

Precision: It tells you what fraction of predictions as a positive class were actually positive. To calculate precision, use the following formula:

TP / (TP + FP).

Recall: It tells you what fraction of all positive samples were correctly predicted as positive by the classifier. It is also known as True Positive Rate (TPR), Sensitivity, and Probability of Detection. To calculate Recall, use the following formula:

TP / (TP + FN).

Specificity: It tells you what fraction of all negative samples is correctly predicted as negative by the classifier. It is also known as True Negative Rate (TNR). To calculate specificity, use the following formula:

TN / (TN + FP).

F1-score: It combines precision and recall into a single measure. Mathematically it's the harmonic mean of precision and recall. It can be calculated as follows:

$$F_1\text{−score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}$$

Precision recall curve

Precision-recall bends plot the relationship between accuracy and review (or affectability). This bend centers on the model's capacity to recognize all the fake news articles, indeed on the off chance that this interprets into a better number of FP. A valuable outline measurement from the precision-recall bend is the AUPRC, which evaluates the capacity of the show to identify fake news articles. This can be thought of as an expectation of the proportion of fake news articles given a particular threshold, and is shown in Eq. (3.1).

An AUPRC yield break even with to the extent of genuine positives would compare to a arbitrary classifier. It has moreover been appeared that when identifying uncommon events (as is the case with fake news), the region beneath precision-recall bend (AUPRC) metric is ideal to the customary zone beneath bend (AUC) metric, which is the region beneath the Review vs Specificity bend, because it way better summarizes the prescient execution of the classifier.

$$\text{AUPRC} = \sum_n (R_n - R_{n-1}) P_n \qquad \textit{Eq. 3.1}$$

where *Pn* and *Rn* are the precision and recall at the n^{th} threshold [19].

3.2. PROPOSED SOLUTION

We test with distinctive word vector representation and neural arrange models as specified within the Strategies sub-section (3.13). Our best performing demonstrate takes Tf-Idf vector representation of words combined with preprocessed designed highlights as concatenated inputs and employments thick neural organize design to foresee the target position. The input highlights for our demonstrate comprises of Tf-Idf word vector representations of article-headline match, Cosine closeness between article-headline sets spoken to utilizing Tf-Idf, and cosine likeness between article-headline combine spoken to utilizing Google's word vectors (Word2Vec). We computed Tf-Idf scores on unigrams and bigrams. To maintain a strategic distance from inclination due to unbalance in dataset, we did not consider words which showed up in more than 50% of all preparing archives and prohibited the words which showed up in less than 50 reports.

Our model endeavors to capture the relative significance of a word show in article feature sets locally (how imperative is the word for that particular headline-article combine) and all inclusive (how common or unprecedented that particular word is in connection to all the words within the corpus). In arrange to capture the closeness between the feature article match, we calculated the Cosine similitude between Feature- Article Tf-Idf sets.

Given the measure of our lexicon and the uneven nature of our information, it is exceptionally simple for a neural organize show to over-fit, meaning incapable to anticipate precisely on the concealed test set. We conveyed L2, dropout and early ceasing as regularization methods to overcome overfitting and progress generalization. The neural arrange designs offer a wide assortment of hyper parameters; the table 6 presents the hyper parameters we centered on.

Hyper parameters	Experiment range	Choice
Final layer activation function	{ReLU, Tanh, Softmax}	Softmax
Batch size	32 – 256	64
Dropout rate	0 – 1	0.1
Epochs	50 – 200	50
L2 penalty	{0.1,0.01,0.001,0.0001}	0.0001

Table 6: Hyper parameter Tuning performance

CHAPTER 4

4. DATA COLLECTION AND ANALYSIS

4.1. Dataset Collection

4.1.1. Collecting dataset

Instead of succeeding this thesis research the first job to do is collecting news from different sources which are true and also false in different categories which means politics, health, social, entertainment and religion. Amharic fake news dataset is collected from different news agencies which are Amharic news reporters, since we are working in Amharic language we need Amharic news agencies like, ኤፍ ቢ ሲ, ቪኦኤ ዜና, ቢቢሲ ዜና, ኢቢሲ, Ethiopian Reporter and also our dataset is based on social media (Facebook) so we need to have some news from individuals who has much followers on Facebook like, ጌጡ ተመስገን, Elias Meseret, and so many others, totally our dataset have 4,000 number of rows collected from 25 Amharic news sources with 5 news categories.

Amharic fake news dataset is one of the great contribution of this thesis research, as we know in the world detecting fake news from social medias is a hot research area, for conducting those researches, there are many fake news dataset resources in English and other languages freely but not in Amharic language so we need to put this dataset as a concrete for other researchers on the area to conduct such researches as a base line. So we need to put this dataset as a main job of the research, of course after that we need to have the classification methodology.

While conducting Amharic fake news dataset we use python programming language to fetch data from news sites on social media. Using automatic data extract using Graph Api on python.

1. Extract the number of likes for the specific news site
2. Extract the number of news page followers
3. Gathering the fake and real news from Facebook page
4. Extract the number of replies for the specific news article
5. Extract the number of shares for the specific news article
6. Extract the number of comments for the specific news article
7. Extract publishing data of the article
8. Categorize the news article (politics, religious ….)

Fake news collection:

In Ethiopia our researchers are struggling to tackle fake-news problem but still there is no fact-checking websites, but there is one which says “fake news Ethiopia” on face book but not that much updated and active as needed. So that we have done a manual cross checking with Ethiopian journalists like Elias Meseret on the sites. After such tedious tasks we collect 2,000 fake news.

Real news collection:

This was the less demanding assignment. We accumulated posts of presumed news organizations, media news writers and indeed a few confirmed clients and bunches. We picked the news, which carried solid opinions (negative as well as positive), looking for higher consideration but were genuine. Hence, the dataset made, held likeness between Fake & Genuine news in term of gathering consideration. This was of course a noteworthy step to degree the execution of the show, since reactions to the news with negative assumption can make clients accept that it is Fake. We collected up to 5,000 news articles for the dataset, out of which 1546 were classified as Fake news and 3,454 as genuine news. The dataset comprised of 16 columns.

Amharic Fake news DataSet

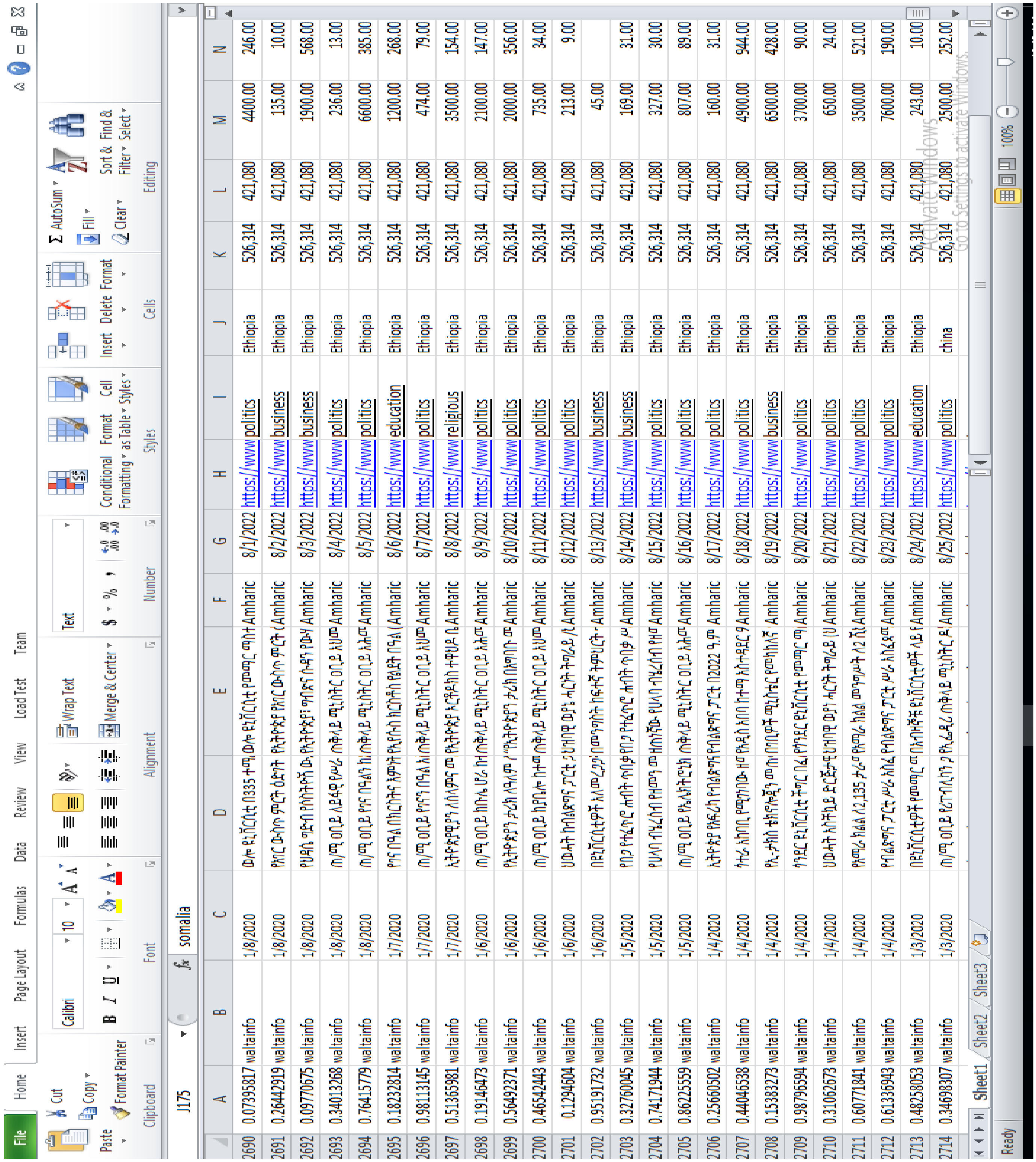
UID	Author	Published_date	Title	Text	Language	Crawled_date	Site_url	Category	Country	Page_follow	Page_likes	Replies_count	Comment	Shares	Type

Dataset attributes with their data type:

Attribute	***Data Type***
Uid	General
Author	Text
Published_date	Date
Title	Text
Text	Text
Language	Text
Crawled	date
Site_url	general
Category	text
Country	text
Page_follow	number
Page_likes	number
Replies_count	number
Comments	number
Shares	number
Stance	text
Type	text

Table 7: Dataset Attributes

Amharic Fake news DataSet snapshoot



Dataset description using tables

Name of the sites with their corresponding number of reviewed news articles both fake and fact news:

Site	Number of reviews
VOA Amharic	370
BBC Amharic	176
FBC	491
ETHIO_DJ	602
H-Axumit	22
WALTA	1350
WOLLO PRESS	34
Ethiopian Reporter	68
EBC	1350
WOLOTUBE	100
DIRETUBE	245
ETEGETUBE	82
GETU TEMESGEN	15
ELIASMESERET	95
Total	**5000**

News category	Label	Number of reviews
POLITICS	FACT/FAKE	2885
BUSINESS	FACT/FAKE	218
ENTERTAINMENT	FACT/FAKE	32
HEALTH	FACT/FAKE	1545
RELIGIOUS	FACT/FAKE	115
TECHNOLOGY	FACT/FAKE	30
EDUCATION	FACT/FAKE	19
SPORT	FACT/FAKE	39
SOCIAL	FACT/FAKE	117
Total		**5000**

Table 8: Dataset description

Number of news articles with their corresponding label:

News Label	Number of reviews
FACT	3454
FAKE	1546

4.2. Data Pre-Processing

Data preprocessing is a data mining technique that involves converting raw data into an understandable format. [15] Real-world data is often incomplete, inconsistent, and / or lacking in certain behaviors or trends, and it is likely that it contains many errors. Data preprocessing is a proven method for solving such problems. Data preprocessing is the main task that needs to be done before doing further work on the model.

After collecting our Amharic news dataset for our research we perform Stop Word Removal, Punctuation Removal and Stemming data preprocessing methods in python then Word Vector Representation has been done in Bag of Word method, which is explained well in third chapter Methodology chapter.

Labeling

In machine learning process data labeling is the base for the output performance. 80 percent of Our dataset are labeled manually (with humans) the reason is in our research the trusted news agencies news articles like Fana, ETV, Walta …etc. acts like true baseline for our model that we have going to apply to detect the fake news. But those trusted news agencies also disseminate fake or biased news which is our research limitation, and also for the reliability and the quality purpose, we have to do the labeling part manually by using manual data labeling tools like tagtog, brat, doccano… based on the news which is announced by Ethiopian trusted news agencies on social media web sites which is Facebook.

The outlines created for Amharic news articles are interpreted utilizing the Google Cloud Interpretation API. This cloud-based benefit interfaces specifically to Google's Neural Machine interpretation demonstrate (GNMT). The GNMT demonstrate employments a procedure called 'Zero-Shot Translation' to bypass the got to store the same data in numerous distinctive dialects (e.g. in a information base), and instep is prepared to get it the relationship between diverse dialects [15].

CHAPTER 5

5. RESULTS AND DISCUSSION

5.1. RESULTS

After a intensive hyper parameter tuning on our best performing show, we assessed the show on test information. Since our deliberate is to precisely degree the closeness of the anticipated position to the initial, we select 'classification accuracy' as our assessment metric. The anticipated exactness for the models portrayed in Segment 4 is displayed in Table 9, 10 and 11.

It begins with preprocessing the information set, by evacuating superfluous characters and words from the information. N-gram highlights are extricated, and a highlights framework is shaped speaking to the records included. The final step within the classification prepare is to prepare the classifier. We explored distinctive classifiers to anticipate the lesson of the records. We examined particularly six diverse machine learning calculations, specifically, Tf-Idf on unigrams and bigrams with cosine similitude encouraged into thick neural organize, BoW without unigrams and bigrams with cosine closeness bolstered into thick neural organize and Pre-trained implanting (Word2Vec) bolstered into thick neural organize. We utilized usage of these calculations from the Python Characteristic Dialect Toolkit (NLTK).

We separate our information into preparing, approval, and testing datasets. Since we number the variable (position) has four distinctive classes and the information isn't adjusted, we utilized it Stratified corresponding assignment and irregular rearranging for our information division strategy. We Devote 67% of our Amharic fake news information to the prepare set and the remaining 33% to the test gather. The preparing information is assist partitioned into approval bunches (part 80/20). Each show tests on preparing and approval sets are performed at 3 overlay approval demonstrate.

Accuracy for different Model variation

Variations	*Accuracy*
Tf-Idf on unigrams and bigrams with cosine similarity fed into dense neural network	**95.21%**
BoW without unigrams and bigrams with cosine similarity fed into dense neural network	**89.23%**
Pre-trained embedding (Word2Vec) fed into dense neural network	**75.67%**

Table 9: compares our model accuracy with the other models presented in the literature. Our model performs slightly better than the second best performing model.

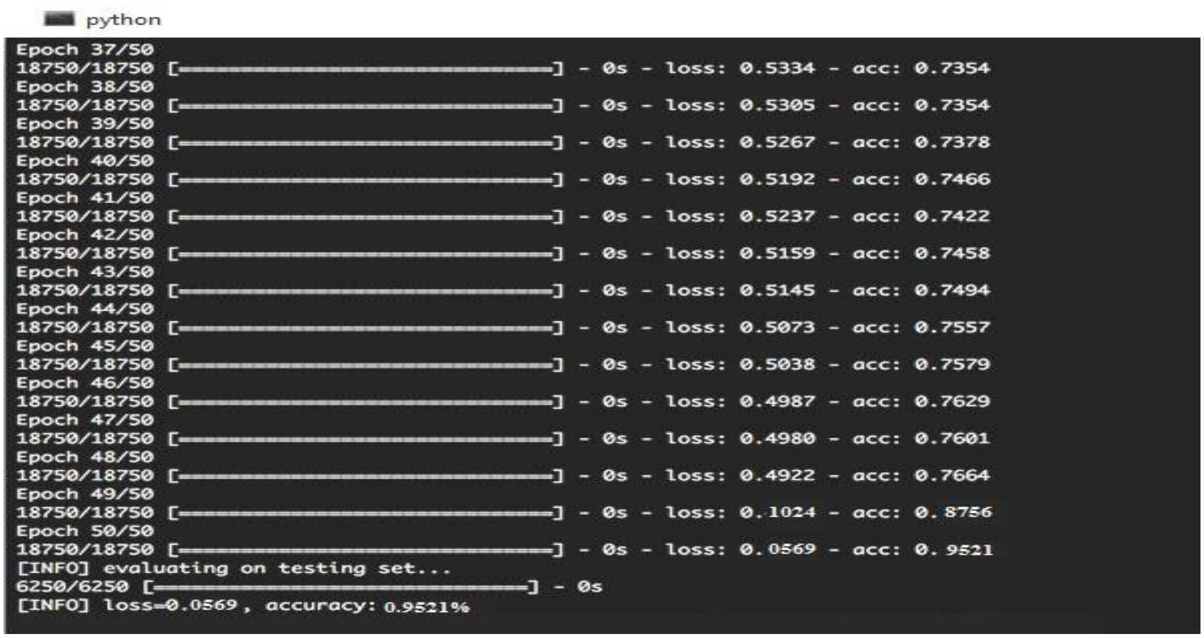
Model Description	**Accuracy**
Tf-Idf on unigrams and bigrams with cosine similarity fed into dense neural network (our model)	**95.21%**
BoW with multilayer perceptron [6]	**92.46%**
BoW with cosine similarity fed into dense neural network [5]	**88.46%**

Table 10: Comparison between best performing model and our model

We recognize that our dataset is an unequal dataset. It is exceedingly important for our show to perform sensibly well on minority positions. Table 11 summarizes the exactness comes about on test information for diverse position classifications.

Stance	Prediction Accuracy on Test data
Agree	77.33%
Discussed	90.06%
Disagree	52.38%
Unrelated	98.33%
Overall	**95.21%**

Table 11: Prediction Accuracy for different Stances



```
python
Epoch 37/50
18750/18750 [==============================] - 0s - loss: 0.5334 - acc: 0.7354
Epoch 38/50
18750/18750 [==============================] - 0s - loss: 0.5305 - acc: 0.7354
Epoch 39/50
18750/18750 [==============================] - 0s - loss: 0.5267 - acc: 0.7378
Epoch 40/50
18750/18750 [==============================] - 0s - loss: 0.5192 - acc: 0.7466
Epoch 41/50
18750/18750 [==============================] - 0s - loss: 0.5237 - acc: 0.7422
Epoch 42/50
18750/18750 [==============================] - 0s - loss: 0.5159 - acc: 0.7458
Epoch 43/50
18750/18750 [==============================] - 0s - loss: 0.5145 - acc: 0.7494
Epoch 44/50
18750/18750 [==============================] - 0s - loss: 0.5073 - acc: 0.7557
Epoch 45/50
18750/18750 [==============================] - 0s - loss: 0.5038 - acc: 0.7579
Epoch 46/50
18750/18750 [==============================] - 0s - loss: 0.4987 - acc: 0.7629
Epoch 47/50
18750/18750 [==============================] - 0s - loss: 0.4980 - acc: 0.7601
Epoch 48/50
18750/18750 [==============================] - 0s - loss: 0.4922 - acc: 0.7664
Epoch 49/50
18750/18750 [==============================] - 0s - loss: 0.1024 - acc: 0.8756
Epoch 50/50
18750/18750 [==============================] - 0s - loss: 0.0569 - acc: 0. 9521
[INFO] evaluating on testing set...
6250/6250 [==============================] - 0s
[INFO] loss=0.0569 , accuracy: 0.9521%
```

Figure 10: result screenshot

The confusion matrix in Figure 11 appears the misclassification rates for different Position sorts. As seen, the misclassification rate of ‘disagree’ lesson as ‘agree’ is tall and raises concern. The taken a toll of mistake can be taken into thought for basic assessment of such an blunder and is past the scope of this inquire about.

True label \ Predicted label	Agree	Disagree	Discuss	Unrelated
Agree	0.7733	0.0711	0.1116	0.0440
Disagree	0.3027	0.5238	0.1262	0.0473
Discuss	0.0296	0.0107	0.9006	0.0591
Unrelated	0.0012	0.0106	0.0049	0.9833

Accuracy = 0.9521 *Misclass* = 0.0569

Figure 11: Confusion Matrix for Final Model

Accuracy VS Dropout

Epoch	*Dropout*	*Validation Accuracy*
10		
30		
40		
50	*0.1*	*~ 95.21 (highest validation Accuracy)*
60	*0.6*	*~ 93.8 validation accuracy*
70		
100	*0.0*	*~ 91.7 (lowest validation Accuracy)*

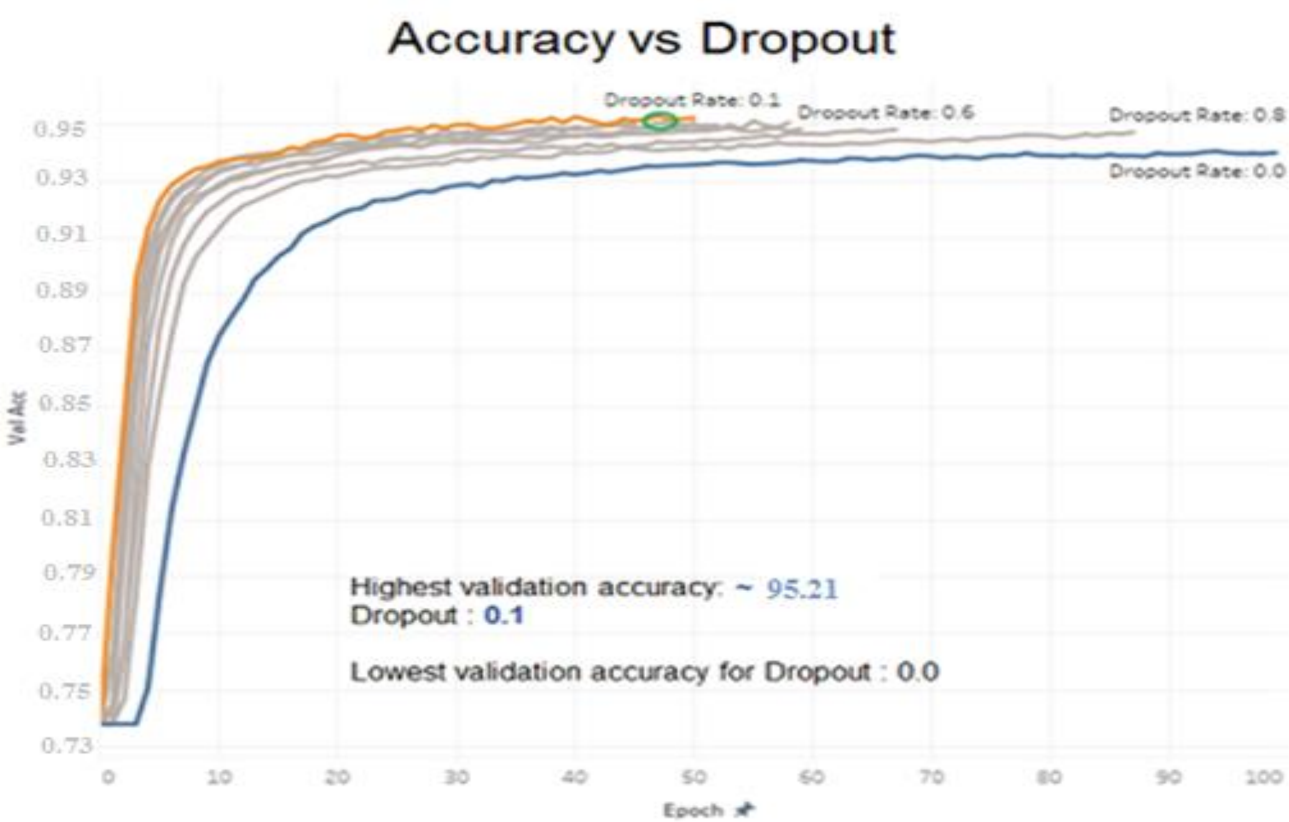


Figure 12: Accuracy for different epochs and dropout

5.2. DISCUSSION

Based on our fake news detection model results we can see some detail results. Based on the stance of the given claim which is agree, disagree, discuss or unrelated classes we can say that if any claim class is either agree or discuss then the claim is automatically Fact and if the claim class is predicted as disagree or unrelated then the given claim is Fake. Based on this conclusion in our country Ethiopia the fake news dissemination can be described as follows.

Fake news in Ethiopia in terms of news category:

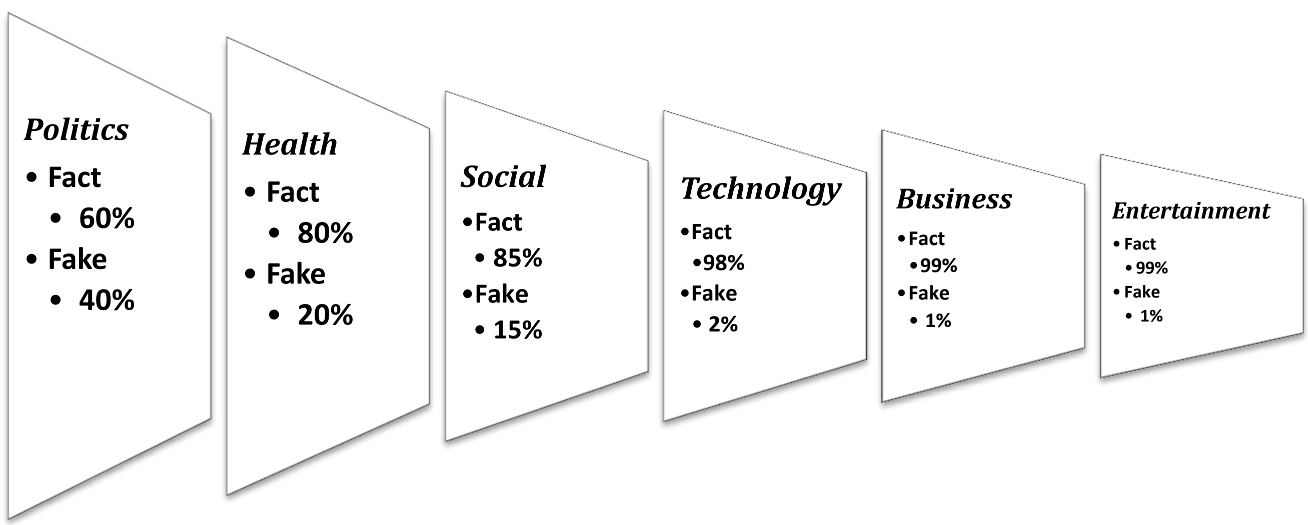
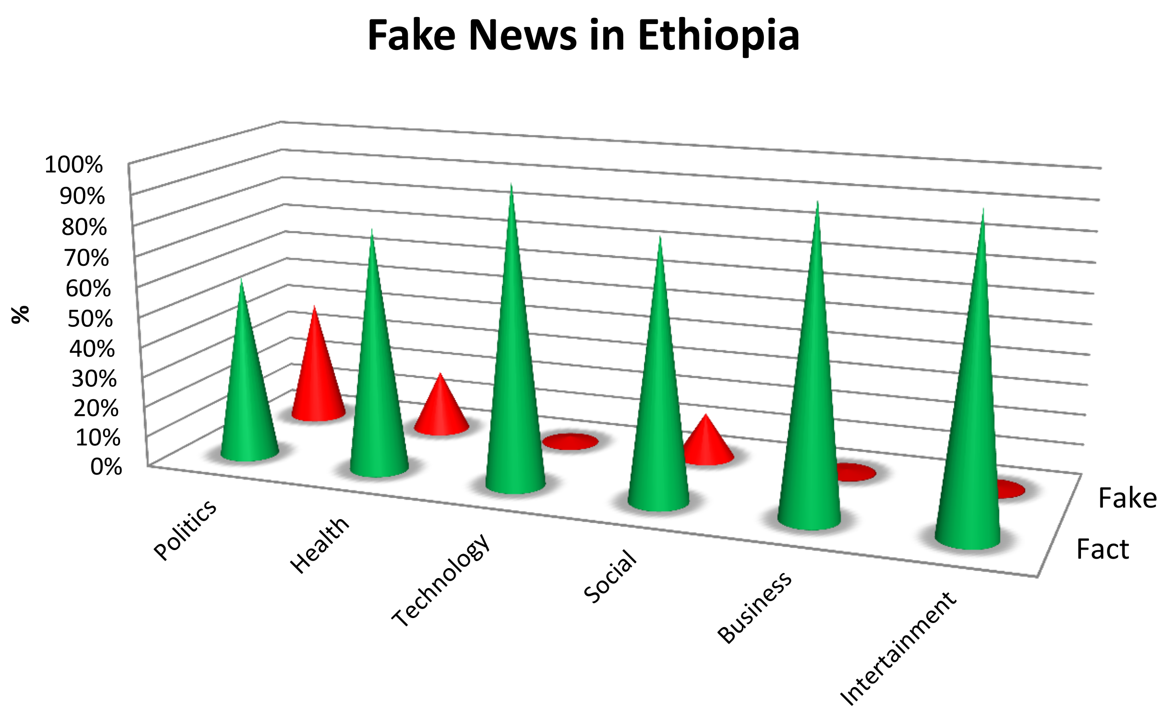


Figure 13: Fake news in Ethiopia in terms of news category



	Politics	Health	Technology	Social	Business	Intertainment
Fact	60%	80%	98%	85%	99%	99%
Fake	40%	20%	2%	15%	1%	1%

Figure 14: Fake news in Ethiopia in terms of news category

As we can see from figure 14, in our country Ethiopia most of the fake claims which are disseminated in social medias are politics news compared to other news categories, which leads our people to unnecessary conflict and some unwanted fight.

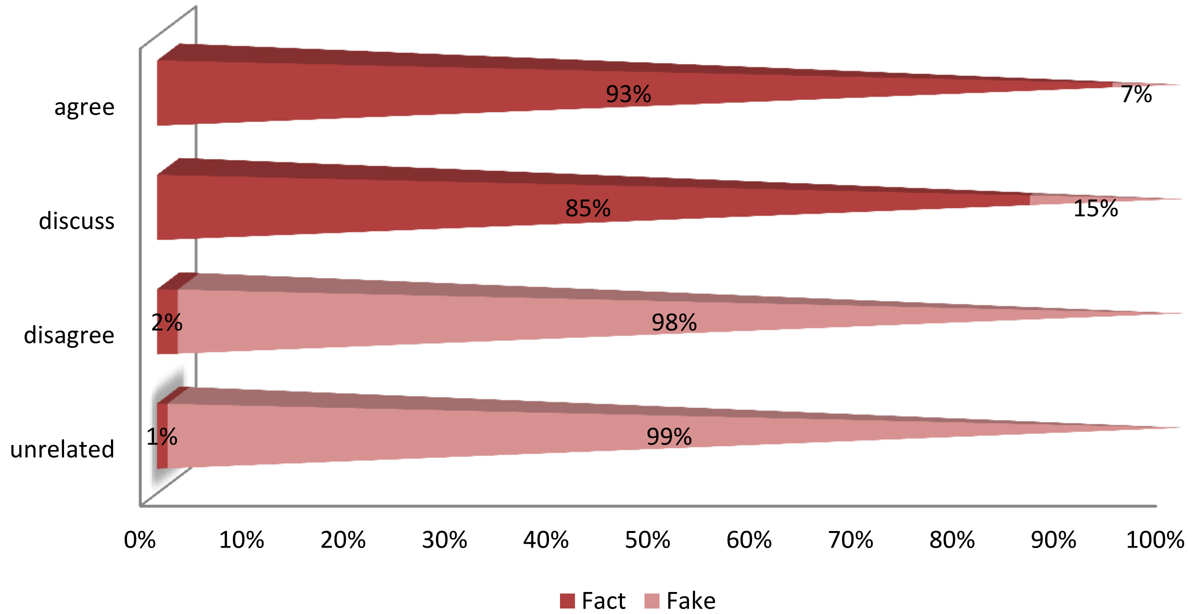


Figure 15: News stance Vs Truthiness

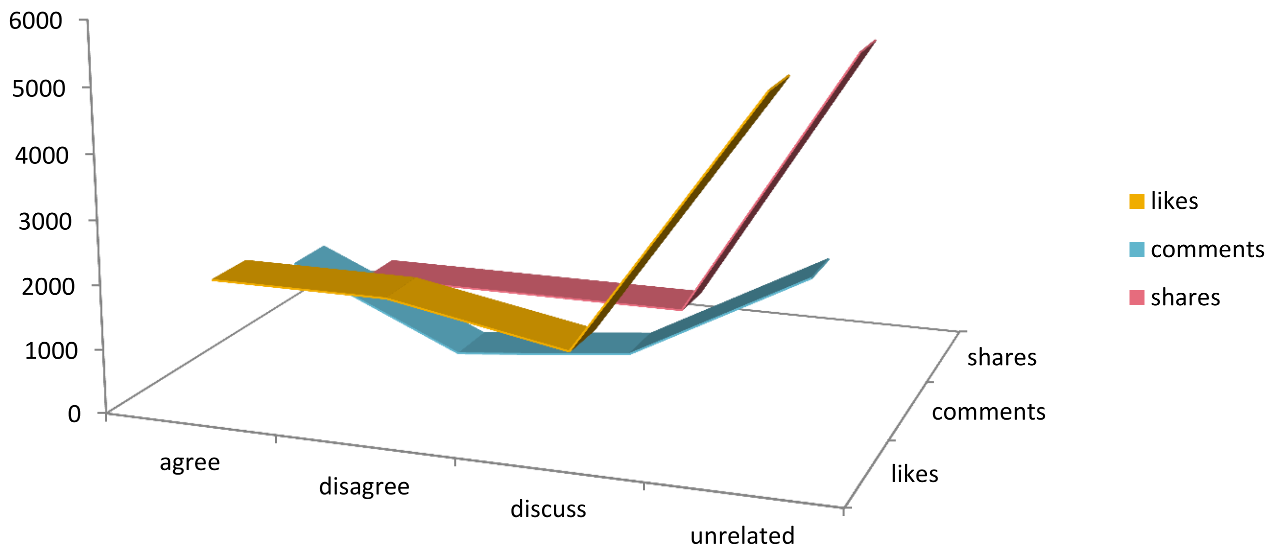


Figure 16: News stance Vs dissemination rate in social media (Facebook)

As we can see from the above figure 16, in our country Ethiopia, News with unrelated news class has the highest dissemination rate than the other news classes which means news with headline like **ሰበር ዜና** and **አሳዛኝ_ዜና** such news has a highest dissemination probability. As figure 15 shows that 99 percent of unrelated class news are totally fake only one percent is true so which means social media users in our country shares fake news broadly.

CHAPTER 6

6. CONCLUSION AND RECOMMENDATIONS

By employing an absolutely tuned Tf-IDF - Dense Neural Network (DNN) demonstrate, we able to accomplish an exactness of 95.12% within the test dataset. Our demonstrate works sensibly well when circumstances between the feature and news article are “unrelated”, “agree”, and “discuss”, but the forecast precision of the “disagree” position is moo (44%). The demonstrated modeled with BoW- DNN is our moment of best execution. It's exceptionally shocking to see that words when spoken to utilizing pre-trained word inserting such as Word2Vec reliably create lower resolutions when compared to the basic Tf-Idf and BoW representations.

There can be a few reasons for this marvel, counting the estimate of the news article. Word2Vec-based idle space representations may not be able to capture the significance of the semantic level of the word on the off chance that the length of the news article is exceptionally huge. Our methodology for calculating Tf-Idf vectors based on unigrams and bigrams has demonstrated to be exceptionally successful. Handcrafted passages such as cosine similitude between the news article and the feature moreover demonstrated to be a important include for our format section. We moreover tested with distinctive hyper parameters. We watched that by sending regularization procedures such as deserting, L2 regularization, cross approval and early end, we were able to realize a really liquid and steady learning handle.

At last, we need to expand this work by running a comparable examination on totally distinctive datasets like with diverse category. By categorizing fake news from social media stages, we trust to require a step closer to building an mechanized stage for identifying fake news. These consider gives a premise for future tests and broadens the run of arrangements that bargain with fake news discovery. Social media information will guarantee that contrasts in dialect are taken care of.

We need the other analysts to go more profound and assess the impacts of spreading the news on our peruses and make basic strategies for quicker forecast. Inquire about can borrow subjective models based on comparative errands by other disciplines and re-evaluate the include building and pre-treatment procedures utilized.

Appendix

Sample code

```
from data_loading.load_data import load_dataset
from feature_extraction.feat_extract import FeatureExtraction
from dimensionality_reduction.umap import get_umap_embedding
from dimensionality_reduction.tsne import get_tsne_embedding
from clustering.mean_shift import mean_shift_clustering
from graph_plots.plot_3d import scatter_plot_3d



if __name__ == "__main__":

    users_list, usernames_list, tweets_list, mentions_list, hashtags_list  =  load_dataset(
            dataset_path="stance_detect/datasets/twitter_covid.csv",
            features=["user_id", "username", "tweet", "mentions", "hashtags"],
            num_top_users=2000,
            min_tweets=10,
            random_sample_size=0,
            rows_to_read=None,
            user_col="user_id",
            str2list_cols=["mentions", "hashtags"])
    ft_extract = FeatureExtraction()
    user_feature_dict = ft_extract.get_user_feature_vectors(
            FEATURES_TO_USE,
            users_list,
            tweets_list,
            mentions_list,
              )
```

```
dataset = []
    with open(dataset_path, encoding="utf8") as csv_file:
        csv_file = DictReader(csv_file)

        for i,row in enumerate(tqdm(csv_file, desc="reading rows", leave=LEAVE_BAR),1):
            if features:
                out = tuple( [row[feat] for feat in features] )
                        dataset.append( out )
                    else:
                        dataset.append( row )

                    if i==rows_to_read:
                        break

            # Select random samples from the list
            if random_sample_size:
  try:
      dataset = sample(dataset, random_sample_size)
  except:
      raise ValueError(f"random_sample_size larger than dataset size: {len(output)} or negative !")
```

```
# Get data to plot from input_data dict
    feature_vectors, labels = zip(*list(input_data.values()))
    x,y,z = zip(* [tuple(x) for x in feature_vectors] )

    trace = [go.Scatter3d(x=x, y=y, z=z,
                          mode='markers',
                          hovertemplate = '%{text}',
                          text = [" ".join(x for x in item) for item in hover_info],
                          marker=dict(size=marker_size,
                                      color=labels,
                                      colorscale='Viridis',
                                      opacity=0.8)
                          )]
```

References

[1] Mapping the-Ethiopian-Media_2018

[2] 'Fake News' and Its Discontent in Ethiopia A Reflection June 2017

[3] Ethiopian Constitution

[4] https://www.thereporterethiopia.com/article/fake-news-alert

[5] The impacts of fake news on peace and development in the world: the case study of Ethiopia July, 2018

[6] http://science.sciencemag.org/content/359/6380/1146/tab-figures-data

[7] Fette, N. Sadeh, and A. Tomasic, "Learning to detect phis hing emails," in Proceedings of the 16th international conference on World Wide Web. ACM, 2007, pp. 649–656.

[8] Chao Yang, Robert Harkreader, Jialong Zhang, Seungwon Shin, and Guofei Gu. Analyzing spammers 'social networks for fun and profit: a case study of cyber criminal ecosystem on twitter. In Proceedings of the 21st international conference on World Wide Web, WWW '12, 2012.

[9] Niall J.Conroy,Victoria L.Rubin, Yimin Chen. deception detection : Methods for finding fake news. Crossref, 2015.

[10] Bajaj, Samir. " "The Pope Has a New Baby!" Fake News Detection Using Deep Learning."

[11] C. Shao and G. L. Ciampaglia. "Hoaxy: A platform for Tracking Online Misinformation." Mar. 2016.

[12] Shu, Kai, et al. "Fake news detection on social media:" ACM SIGKDD Explorations Newsletter (2017): 22-36.

[13] Janze, Christian, and Marten Risius. "Automatic Detection of Fake News on Social Media Platforms." (2017).

[14] W. Galuba, K. Aberer, D. Chakraborty, Z. Despotovic, and W. Kellerer. Outtweeting the Twitterers -Predicting Information Cascades in Microblogs. In Usenix Workshop on Online

[15] Ponnurangam Kumaraguru Anupama Aggarwal, Ashwin Rajadesingan. Phishari: Automatic realtime phishing detection on twitter. 7th IEEE APWG eCrime Researchers Summit (eCRS), 2012.

[16] Cook, David M., et al. "Twitter deception and influence: Issues of identity, slacktivism, and puppetry." Journal of Information Warfare13.1 (2014): 58-IV.

[17] Pisarevskaya, Dina. "Deception Detection in News Reports in the Russian Language: Lexics and Discourse." Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism. 2017

[18] shivam B. Parikh and pradeep K. Atrey "Media-Rich Fake News Detection: " ieeexplore.ieee.org/ 28 June 2018

[19]Automatic-Generation-of-Stopwords-in-the-Amharic-Miretie hedkar/d89a3f99c9495497eff7f8f043d8e6778c24e486

[20] https://www.google.com/search?client=firefox-b-d&q=classification+model

[21] https://www.kdnuggets.com/2017/12/general-approach-preprocessing-text-data.html

[22]] K. Jarmul, "Detecting fake news with scikit-learn," DataCampCommunity, 2017. [Online]. Available: https://www.datacamp.com/community/tutorials/scikit-learn-fake-news

[23] A. Thota, “Fake news detection: A deep learning approach,” 2018.[Online]. Available: https://scholar.smu.edu/datasciencereview/vol1/iss3/10

[24] M. Rao, “Exploiting network structure to detect fake news,” 2018 [Online]. Available: http://cs229.stanford.edu/proj2018/report/125

[24] https://journalcra.com/article/impacts-fake-news-peace-and-development-world-case-study-ethiopia

[25] https://www.researchgate.net/publication/347508623_Afaan_Oromo_Text_Content-Based_Fake_News_Detection_using_Multinomial_Naive_Bayes