ADDIS ABABA UNIVERSITY
COLLEGE OF NATURAL SCIENCES
DEPARTMENT OF COMPUTER SCIENCE

Deep Learning Based Emotion Detection Model for Amharic Text

Eyob Tesfu Merine

A Thesis Submitted to the Department of Computer Science in Partial Fulfillment For The Degree of Masters of Science in Software Engineering.

Addis Ababa, Ethiopia
2021

ADDIS ABABA UNIVERSITY
COLLEGE OF NATURAL SCIENCES
DEPARTMENT OF COMPUTER SCIENCE

Deep Learning Based Emotion Detection Model for Amharic Text

Eyob Tesfu Merine

Advisor: Ayalew Belay (Ph.D.)

Signed by the Examining Committee:

Name	**Signature**	**Date**
1. Ayalew. B (PhD),	Advisor _____	_____
2. Minale. A (PhD),	Examiner _____	_____
3. Mesfin. K (PhD),	Examiner _____	_____

Abstract

Emotions are so important that whenever we need to make a decision, we want to feel other's emotions. This is not only true for individuals but also for organizations. Due to the rapid growth of internet peoples expirees their emotions using different social media networks, reviews, blogs, online and so on. The need for finding relevant sources, extracts related sentences with emotion, summarizes them and organize them to useful form is becoming very high. Emotion detection can play an important role in satisfying these needs. The process of emotion detection involves categorizing emotional sentences into predefined categories such as sadness, anger, disgust, happiness, so on based on the emotional terms that appear within the comment. So that it's difficult to manually identifying emotion of a million of users and aggregating them towards a rapid and efficient decision is quite a challenging task due to the rapid growth of Amharic language usage in social media. In this research work, an emotion detection model is proposed for determining the emotion expressed in the Amharic texts or comment.

In this study, we proposed deep learning based emotion detection model for Amharic text using CNN with word embedding. The proposed model includes different tasks. The first task is text pre-processing which consists of commonly used text pre-processing tasks in many natural language processing applications. We perform text pre-processing in Amharic text and train the document using a word embedding in order to generate word embedding model. The embedding result provides a contextually similar word for every word in the training set then we implement our CNN model for emotion classification.

The common evaluation metrics such as accuracy, recall, F1 score and precision were used to measure our proposed model performance. Deep learning based emotion detection model for Amharic text prototype is developed and used to tests the system performance using the collected Amharic text comments. Finally, this study with four categories (sadness, anger, disgust, and happiness) of classification shows a result of 71.11% accuracy. Also did better when the number of classification is two (positive and negative) shows result of 87.46% accuracy. We also evaluate our model using RNN to compare with our CNN model.

Keywords – Emotion Detection, Pre-processing, Deep learning, CNN, RNN

Dedication

For My baby boy **“Kidus”**

Acknowledgment

I would like to thank Dr. Ayalew Belay, My supervisor, for his encouragement, patience and expert advice to finalize this thesis work. I would like to thank everybody that helped and supported me in one way or another. Special thanks to Samuel Tesfaye from our IT Staff for his valuable support in the course of the research.

Table of contents

List of Tables

List of Figures

List of Algorithm

Abbreviations

CNN	Amharic Text Convolutional Neural Network.
ANN	Artificial Neural Networks
CNN	Convolutional Neural Networks
CBOW	Continuous Bag of Word Method
DL4J	Deep Learning for Java
TF-IDF	Term Frequency- Inverse Document Frequency
FBC	Fana Broadcasting Corporation
FP	False Positives
FN	False Negatives
IDF	Inverse Document Frequency
KNN	K-nearest Neighbors
LSTM	Long Short-Term Memory
ML	Machine Learning
MLP	Multi-Layer Neural Networks
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
NR	Neural Network
NB	Naïve Bayes
RNN	Recurrent Neural Networks
SVM	Support Vector Machine
TF	Term Frequency
TP	True Positives
TN	True Negatives

Chapter 1: Introduction

1.1 Background

The Web is no longer as it was in the past. Recently, it supports direct interaction between users and its various resources and services such as social media networks, reviews, blogs, online commercial shops and more. The number of reviews especially for popular products grows rapidly. People´s emotions have become an important issue for making decisions, not only for individuals but also for government and commercial sectors. The abundance of information available on the Web raises the need for techniques that are able to analyze and make a better use of such huge information. Recently social media platforms have become a main medium for people to express their daily activities, reactions and emotions. Blogs and micro blogs are the most common form of social media. Blogs are informal sites on the worldwide web where users are used to post ideas and discussions thoughts on a particular issue.

In our daily life, every one of us faces different situations and the outcome of it is developing a feeling about it. Emotion is a strong feeling about human's situation or relation with others. It has a big role in customer decision in many domains including e-commerce, restaurants, movies, interests, and satisfaction with a service or a product. Moreover, Facebook added some reactions including angry, happiness, love, and surprise to allow users to express their emotions toward a comment, picture or an event. Emotion analysis is the area which deals with judgments, responses as well as feelings, which is generated from texts, being extensively used in fields like data mining, web mining, and social media analytics because emotions are the most essential characteristics to judge the human behavior. It becomes popular task in social network texts which express opinions and emotions on different issue to analyze and produce useful knowledge. Sentiment analysis models focus on polarity (positive, negative, neutral) but also on feelings and emotions (angry, happy, sad, etc.), and urgency (urgent, not urgent) [1].

Natural Language Processing used in the area of machine learning that focuses on the generation and understanding of language. Its main objective is to enable machines to understand, communicate and interact with humans in a natural way. NLP has many tasks such as Text Generation, Text Classification, Machine Translation, Speech Recognition, Sentiment Analysis, etc. Deep learning has emerged as a powerful machine learning technique that learns multiple layers of representations or features of the data and produces state-of-the-art prediction results.

Along with the success of deep learning in many other application domains, deep learning is also popularly used in emotion analysis in recent years.

Processing Amharic language text using different techniques and models has become popular field of study on social media. The sentiment of a post can be seen in the tone or emotion conveyed in a text. We can understand exactly how people feel about political, social, religious situations in the country, by effortlessly analyzing the emotion of each and every post and monitoring how the community is responding to contents. A proper social media emotion analysis could categorize social media mentions into the right category. Social media emotion analysis uses some revolutionary machine learning and deep learning algorithms and analyses the text posted online. Emotion analysis is a challenging task especially for languages having low resources and complex linguistic structures like Amharic. A small amount of resources like training data for emotion analysis highly impacts the accuracy of the system. The complex structure of languages also needs the design of important features and the best combination from these features. Amharic as one of low resourced and morphologically rich languages shares the above challenges [2].

The increasing usability of social media creates a difficult of emotion analysis to manually identifying emotion of a million of users and aggregating them towards a rapid and efficient decision. In terms of data required to train and test the model the deep learning efforts to learn high level abstraction by exploiting the hierarchical architectures. It is a promising approach and has been extensively applied learning, semantic parsing, natural processing and many more. We select deep learning approach because of improved ability of chip processing, extensive lower expenditure of hardware and significant enhancement in machine learning algorithm. Our goal is to investigate which deep learning methods are best suited to solve our problem.

1.2 Motivation

The main motivation of this work came from the popularity of emotion analysis in foreign language due to its various useful applications in a different area. However, there is lack of efficient emotion detection model for Amharic. The related work done on this area for Amharic language [3-9] that only determining the polarity of the opinion it can be positive, negative or neutral sentiment differentiation but didn’t describe the exact feeling of the public and the intensity of their reaction, emotion detection performs whether the public reaction with different emotional states happiness, sadness, fear, anger, surprise, disgust etc.

Amharic being the official language of Ethiopia which is spoken by a substantial segment of the population and it is ranked as the second language that has native speakers. In the 2007 census, 21.6 million speak Amharic which is 29.33% of the population [10]. The language is used in business, government and education. Also, about 16,037,811 Internet users on June/2017, 15.4% of the Ethiopian population per ITU 4,500,000 Facebook users on June/2017, 4.3% penetration rate [11].So most of the users share their emotion in different social media using Amharic language this has changed the manner in which people communicate and influence behavior and emotion of other people in the social network. Emotions of users that are expressed on the web have great influence on the readers, product vendors and politicians. The unstructured form of data from the Amharic social media is needed to be analyzed and well-structured for this purpose emotion detection has recognized significant attention. Therefore, it's critical importance to analyses and identifies emotion detection for Amharic text on the internet.

1.3 Statement of the problem

The growing use of Internet as a tool to help governments, companies and organizations do their daily works and gives the companies the opportunity to providing online shopping websites and using social media to increase their sales. In this new era, besides the content offered by companies and organizations, individuals have come to share their emotion, knowledge, experiences ,reviews and opinions to the world via personal blogs, networking sites, and micro blogs, Facebook pages, twitter just to name a few.

So that it's difficult to manually identifying emotion of a million of users and aggregating them towards a rapid and efficient decision is quite a challenging task due to the rapid growth of Amharic language usage in social media. Language is used not only for communication but also for imparting emotion associated with it. People express their emotion in different way, in video and photos, whether it can be speech or in writing. Written text is one of a good source for expressing of public views, ideas, emotions, and filling. This study aimed to develop Amharic text emotion detection model in order to tackle this challenge [12-18].

Works on emotion analysis and social media on different languages exist but as Amharic is a low-resource language, in regards to digitization, there is little attempt on this area. Therefore, in this work the performance of word2vec and deep learning algorithms on Amharic language is analyzed and explored. Qualitative analysis using deep learning algorithms, determining word analogies using word2vec and analysis emotion on social media based on prediction of the

algorithms. Investigating the impact of automatically generated word vector features in Amharic emotion detection task and comparing different deep learning algorithms for Amharic emotion analysis and new emotion detection model based on deep learning are focus of this paper.

1.4 Objectives

General

The general objective of this thesis is to develop an emotion detection model for Amharic text using deep learning.

Specific Objectives

To achieve the above general objective, the following specific objectives are identified.

- Conduct a literature review on emotion detection.
- Preparing emotional dataset.
- Select appropriate deep learning technique using systematic review.
- Identify basic components and characteristics of emotion detection model.
- Develop emotion detection model.
- Evaluate the performance of the model through prototyping.

1.5 Methods

In order to accomplish the objective of the research, we use design science research methodology, which offers specific guidelines for evaluation and iteration within research. It focuses on the development and performance of designed with the explicit intention of improving the functional performance of the design. In this model all design begins with awareness of a problem. This study will first conduct a comprehensive review of literatures to understand the research area and its problem domains [19].

Based on this understanding, we identify the importance of the works done in the area and analysis of the different technique, approaches and algorithms suitable and applicable for emotion detection. This designation emphasizes the problem-solving or performance-improvement of the model then suggestions for a problem solution are drawn from related works done on this research area by assessing to identify and point direction in order to provide solution to identified problems or improve the performance from the existing knowledge base for the problem area.

Finally, based on the literature survey we select appropriate tools, techniques and algorithms for our problem. After identifying those requirements we will design our model and identify its components for an attempt in implementing an artifact according to the suggested solution is performed. Next preparing data properly, it's an important and critical step in our research for data analysis and it has an immense impact on the success of our research. We use systematic sample techniques. Therefore, we need to collect data from different official Facebook pages for partially or fully successful implementations and evaluation of the system using some performance evaluation metrics like precision, F1 score, recall and accuracy. The overall methodology is to produce artifacts including model, prototype and algorithm we also evaluate the performance of the model through prototype.

1.6 Scope and limitation

The scope of the study is to analysis and develops an emotion detection model for Amharic text using deep learning technique. In social media, people generally use informal or semi-formal language for communication. In everyday life conversation, people do not care about the spellings and accurate grammatical construction of a sentence which makes emotion detection a complex task, the following are considered as the limitation of this research work:-

- Specific comments on a post in textual contents only not include abbreviations.
- Limited to Amharic texts comment for specific post on social network.
- Limited on four emotion state category (anger, disgust, happiness, and sadness).

1.7 Application of result

It's estimated that 80% of the world's data is unstructured, in other words it's unorganized [20]. A huge volume of text data is created every day emails, support tickets, chats, social media conversations, surveys, articles, documents, etc. But it's hard to analyze, understand, and sort through, time consuming and expensive. Emotion detection techniques have enabled to make sense or emotion of big social media data to make more informed decisions and understand social events, product marketing or political events or tendencies. The applications of emotion detection are endless. Public opinion detection though emotion or feelings about almost anything in the social network. Just to name few of them:-

- Emotion detection about political tendencies.
- Tracking product perception.

- Tracking customer's feedback.
- Market research and product analytics so on.

1.8 Organization of the thesis

This thesis is organized into five chapters including the current one:

- Chapter Two introduces an overview of emotion classification and the different techniques used in sentiment analysis researches. Moreover, the general steps in sentiment analysis are also discussed in this chapter.
- Chapter Three presents reviews of related researches conducted on sentiment analysis. In this chapter, an overview and definition of emotional words, in-depth reviews of researches done on emotion detection using different techniques for different languages is presented.
- Chapter Four describes the proposed model for the Amharic text emotion detection. In addition, implementation-related issues such as pre-processing, stop word removal, light steaming and classification are also explained in the same chapter.
- Chapter Five presents the experimental results of the proposed model in general and the different algorithms in particular.
- Finally, future works, recommendations, and conclusions are given in the last chapter.

Chapter 2: Literature Review

2.1 Introduction

In literature review chapter, various studies of sentiment analysis both on local language and other language using different techniques are discussed. Also, the current state of the art on sentiment analysis and clearly define what are the gap of Amharic sentiment analysis we done extensive review on different thesis.

2.2 Sentiment analysis

Sentiment analysis is a term that refers to the use of natural language processing, text analysis, and computational linguistics in order to ascertain the attitude of a speaker or writer toward a specific topic. Basically, it helps to determine whether a text is expressing sentiments that are positive, negative, or neutral. Sentiment analysis is an excellent way to discover how people, particularly consumers, feel about a particular topic, product, or idea.

The origin of sentiment analysis can be traced to the 1950s, when sentiment analysis was primarily used on written paper documents. Today, however, sentiment analysis is widely used to mine subjective information from content on the internet, including texts, tweets, blogs, social media, news articles, reviews, and comments. This is done using a variety of different techniques, including NLP, statistics, and machine learning methods [21]. Sentiment analysis can occur at different levels document level, sentence level and aspect/feature level [22].

Document level classification

In this process, sentiment is extracted from the entire review, and a whole opinion is classified based on the overall sentiment of the opinion holder. The goal is to classify a review as positive, negative, or neutral. Document level classification works best when the document is written by a single person and expresses an opinion/sentiment on a single entity.

Sentence level classification

This process usually involves two steps the first subjectivity classification of a sentence into one of two classes' objective and subjective and second sentiment classification of subjective sentences into two classes positive and negative. An objective sentence presents some factual information, while a subjective sentence expresses personal feelings, views, emotions, or beliefs. Subjective sentence identification can be achieved through different methods such as Naïve Bayesian classification and different deep leering methods. However, just knowing that

sentences have a positive or negative opinion is not sufficient. This is an intermediate step that helps filter out sentences with no opinions and helps determine to an extent if sentiments about entities and their aspects are positive or negative. A subjective sentence may contain multiple opinions and subjective and factual clauses. Sentiment classifications at both the document and sentence levels are useful.

Feature level classification

In this process, the goal is to identify and extract object features that have been commented on by the opinion holder and determine whether the opinion is positive, negative, or neutral. Feature synonyms are grouped, and a feature-based summary of multiple reviews is produced. Words express various kinds of sentiments that may be positive, negative, strong, or weak.

To perform sentiment analysis, it is important to understand the polarity of words and classify sentiments into categories such as positive, negative, or neutral. This task can be accomplished through the use of different types of sentiment. Let's look at the most important types of sentiment analysis.

Fine-grained

This sentiment analysis model helps us to derive polarity precision. We can conduct a sentiment analysis across the following polarity categories very positive, positive, neutral, negative, or very negative. Fine-grained sentiment analysis is helpful for the study of reviews and ratings. For a rating scale from 1 to 5, you can consider 1 as very negative and five as very positive. For a scale from 1 to 10, you can consider 1-2 as very negative and 9-10 as very positive.

Aspect-based

While fine-grained analysis helps you determine the overall polarity of your customer reviews, aspect-based analysis delves deeper. It helps you determine the particular aspects people are talking about. With aspect-based analysis, you can determine that the reviewer has commented on something negative about the review.

Emotion detection

As the name suggests, emotion detection helps you detect emotions. This can include anger, sadness, happiness, frustration, fear, worry, panic, etc. Emotion detection systems typically use a collection of words that convey certain emotions. Some advanced classifiers also utilize robust ML algorithms. It's recommended to use ML over lexicons because people express emotions in a myriad of ways. Take this line, for example: "This product is about to kill me." This line may

express feelings of fear and panic. A similar line “this product is killing it for me” has an entirely different and positive meaning. But the word “kill” might be associated with fear or panic in the lexicon. This may lead to inaccurate emotion detection.

Intent analysis

Accurately determining consumer intent can save companies time, money, and effort. So many times, businesses end up chasing consumers that don’t plan to buy anytime soon. Accurate intent analysis can resolve this hurdle. The intent analysis helps you identify the intent of the consumer whether the customer intends to purchase or is just browsing around. If the customer is willing to purchase, you can track them and target them with advertisements. If a consumer isn’t ready to buy, you can save your time and resources by not advertising to them. Modern-day sentiment analysis approaches are classified into three categories knowledge-based, statistical, and hybrid.

Sentiment analysis algorithms

There are three major Sentiment Analysis methods.

- **Rule-based approach**: is based on an algorithm with a clearly defined description of an opinion to identify. Includes identify subjectivity, polarity, or the subject of opinion.
- **Automatic approach**: is more of a toy than a real tool, automated sentiment analysis is the real deal. It is the one approach that truly digs into the text and delivers the goods. Instead of clearly defined rules this type of sentiment analysis uses machine learning to figure out the gist of the message. Because of that, the precision and accuracy of the operation drastically increase and you can process the information on numerous criteria without getting too complicated.
- **Hybrid approach:** that combines both rule-based and automatic approaches.

Sentiment analysis metrics and evaluation

There are many ways in which we can obtain performance metrics for evaluating a classifier and to understand how accurate our emotional analysis model is. One of the most frequently used is known as cross-validation.

What cross-validation does is splitting the training data into a certain number of training folds (with 75% of the training data) and the same number of testing folds (with 25% of the training data), use the training folds to train the classifier, and test it against the testing folds to obtain performance metrics. The process is repeated multiple times and an average for each of the metrics is calculated.

If the testing set is always the same, it might be over fitting to that testing set, which means it might be adjusting the analysis to a given set of data so much that it might fail to analyze a different set. Cross-validation helps prevent that. The more data we have, the more folds we will be able to use.

Precision, Recall, F1 score and Accuracy

Precision, recall, and accuracy are standard metrics used to evaluate the performance of a classifier.

- **Precision** measures how many texts were predicted correctly as belonging to a given category out of all of the texts that were predicted (correctly and incorrectly) as belonging to the category.
- **Recall** measures how many texts were predicted correctly as belonging to a given category out of all the texts that should have been predicted as belonging to the category. We also know that the more data we feed our classifiers with, the better the recall will be.
- **Accuracy** measures how many texts were predicted correctly (both as belonging to a category and not belonging to the category) out of all of the texts in the corpus.
- **F1 score** is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. Intuitively it is not as easy to understand as accuracy, but F1 is usually more useful than accuracy, especially if you have an uneven class distribution.

Accuracy works best if false positives and false negatives have similar cost. If the cost of false positives and false negatives are very different, it's better to look at both Precision and Recall.

Most frequently, precision and recall are used to measure performance since accuracy alone does not say much about how good or bad a classifier is. For a difficult task like analyzing emotion, precision and recall levels are likely to be low at first. As we feed the classifier with more data, performance is improved.

2.3 Amharic Language

Amharic which belongs to the Semitic family of languages, is written in the unique and ancient Ethiopic script inherited from Geez language. Amharic is the second most spoken Semitic language in the world, next to Arabic. Amharic language serves as the working language of Ethiopia, and is also the working language of several of the states within the Ethiopian federal

system. With 21,811,600 total speakers as of 2007, including around 4,000,000 second language speakers, Amharic is the second-most common language of Ethiopia (after Oromo) and second-most commonly spoken Semitic language in the world (after Arabic) [23].

Unlike Arabic and Hebrew, Amharic is written left-to-right using a system that grew out of the Geʿez script. The writing system is called fidäl (**ፊደል**) in Ethiopian Semitic languages. Fidäl means "script", "alphabet", "letter", or "character". The writing system is also called abugida (**አቡጊዳ**), from the first four symbols; from this the modern term abugida is derived [23].

Recently, the introduction of Amharic Unicode font and its integration in different technologies has paved a way for many online publishers and users to interact using the native Ethiopic script. This trend of Amharic being used as a medium of online communication among speakers is contributing to the enrichment of the web with Amharic contents thereby opening further study opportunities for the language. This is because it lacks basic tools and resources for carrying out natural language processing research and applications.

Amharic has thirty four consonants with seven vowels. Redundancy of Amharic symbols results in some confusion of written Amharic words. For example tsibay (**ፀባይ**, **ጸባይ**) literal meaning “conduct”; hymanot (**ሀይማኖት**, **ኃይማኖት**, **ሐይማኖት**, **ሃይማኖት**) literal meaning “religion” are various writing forms of a single word with different symbols existing in the Amharic language.

Across the time, social media keeps growing, pioneered by Friendster which dominated the usage of Amharic language on social media. The language is used in business, government and education. Also, about 16,037,811 Internet users Amharic language on June/2017, 15.4% of the Ethiopian population per ITU 4,500,000 Facebook users on June/2017, 4.3% penetration rate [24]. So most of the users share their emotion in different social media using Amharic language this has changed the manner in which people communicate and influence behavior and emotion of other people in the social network. Emotions of users that are expressed on the web have great influence on the readers, product vendors and politicians. The unstructured form of data from the Amharic social media is needed to be analyzed and well-structured for this purpose emotion detection has recognized significant attention.

2.4 Defining emotional data

Social media platforms, such as Facebook, Instagram, Twitter or others, are now part of everyday life for thousands of people who use them to share their emotion, knowledge, experiences, reviews, opinions and photos of a vacation, or from a special event using Amharic language. Social media and its content also increasingly serve as a source of information, and have even replaced traditional media such as television or newspapers. Thus social media content can play an important role in a government and company's overall content marketing strategy. Content on social media channels can have a considerable reach, and because of its close proximity to the user can often prove more effective than classic advertising. In certain circumstances, comments on social media can even boost Google rankings. However, social media content also follows its own rules, and therefore requires its own dedicated strategy. The high degree of interaction is a special feature of social media users can give direct feedback or comment on specific content by comments or something similar. In addition, by sharing content, users can directly contribute to the spreading of emotional content right across social media.

Humans are inherently emotional beings. They like to feel and, when they do, they also like to share it with their friends and family across the social networks. That's the most fundamental reason why social media networks thrive. The underlying reason is because they instigate the reader to feel a particular emotion strongly. Maybe it's making our target audience jubilant. Maybe it makes them sad or even angry. There are many different types of emotions that have an influence on how we live and interact with others. At times, it may seem like we are ruled by these emotions. The choices we make, the actions we take, and the perceptions we have are all influenced by the emotions we are experiencing at any given moment [25].

Psychologists have also tried to identify the different types of emotions that people experience. A few different theories have emerged to categorize and explain the emotions that people feel. The basic types of emotions are happiness, sadness, disgust, fear, surprise, and anger. In our research context, we define any textual message or comments calcified in the following emotional categories.

Happiness

Given textual comments on social media is defined as happiness if its intent is any of the following. If the comment describe a pleasant emotional state that is characterized by feelings of contentment, joy, gratification, satisfaction, and well-being. This type of emotion is sometimes

expressed through Facial expressions such as smiling, Body language such as a relaxed stance and Tone of voice an upbeat, pleasant way of speaking.

Sadness

Given textual comments or contents on social media is defined as sadness if its intent is any of the following. If the content has a type of emotion that has a transient emotional state characterized by feelings of disappointment, grief, hopelessness, disinterest, and dampened mood. Sadness can be expressed in a number of ways including Crying, Dampened mood, Lethargy, Quietness, Withdrawal from others

Fear

Given textual comments or contents on social media is defined as fear if its intent is any of the following. The contents that expires some sort of danger and experience fear, you go through what is known as the fight or flight response. Expressions of this type of emotion can include Facial expressions such as widening the eyes and pulling back the chin, Body language attempts to hide or flea from the threat, Physiological reactions such as rapid breathing and heartbeat.

Disgust

Given textual comments or contents on social media is defined as disgust if its intent is any of the following. Disgust can be displayed in a number of ways including Body language turning away from the object of disgust, Physical reactions such as vomiting or retching and facial expressions such as wrinkling the nose and curling the upper lip.

Anger

Given textual comments or contents on social media is defined as anger if its intent is any of the following. The content characterized by feelings of hostility, agitation, frustration, and antagonism towards others. Like fear, anger can play a part in your body's fight or flight response. Anger is often displayed through Facial expressions such as frowning or glaring , Body language such as taking a strong stance or turning away, Tone of voice such as speaking gruffly or yelling, Physiological responses such as sweating or turning red and Aggressive behaviors such as hitting, kicking, or throwing objects.

Surprise

Given textual comments or contents on social media is defined as surprise if its intent is any of the following. Contents quite brief and is characterized by a physiological startle response following something unexpected. This type of emotion can be positive, negative, or neutral.

Surprise is often characterized by Facial expressions such as raising the brows, widening the eyes, and opening the mouth, Physical responses such as jumping back and Verbal reactions such as yelling, screaming, or gasping.

Based on the above definitions, when we say emotion, implicitly we are talking about every context that falls into one or more of the defined cases. Characterizing, certain expression as different emotion have an important role in identifying exact feeling of the public and the intensity of their reaction. We consider Amharic text collected from social media, which can provide information having utility in a variety of ways, especially opinion mining. Social media like Facebook is full of emotions, feelings and opinions of people all over the world. However, analyzing and classifying Amharic text on the basis of emotions is a big challenge and can be considered as an advanced form of Sentiment Analysis.

2.5 Types of emotion detection modalities

Emotion detection is probably to gain the best outcome if applying multiple modalities by combining different objects, including text, audio, video, and physiology to detect emotions. In general, five types of emotion detection modalities they are emotion from speech, text, facial expression, body gesture and movement and physiological state [26].

Emotion from speech

Detecting emotion from speech has become one the active research themes in speech processing and in applications based on human-computer interaction and it is one of the exploitable channels for gathering emotional information from the user of a system is their voice. When a person starts talking, they generate information in two different channels primary and secondary [26]. The primary channel is linked to the syntactic-semantic part of the talking (what the person is literally saying), while the secondary channel is linked to paralinguistic information of the speaker (tone, emotional state, and gestures). E.g., someone says “That’s so funny” (primary channel) with a serious tone (second channel). By looking at the information of the primary channel, we get the message that the speaker thinks that something is funny, and by looking at the information received by the second channel, we get to know that the real meaning of the message is that the speaker is lying or being sarcastic.

Emotion from text

There are certain situations in which the communication between two people, or between one person and one machine, has not the visual component given by the face-to-face communication.

Detecting emotional state of a person by analyzing a text written by him/her appear challenging but also essential many times due to the fact that most of the times textual expressions are not only direct using emotion words but also result from the interpretation of the meaning of concepts and interaction of concepts which are described in the text. Although emotion detection from text must face more obstacles like spelling errors, languages, slang, and so on. Since emotion detection from texts analyses the words contained on a message, the process to analyses a text take some more steps than the analysis of a face or a voice. There is still a model that needs to be trained, but now text must be processed in order to use it to train a model. This processing involves tasks of tokenization, parsing and part-of-speech tagging, lemmatization, stemming, among others.

Emotion from facial expression

Detecting human emotions from facial expressions and classifying them into one of the six basic emotions. As in the case of speech, facial expressions reflect the emotions that a person can be feeling. Eyebrows, lips, nose, mouth, and muscles of the face: they all reveal the emotions were feeling. Even when a person tries to fake some emotion, still their own face is telling the truth. The technologies used in this field of emotion detection work in an analogous way to the ones used with speech detecting a face, identifying the crucial points in the face which reveal the emotion expressed and processing their positions to decide what emotion is being detected.

Emotion from body gesture and movement

Human motion analysis is currently one of the most active research fields in computer vision. The aim is to detect, track and identify people, and more generally, to interpret human behaviors, from image sequences involving humans. Even though people do not use it to communicate information in an active way, their body is constantly broadcasting affective information. Tapping with the foot, crossing the arms, tilting the head, changing our position a lot of times while being sat, etc. Body language reveals what a person is feeling the same way our voice does [27]. However, this field is pretty new and there is not a clear understanding about how to create systems which read emotions in a body gesture. Most researchers have focused on facial expressions (over 95 per cent of the studies carried out on emotions detection have used faces as stimuli), almost ignoring the rest of channels through which people reveal affective information [28].

Emotion from physiological state

Emotion, mood, and emotion recognition has been studied in laboratory settings for decades. In particular, physiological signals are widely used to detect and classify affective states in lab conditions. However, physiological reactions to emotional stimuli have been found to differ in laboratory and natural settings. There are states of feeling that result in physical and psychological changes that influence our behavior. The physiology of emotion is closely linked to arousal of the nervous system with various states and strengths of arousal relating, apparently, to particular emotions. Physiologically speaking, emotions originate on the limbic system. Within this system, the amygdala generates emotional impulses which create the physiological reactions associated with emotions [29]. Electric activity on face muscles, electro dermal activity (also called galvanic skin response), pupil dilatation, breath and heart rate, blood pressure, brain electric activity, etc. Emotions leave a trace on the body, and this can be measured with the right tools. Nevertheless, information coming directly from the body is harder to classify, at least with the category system used in other emotion detection technologies. When working with physiological signals, the best option is to adopt a classification system based on a dimensional approach. An emotion is not just "happiness" or "sadness" anymore, but a state determined by various dimensions, like valence and arousal. It is because of this that the use of physiological signals is usually reserved for research and studies.

2.6 Emotion detection techniques

We have identified four major categories of techniques that have been proposed for building text based emotion detection models. They are corpus-based methods, machine learning based techniques, knowledge-based techniques and hybrid techniques [30].

Corpus based technique

This technique includes the use of an emotion lexicon with weighted scores from training documents which are then used to build an emotion prediction model. Corpus-based classification uses unigrams (bag-of-words). Its key features are that it employs an emotion lexicon with weighted scores from training documents and uses unigrams.

Machine learning based technique

This technique makes use of an annotated corpus to train an emotion classifier. Its key features are that it employs an annotated corpus to train the emotion classifier. It also uses a supervised or unsupervised method to classify emotions and relies on a classifier for emotion detection.

Knowledge based technique

This technique applies linguistic rules through exploiting the knowledge of sentence structures in conjunction with sentiment resources for emotion classification. Its key features are that it applies linguistic rules and exploits sentence structures in conjunction with sentiment resources like.

Hybrid techniques

The hybrid method uses both the Corpus based and machine learning approaches to be utilized which allows the advantages of both methods.

2.7 Emotion detection methods

Emotion detection methods can be broadly classified into keyword based, learning-based and hybrid methods. We further distinguish them based on whether they employ any affect lexicons [31].

Keyword-based methods

Keyword-based methods are the most intuitive ways to detect textual emotions. Keyword-based emotion detection serves as the starting point of textual emotion recognition. Once the set of emotion labels (and related words) is constructed, it can be used exhaustively to examine if a sentence contains any emotions.

However, while detecting emotions based on related keywords is very straight forward and easy to use, the key to increase accuracy falls to two of the preprocessing methods, which are sentence parsing to extract keywords, and the construction of emotional keyword dictionary. Parsers utilized in emotion detection are almost ready-made software packages, whereas their corresponding theories may differ from dependency grammar to theta role assignments. On the other hand, constructing emotional keyword dictionary would be naval to other fields. As this dictionary collects not only the keywords, but also the relations among them, this dictionary usually exists in the form of thesaurus, or even ontology, to contain relations more than similar and opposite ones. As was observed in [32], keyword-based emotion detection methods have three limitations described below.

- **Ambiguity in keyword definitions**

Though using emotion keywords is a straightforward way to detect associated emotions, the meanings of keywords could be multiple and vague. Except those words standing for emotion labels themselves, most words could change their meanings according to different usages and

contexts, and it is just not feasible to include all possible combinations. Moreover, even the minimum set of emotion labels (without all their synonyms) could have different emotions in some extreme cases such as ironic or cynical sentences.

- **Incapability of recognizing sentences without keywords**

Keyword-based approach is totally based on the set of emotion keywords. Therefore, sentences without any keywords would imply they do not contain any emotions at all, which is obviously wrong. For example, “I passed my qualify exam today” and “Hooray! I passed my qualify exam today” should imply the same emotion (joy), but the former without “hooray” could remain undetected if “hooray” is the only keyword to detect this emotion.

- **Lack of linguistic information**

Syntax structures and semantics also have influences on expressed emotions. For example, “I laughed at him” and “He laughed at me” would suggest different emotions from the first person’s perspective. As a result, ignoring linguistic information also poses a problem to keyword-based methods.

In summary, keyword-based methods should also detect not only the existence of keywords, but also their linguistic information to detect emotions more accurately.

Learning-based methods

Researchers using learning-based methods attempt to formulate the problem differently. The original problem that determining emotions from input texts has become how to classify the input texts into different emotions. Unlike keyword-based detection methods, learning-based methods try to detect emotions based on a previously trained classifier, which apply various theories of machine learning such as support vector machines and conditional random fields to determine which emotion category should the input text belongs.

However, comparing the satisfactory results in multimodal emotion detection, the results of detection from texts drop considerably. The reasons are addressed below:

- **Difficulties in determining emotion indicators**

The first problem is, though learning-based methods can automatically determine the probabilities between features and emotions, learning-based methods still need keywords, but just in the form of features. The most intuitive features may be emoticons, which can be seen as author’s emotion annotations in the texts. The cascading problems would be the same as those in keyword-based methods.

- **Over-simplified emotion categories**

Nevertheless, lacking of efficient features other than emotion keywords, most learning-based methods can only classify sentences into two categories, which are positive and negative. Although the number of emotion labels depends on the emotion model applied, we would expect to refine more categories in practical systems.

Hybrid methods

Since keyword-based methods with thesaurus and naive learning-based methods could not acquire satisfactory results, some systems use a hybrid approach by combining both or adding different components, which help to improve accuracy and refine the categories.

2.8 Emotion detection approaches

The existing approaches in emotion detection to classify certain emotion types can be generally classified into three main categories knowledge-based, statistical, and hybrid approaches [33].

Knowledge-based approaches

Knowledge-based techniques (sometimes referred to as lexicon-based techniques), utilize domain knowledge and the semantic and syntactic characteristics of language in order to detect certain emotion types. In this approach, it is common to use knowledge-based resources during the emotion classification process. One of the advantages of this approach is the accessibility and economy brought about by the large availability of such knowledge-based resources. A limitation of this technique on the other hand, is its inability to handle concept nuances and complex linguistic rules.

Knowledge-based approaches can be mainly classified into two categories dictionary-based and corpus-based approaches. Dictionary-based approaches find opinion or emotion seed words in a dictionary and search for their synonyms and antonyms to expand the initial list of opinions or emotions. Corpus-based approaches on the other hand, start with a seed list of opinion or emotion words, and expand the database by finding other words with context-specific characteristics in a large corpus. While corpus-based approaches take into account context, their performance still vary in different domains since a word in one domain can have a different orientation in another domain

Statistical approaches

Statistical approaches, such as Bayesian inference and support vector machines, have been popular for affect classify caption of texts. By feeding a machine learning algorithm a large

training corpus of affectively annotated texts, it's possible for the system to not only learn the affective valence of affect keywords, but also to take into account the valence of other arbitrary keywords (like lexical), punctuation, and word co-occurrence frequencies. However, traditional statistical methods are generally semantically weak, meaning that, with the exception of obvious affect keywords, other lexical or co-occurrence elements in a statistical model have little predictive value individually. As a result, statistical emotion classifiers only work with acceptable accuracy when given a sufficiently large text input. So, while these methods may be able to affectively classify user's text on the page- or paragraph level, they do not work well on smaller text units such as sentences or clauses.

Hybrid approaches

Hybrid approaches in emotion recognition are essentially a combination of knowledge-based and statistical approaches, which exploit complementary characteristics from both techniques. The role of such knowledge-based resources in the implementation of hybrid approaches is highly important in the emotion classification process. Since hybrid techniques gain from the benefits offered by both knowledge-based and statistical approaches, they tend to have better classification performance as opposed to employing knowledge-based or statistical methods independently. A downside of using hybrid techniques however, is the computational complexity during the classification process.

2.9 How to detect emotion

There are lots of methods to emotion detection. In most researches, methods widely used for social media emotion detection include machine learning based, deep learning, rule-based and lexicon based. In this section, we will discuss each and every technique one by one.

2.9.1 Machine learning based

Machine learning methods divided into supervised learning and unsupervised learning.

Supervised learning

This approach is used when there is labeled data available for training the model. Two steps are used in supervised learning: first is to train the model and another is prediction. During training, data set with its labels is fed to the classification algorithm which gives a model as an output. After that test data is fed into the model to predict the category. There are various supervised classification algorithm are:

Naïve Bayes.

It is a probabilistic classification algorithm. It considers each word independent as it does not consider the location of a term in the sentence. Naïve Bayes based on Bayes theorem to calculate the probability of each term which corresponding to a label.

Advantages

- It is not only a simple approach but also a fast and accurate method for prediction.
- Naive Bayes has very low computation cost.
- It can efficiently work on a large dataset.
- It performs well in case of discrete response variable compared to the continuous variable.
- It can be used with multiple class prediction problems.
- It also performs well in the case of text analytics problems.
- When the assumption of independence holds, a Naive Bayes classifier performs better compared to other models like logistic regression.

Dis advantages

- The assumption of independent features. In practice, it is almost impossible that model will get a set of predictors which are entirely independent.
- If there is no training tuple of a particular class, this causes zero posterior probability. In this case, the model is unable to make predictions. This problem is known as Zero Probability/Frequency Problem.

Support Vector Machine

SVM is initialized first time to solve the problems of binary classification. Its focuses on determining best hyper planes which act as a separator to describe the decision boundaries among the data points which are from different classes. A hyper plane should be selected which can maintain the maximum distance between two support vectors of different classes as shown in the figure. The SVM has the capability to manage the linear and non-linear classification tasks. SVM for classification with various weighting schemes like TF-IDF, term occurrence, Binary Occurrence.

Advantages

- SVM works relatively well when there is a clear margin of separation between classes.
- SVM is more effective in high dimensional spaces.

- SVM is effective in cases where the number of dimensions is greater than the number of samples.
- SVM is relatively memory efficient

Disadvantages

- SVM algorithm is not suitable for large data sets.
- SVM does not perform very well when the data set has more noise i.e. target classes are overlapping.
- In cases where the number of features for each data point exceeds the number of training data samples, the SVM will underperform.
- As the support vector classifier works by putting data points, above and below the classifying hyper plane there is no probabilistic explanation for the classification.

Decision tree

It is a tree like structure where the non- terminal nodes represent a feature and terminal node represents the label. The path is taken on the basis of a condition. This is a recursive process and ultimately reaches a terminal node which gives a label to an input. The main challenge in the decision tree is to find which attribute is to be chosen as a root node. This can be solved by using some statistical approach such as information gain and Gini index. A Decision tree is a good method for sentiment analysis because it also provides a good result on a large amount of data. Commonly used decision tree algorithms are CART, CHAID, and C5.0.

Advantages

- Compared to other algorithms decision trees requires less effort for data preparation during pre-processing.
- A decision tree does not require normalization of data.
- A decision tree does not require scaling of data as well.
- Missing values in the data also do NOT affect the process of building a decision tree to any considerable extent.
- A Decision tree model is very intuitive and easy to explain to technical teams as well as stakeholders.

Disadvantage

- A small change in the data can cause a large change in the structure of the decision tree causing instability.

- For a Decision tree sometimes calculation can go far more complex compared to other algorithms.
- Decision tree often involves higher time to train the model.
- Decision tree training is relatively expensive as the complexity and time has taken are more.
- The Decision Tree algorithm is inadequate for applying regression and predicting continuous values.

Unsupervised learning

This method is used when the reliability of labeled data is difficult. It is easy to collect the unlabeled data than labeled data. The sentence is categorized on the basis of key-word lists of each category. In order to analyses the domain dependent data, it is easier using the unsupervised approach. Unsupervised learning algorithms allow you to perform more complex processing tasks compared to supervised learning. Unsupervised learning problems further grouped into clustering and association problems.

- **Clustering**

Clustering is an important concept when it comes to unsupervised learning. It mainly deals with finding a structure or pattern in a collection of uncategorized data. Clustering algorithms will process your data and find natural clusters (groups) if they exist in the data. You can also modify how many clusters your algorithms should identify. It allows you to adjust the granularity of these groups.

- **Association**

Association rules allow you to establish associations amongst data objects inside large databases. This unsupervised technique is about discovering exciting relationships between variables in large databases. For example, people that buy a new home most likely to buy new furniture.

2.9.2 Rule-based

Rule-based approaches usually defined a number of rules in a certain type of scripting language, which indicates subjectivity, polarity or an opinion. There are various inputs available in the rules, such as conventional NLP techniques such as stemming, tokenization, voice tagging, and parsing. A basic example of these techniques is two lists of polarized words, e.g. negative words such as bad, worse, ugly, etc., and positive words such as good, best and beautiful.

The number of positive words in the text is counted. In the text, count the number of negative words. When the number of positive word appearances is higher than that of negative word appearances, a positive feeling gives the contrary a negative feeling. More advanced processing may be possible, but such systems become very complex very quickly. Additionally, the addition of new rules could lead to undesirable results as such systems require substantial investment in the manual tuning and observance of the rules.

2.9.3 Lexicon based

The lexicon-based approach involves calculating orientation for a document from the semantic orientation of words or phrases in the document. The text classification approach involves building classifiers from labeled instances of texts or sentences, essentially a supervised classification task. The latter approach could also be described as a statistical or machine-learning approach. We follow the first method, in which we use dictionaries of words annotated with the word's semantic orientation, or polarity.

Dictionaries for lexicon-based approaches can be created manually, as we describe in this article, or automatically, using seed words to expand the list of words. Much of the lexicon-based research has focused on using adjectives as indicators of the semantic orientation of text. First, a list of adjectives and corresponding values is compiled into a dictionary. Then, for any given text, all adjectives are extracted and annotated with their value, using the dictionary scores. The scores are in turn aggregated into a single score for the text.

2.9.4 Deep learning

Neural Networks is a method that tries to optimize some weights, the body of the neuron, that are multiplied by the vector of features, the dendrites. The result of this multiplication is the prediction made by this neuron, the axon terminal. It can be used as a result or as a feature to the next set of neurons, called MLP. The objective is train the internal weights using the Gradient Descent and Back-propagation method where a cost function is computed and the result is propagated back to the neurons weights that gets updated to minimize the objective function on each round.

The different types of neural networks in deep learning, such as CNN, RNN, ANN, etc. are changing the way we interact with the world. These different types of neural networks are at the core of the deep learning revolution, powering applications like unmanned aerial vehicles, self-

driving cars, speech recognition, emotion detection etc. Feature engineering is a key step in the model building process. It is a two-step process feature extraction and feature selection. In feature extraction, we extract all the required features for our problem statement and in feature selection, we select the important features that improve the performance of our machine learning or deep learning model. This section we focuses on three important types of neural networks that form the basis for most pre-trained models in deep learning.

Convolution Neural Networks

When it comes to image classification, the most used neural networks are Convolution Neural Networks in recent time in most research used CNN for text classification. CNN contain multiple convolution layers which are responsible for the extraction of important features from the image/text. The earlier layers are responsible for low-level details and the later layers are responsible for more high-level features. The Convolution operation uses a custom matrix, also called as filters, to convolute over the input text. These filters are initialized randomly and then are updated via back propagation. After the convolution layer, there is a pooling layer which is responsible for the aggregation of the maps produced from the convolutional layer. It can be Max Pooling, Min Pooling, etc. For regularization, CNNs also include an option for adding dropout layers which drop or make certain neurons inactive to reduce over fitting and quicker convergence.

Advantages

- Used for deep learning with few parameters.
- Less parameter to learn as compared to fully connected layer.

Disadvantages

- Comparatively complex to design and maintain.
- Comparatively slow [depends on the number of hidden layers].

Recurrent Neural Networks

Recurrent Neural Networks come into picture when there's a need for predictions using sequential data. Sequential data can be a sequence of images, words, etc. The RNN have a similar structure to that of a Feed-Forward Network, except that the layers also receive a time-delayed input of the previous instance prediction. This instance prediction is stored in the RNN cell which is a second input for every prediction.

Advantages

- Model sequential data where each sample can be assumed to be dependent on historical ones is one of the advantages.
- Used with convolution layers to extend the pixel effectiveness.

Disadvantages

- Gradient vanishing and exploding problems.
- Training recurrent neural nets could be a difficult task.
- Difficult to process long sequential data using ReLU as an activation function.

Artificial Neural Network

A single perceptron (or neuron) can be imagined as a Logistic Regression. Artificial Neural Network, or ANN, is a group of multiple perceptron's/ neurons at each layer. ANN is also known as a Feed-Forward Neural network because inputs are processed only in the forward direction. ANN consists of 3 layers input, hidden and output. The input layer accepts the inputs, the hidden layer processes the inputs, and the output layer produces the result. Essentially, each layer tries to learn certain weights.

Advantages

- Artificial Neural Network is capable of learning any nonlinear function.
- ANNs have the capacity to learn weights that map any input to the output.

Disadvantages

- Vanishing and Exploding Gradient.
- ANN cannot capture sequential information in the input data.

Word Embedding

Word Embedding is a representation of text where words that have the same meaning have a similar representation. In other words it represents words in a coordinate system where related words, based on a corpus of relationships, are placed closer together. In the deep learning frameworks this part is usually handled by an embedding layer which stores a lookup table to map the words represented by numeric indexes to their dense vector representations. Computers are unable to understand the concepts of words. It requires data to be converted into a numeric format to perform any machine learning task. In order to perform such tasks, various word embedding techniques are being used i.e. Bag of Words, TF-IDF, Word2vec to encode the text data. In order to process natural language, a mechanism for representing text is required. Word

embedding's are commonly used in many Natural Language Processing tasks because they are found to be useful representations of words and often lead to better performance in the various tasks performed. Given its widespread use, in this paper, we will try to introduce the concept of word embedding's with respect to Amharic language.

The word embedding of our dataset can be learned while training a neural network on the classification problem. Before it can be presented to the network, the text data is first encoded so that each word is represented by a unique integer. The vectors created by Word Embedding preserve these similarities, so words that regularly occur nearby in the text will also be in close proximity in vector space. In general, word embedding is a means of building a low-dimensional vector representation from the corpus of text, which preserves the contextual similarity of words; the semantic relationships between words are reflected in the distance and direction of the vectors. The standard mechanism for text representation is word vectors where words or phrases from a given language vocabulary are mapped to vectors of real numbers.

The vectors are very good at answering analogy questions of the form a is to b as c is to? For example, wonde (**ወንድም**) literal meaning "male" is to ehite (**እህት**) literal meaning "sister" as abate (**አባት**) is to? Enat (**እናት**) literal meaning "mother" using a simple vector offset method based on cosine distance. Using different word embedding's we can represent the sentence differently in numbers. Here we will discuss TF-IDF and Word2Vec.

- **TF-IDF**

TF-IDF is a statistical method for efficiently learning a standalone word embedding from a text corpus. It was developed by Tomas Mikolov, et al. at Google in 2013 as a response to make the neural-network-based training of the embedding more efficient and since then has become the de facto standard for developing pre-trained word embedding. Additionally, the work involved analysis of the learned vectors and the exploration of vector math on the representations of words.

Using TF-IDF embedding's, word will be represented as a single scalar number based on TF-IDF scores. TF-IDF is the combination of TF and IDF. TF gives the count of word t in document d. mathematically we can write tf (t, d). IDF gives information about how the word is common or rare across all documents. It is the logarithmically scaled inverse fraction of the documents that contain the word. Mathematically, idf (t,D) = log (N/dfi) , where N or |D| = Total Number of Document, and dfi = Number of document where the term t appears.

- **Word2Vec**

Word2Vec is one of the most popular techniques to learn word embedding's using shallow neural network. Consider the following similar sentences mlekam (**መልካም**) literal meaning "well" wulo (**ውሉ**) literal meaning "day" and melekam (**መልካም**) literal meaning "well" kene (**ቀን**) literal meaning "day". They hardly have different meaning. If we construct an exhaustive vocabulary (let's call it V), it would have V = {**መልካም**, **ውሉ**, **ቀን**}. Now, let us create a one-hot encoded vector for each of these words in V. Length of our one-hot encoded vector would be equal to the size of V (=5). We would have a vector of zeros except for the element at the index representing the corresponding word in the vocabulary. That particular element would be one. The encodings below would explain this better.

መልካም = [1,0,0,0,0]`; **ውሉ** =[0,1,0,0,0]` ; **ቀን** =[0,0,1,0,0]` (`represents transpose). If we try to visualize these encodings, we can think of a 5 dimensional space, where each word occupies one of the dimensions and has nothing to do with the rest (no projection along the other dimensions). This means '**ውሉ**' and '**ቀን**' are as different, which is not true. Our objective is to have words with similar context occupy close spatial positions. Mathematically, the cosine of the angle between such vectors should be close to 1, i.e. angle close to 0.

Word2vec is a group of related models that are used to produce word embedding's. These models are shallow, two-layer neural networks that are trained to reconstruct linguistic contexts of words. Word2vec takes as its input a large corpus of text and produces a vector space, typically of several hundred dimensions, with each unique word in the corpus being assigned a corresponding vector in the space. Word vectors are positioned in the vector space such that words that share common contexts in the corpus are located close to one another in the space.

Word2vec treats each word in corpus like an atomic entity and generates a vector for each word. For example, the word zegeba (**ዘገባ**) literal meaning "report", lezegeba (**ለዘገባ**) literal meaning "for report", bzegeba (**በዘገባ**) literal meaning "by report", selezegeba (**ስለዘገባ**) literal meaning "about the report", etc are treated as atomic unless we apply morphology analysis before providing dataset to model.

Two different learning models were introduced that can be used as part of the word2vec approach to learn the word embedding they are continuous Bag-of-Words, or CBOW model and continuous Skip-Gram Model [34]. The CBOW model learns the embedding by predicting the

current word based on its context. The continuous skip-gram model learns by predicting the surrounding words given a current word. The continuous skip-gram model learns by predicting the surrounding words given a current word.

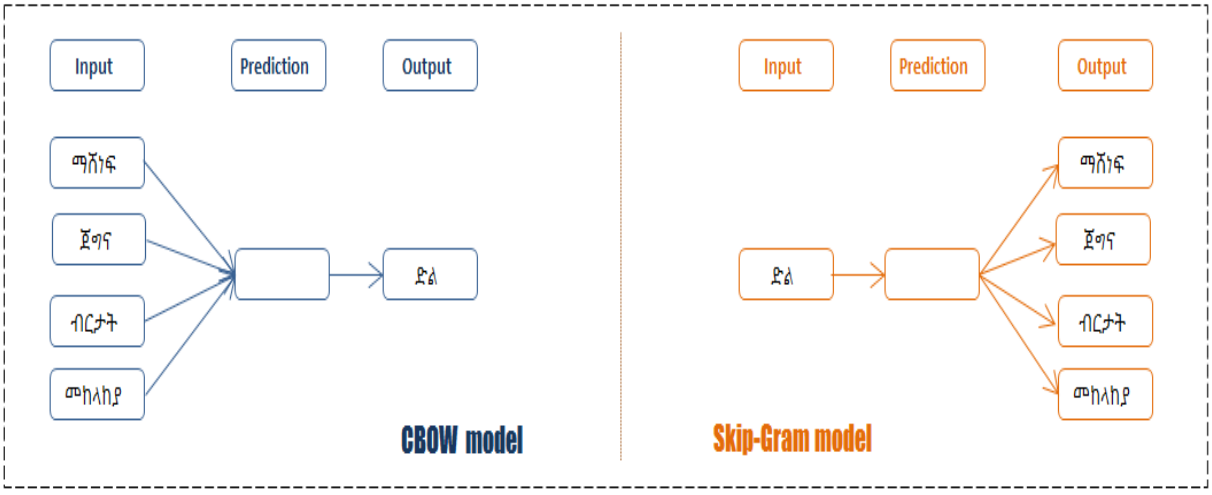

Figure 2-1 Word2Vec Models

Both models are focused on learning about words given their local usage context, where the context is defined by a window of neighboring words. This window is a configurable parameter of the model. The key benefit of the approach is that high-quality word embedding's can be learned efficiently (low space and time complexity), allowing larger embedding's to be learned (more dimensions) from much larger corpora of text (billions of words).

The CBOW model predicts the target word according to its context. For instance, given the sentence "**የህዝብ ግንኙነት ለህዝብ መረጃ በመስጠት የህዝብን አዝማሚያዎችና ተግባራት ለማስተካከል የሚከሄድ የማሳመን ሥራ ነዉ**" literal meaning "Public relations is a means of persuading people to change their attitudes and actions by providing information to the public" and the target word "**ተግባራት**".The Skip-gram model predicts the target using a random close-by word, like "**አዝማሚያዎችና**" or "**የማሳመን**". Whereas the CBOW model takes all the words in a surrounding window, like {**የህዝብን**: **አዝማሚያዎችና**, **ለማስተካከል**: **የሚከሄድ**}, and uses the sum of their vectors to predict the target word "**ተግባራት**" [8].

In short, Word Embedding turns text into numbers. This transformation is necessary because many machine learning algorithms (including deep nets) require their input to be vectors of continuous values they just won't work on strings of plain text. In relation to this, TF-IDF is a

word-document mapping (with some normalization). TF-IDF is obtained from straight forward linear algebra. Word2Vec is obtained from the hidden layer of a two-layered neural network. TF-IDF can be used either for assigning vectors to words or to documents. Word2Vec can be directly used to assign a vector to a word but to get the vector representation of a document further processing is needed. Unlike TF-IDF Word2Vec takes into account the placement of words in a document (to some extent).

2.10 Applications

Human beings have various emotions, which can now be recognized by machines and computers. One developer of such algorithms, emotion recognition as being able to help a range of industries, from retail to healthcare, achieve their business objectives. Software that purportedly reads emotions in faces is being deployed or tested for a variety of purposes, including surveillance, hiring, clinical diagnosis, and market research. But a new scientific report finds that facial movements are an inexact gauge of a person's feelings, behaviors or intentions. It is not possible to confidently infer happiness from a smile, anger from a scowl or sadness from a frown, as much of current technology tries to do when applying what are mistakenly believed to be the scientific facts, a group of leading experts in psychological science, neuroscience and computer science write in their comprehensive research review.

Emotion detection is used in society for a variety of reasons. This provides artificial intelligence software that makes it more efficient to do tasks previously done manually by people, mainly to gather facial expression, form textual and vocal expression information related to specific contexts where viewers have consented to share this information. For example, instead of filling out a lengthy survey about how you feel at each point watching an educational video or advertisement, you can consent to have a camera watch your face and listen to what you say, and note during which parts of the experience you show expressions such as boredom, interest, confusion, or smiling the same as in textual ways.

Other uses include helping children with autism, helping people who are blind to read facial expressions, helping robots interact more intelligently with people, and monitoring signs of attention while driving in an effort to enhance driver safety. Many products also exist to aggregate information from emotions communicated online comments and posts, including via "like" button presses and via counts of positive and negative phrases in text and affect recognition is increasingly used in some kinds of games and virtual reality.

2.11 Evaluation methods

Once we built our model, the most important question that arises is how good is our model? So, evaluating our model is the most important task in the data science project which delineates how good our predictions are. As we discussed realer accuracy is not the only metric for evaluating the effectiveness of our model. Two other useful metrics are precision and recall. These two metrics can provide much greater insight into the performance characteristics of our model.

TP - These are the correctly predicted positive value which means that the value of actual class is yes and the value of predicted class is also yes.

TN - These are the correctly predicted negative values which means that the value of actual class is no and value of predicted class is also no.

False positives and false negatives, these values occur when your actual class contradicts with the predicted class.

FP – When actual class is no and predicted class is yes. E.g. if actual class says this passenger did not survive but predicted class tells you that this passenger will survive.

FN – When actual class is yes but predicted class in no. E.g. if actual class value indicates that this passenger survived and predicted class tells you that passenger will die.

Once we understand these four parameters then we can calculate Accuracy, Precision, Recall and F1 score.

Accuracy - Accuracy is the most intuitive performance measure and it is simply a ratio of correctly predicted observation to the total observations. One may think that, if we have high accuracy then our model is best. Yes, accuracy is a great measure but only when you have symmetric datasets where values of false positive and false negatives are almost same. Therefore, you have to look at other parameters to evaluate the performance of your model.

Accuracy = TP+TN/TP+FP+FN+TN

Precision - Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. The question that this metric answer is of all passengers that labeled as survived, how many actually survived? High precision relates to the low false positive rate.

Precision = TP/TP+FP

Recall - Recall is the ratio of correctly predicted positive observations to the all observations in actual class yes. The question recall answer is of all the passengers that truly survived, how many did we label?

Recall = TP/TP+FN

F1 score - F1 Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. Intuitively it is not as easy to understand as accuracy, but F1 is usually more useful than accuracy, especially if you have an uneven class distribution. Accuracy works best if false positives and false negatives have similar cost. If the cost of false positives and false negatives are very different, it's better to look at both Precision and Recall.

F1 Score = 2*(Recall * Precision) / (Recall + Precision)

Research on social media text has used emotion analysis, the automated extraction of emotions from texts posted on social media. But little attention was given to emotional posts that appear on Ethiopian social media. Different techniques, approaches, evaluation methods and etc. have never been discussed in relation to Ethiopian social media. In our work, we will follow the sentiment analysis approach to perform emotion analysis and evaluation. Amharic posts on Facebook social media will be our main focus while designing our model. In summary, we have seen how precision, F1 score, accuracy and recall naturally lead to probabilistic interpretation, and how one can derive probabilistic distributions of them.

Chapter 3: Related work

3.1. Introduction

In recent years a lot of work has been done in the field of Sentiment analysis by number of researchers. In fact work in the field started since the beginning of the century. In its early stage it was intended for binary classification, which assigns opinions or reviews to bipolar classes such as positive or negative. Many efforts [9-17] done previously in creating sentiment analysis that can be related to the research challenges that we will be addressing in our study, which is deep learning based emotion detection for Amharic text. There are a number of studies on sentiment analysis and some are closely related to our study. We focused on two aspects for selecting studies of our discussion, namely the traditional emotion detection approaches and based on deep learning approaches.

There are number of research works on emotion detection in foreigner language and development and refining the automated techniques of emotion classification and analysis. Many researchers have worked on sentiment analysis techniques via different approaches however, in depth analysis and review on emotion analysis or detection with machine learning algorithms for under resourced language like Amharic is still need research.

3.2. Traditional approaches

By learning methods, traditional emotion classification methods can be divided into three machine learning categories, namely, supervised, unsupervised, and semi-supervised learning methods. Supervised machine learning is immensely helpful in solving real-world computational problems. The algorithm predicts outcomes for unforeseen data by learning from labeled training data. Therefore, it takes highly-skilled data scientists to build and deploy such models. Over time, data scientists also use their technical expertise to rebuild the models to maintain the integrity of the insights given. Unsupervised learning methods mainly classify emotions based on prior knowledge. These methods are less effective than their supervised counterparts but have attracted wide attention from researchers because they do not entail a large labeled corpus. The semi-supervised learning method is a suitable option when there are only few labeled datasets. It consumes less time and manpower than supervised learning methods and achieves better classification effect than unsupervised learning methods. Therefore, semi-supervised learning methods demonstrate advantages over the other two types of methods.

Hiwot Wonago Kululo [2] proposes Information filtering of social media Amharic texts based on Sentiment Analysis. In this work the concept of sentiment analysis on Amharic text on social media and presents a comparative study on machine learning algorithms. They created social media content filtering system has been tested on Facebook posts of each class, and it has been observed that SVM with word2vec has performed best in comparison to other classifiers, achieving average precision of (72%), but did worse on recall (63.4%). The contents classified into “offensive” and “non-offensive” categories.

Zewdie Mossie and Jenq-Haur Wang [13] developed an apache spark based model to classify Amharic Facebook posts and comments into “hate” and “not hate” categories. Authors employed supervised Random forest and Naïve Bayes machine learning and Word2Vec and TF-IDF for feature selection. The results showed that accuracy of 79.83 % and 65.34% for Naïve Bayes with word2vec feature vector and Random Forest with TF-IDF feature modeling approach respectively.

Wondwossen Philemon and Wondwossen Mulugeta [15] propose a multi-scale approach to sentiment analysis provides a more refined breakdown than the traditional positive-negative binary scheme. The approach is preferred in cases where comparison and ranking of opinions is vital. They employed Naïve Bayes machine learning algorithm and used unigram, bigram and hybrid variants as features. There Naïve Bayes implementation to multi-scale sentiment analysis was successful and they achieved a promising performance accuracy of 43.6%, 44.3% and 39.5% for unigram, bigram and hybrid language models, respectively despite the few training data used. To achieve the objective, the author has prepared a sample corpus that contains 608 posts. The corpus was collected from social media sources such as Facebook, Twitter, Dire Tube and Ethiopian reporter websites. The author employed preprocessing activities in the corpus dataset before the actual sentiment classification. After the preprocessing activities were done, the corpus was manually annotated by giving polarity values and sentiment intensity scale values.

3.3. Based on deep learning

Deep learning approaches have recently emerged from machine learning and soft computing techniques. Since then, several deep learning algorithms have been recently introduced to scientific communities and are applied in various application domains. Today the usage of DL has become essential due to their intelligence, efficient learning, accuracy and robustness in model building.

Dongliang Xua,b, Zhihong Tiana, Rufeng Laic, Xiangtao Kongb, Zhiyuan (Thomas) T and Wei Shie [12] proposes a microblog emotion classification model, namely, CNN_Text_Word2vec, on the basis of CNN to solve the above-mentioned problems. CNN_Text_Word2vec introduces a word2vec neural network model to train distributed word embedding's on every single word. The trained word vectors are used as input features for the model to learn microblog text features through parallel convolution layers with multiple convolution kernels of different sizes. Experiment results show that the overall accuracy rate of CNN_Text_Word2vec is 7.0% higher than that achieved by current mainstream methods, such as SVM, LSTM and RNN. Moreover, this study explores the impact of different semantic units on the accuracy of CNN_Text_Word2vec, specifically in processing of Chinese texts. The experimental results show that with positive, negative classification they get accuracy of 97.6% comparing to using feature vectors obtained from training words feature vector obtained from training Chinese characters yields a better performance.

Eman Hamdi(&) , Sherine Rady, and Mostafa Aref [35] on this research they applying the state of the art in sentiment analysis using convolutional neural network model for Emotion Detection from Tweets. They propose a deep learning model for emotion detection from short informal sentences. The model consists of three Convolutional Neural Networks. Each CNN contains a convolutional layer and a max-pooling layer, followed by a fully-connected layer for classifying the sentences into positive or negative. This model has been tested on the Stanford Twitter Sentiment dataset for classifying sentiment into two classes positive and negative. The presented model achieved to record 80.6% accuracy as a prove that even with randomly initialized word vectors, it can work very well in text classification tasks when trained with CNNs.

3.4. Summary

In this chapter, gaps of different researches on Amharic language sentiment analysis and emotion detection are investigated and analyzed. From the analysis of the related works, we came up with the following major problems of present day emotion detection and sentiment analysis as a social media monitoring tool for Amharic language.

Most foreign languages have been the target language in most sentiment analysis research. Recent efforts extend the focus to other languages such as Amharic. Basic machine learning techniques as Naive Bayes and SVM have been used to achieve baseline results [2][13][15]. However, these systems require lots of feature engineering work prior to applying any machine learning method. There is a lack of widely available benchmark for comparing different machine learning algorithms for Amharic Sentiment Analysis of social media text.

Researches on sentiment analysis of social media most time investigated and analyzed on the foreign languages. Only a few researches, try to solve the problems of morphologically rich languages such as Amharic Language [2]. Despite its large number of speakers, the Amharic language hasn‘t got much attention from researchers. Furthermore, social media is a source of raw data which includes the following metrics shares, likes, conversations, comments, mentions, impressions, and most important status updates. It is a place where unstructured information exists. Simply put, for the purposes of social media, it is the process of determining the author's opinion conveyed in a post. With social media monitoring, sentiment analysis is much harder because there isn‘t a defined contextualization process. People talk about everything and anything under the sun and their feelings and opinions toward certain topics are almost impossible to contextualize for a computer.

Most of the users share their emotion in different social media using Amharic language this has changed the manner in which people communicate and influence behavior and emotion of other people in the social network. Emotions of users that are expressed on the web have great influence on the readers, product vendors and politicians. The unstructured form of data from the Amharic social media is needed to be analyzed and well-structured for this purpose emotion detection has recognized significant attention. The absences of Amharic text emotion detection on social media motivate us to build an Amharic text emotion detection model. Our work basically differ from above related work special from [2][13][15] we use deep learning algorithm and they use machine learning algorithms but also we use feature extraction to get comparatively

high performance difference on emotion detection for Amharic texts. On those related work they focused on sentiment analysis with classification of opining with positive, negative and neutral our model classifies emotion with anger, disgust, happiness, and sadness. As we know sentiment analysis can be applied at different levels of scope document-level, sentence, and aspect (feature) level in the above related works except [12][35] they applied document-level sentiment analysis to obtain the sentiment of a complete document. In our study we apply sentence level sentiment analysis to detect emotional state of every sentence. As we can understand from the above related work , for the emotion detection analysis there is no research done on Amharic language, there are lots of works done in different non-Amharic language [12][35] using different techniques and the different accuracy gets, also there is no publication on text based Amharic language emotion detection using deep machine learning techniques. We have proposed model using CNN with word2vec for emotion detection Amharic texts.

Chapter 4: The Proposed Emotion Detection Model for Amharic Text

4.1 Introduction

In recent years, there has been a clear shift in state of the art approaches from statistical machine learning to deep learning based on text categorization models. These have been mainly used to develop an end-to-end deep neural network to extract contextual features from raw text. In this chapter, the design and implementation of the proposed emotion detection model for Amharic texts is described in detail. We illustrate the main components of our system. First, the Amharic texts pass through a preprocessing, stemming and cleaning phase, to remove unwanted symbols and tokens. Then, the texts are prepared to be ready for the training phase. Then we will describe how data preprocessing is performed, how deep learning algorithms classified emotions components along with sub components which are the building blocks of the system.

4.2 Amharic text emotion detection model

This study designed pipeline model of deep learning based emotion detection for Amharic text using CNN called CNN. CNN was proposed based on the basic model of a CNN. After data preprocessing, the comment text was firstly input into the CNN network model to judge the emotion category. Finally, the emotions categories of the different comment texts under the emotional category were obtained. The design of the proposed model has three major components data pre-processing (tokenization, normalization, stop word removal and light steaming), feature extraction, and emotion classifier. We will begin by showing the architectural design of the emotion detection then we will be discussing the main components of the model along with the resources needed for each module and in what way they will interact with each other. The main sources of our model bellow in Figure 4-1are literatures review as we describe we conduct a comprehensive review of literatures to understand the research area and its problem domains, and based on our different experimentation.

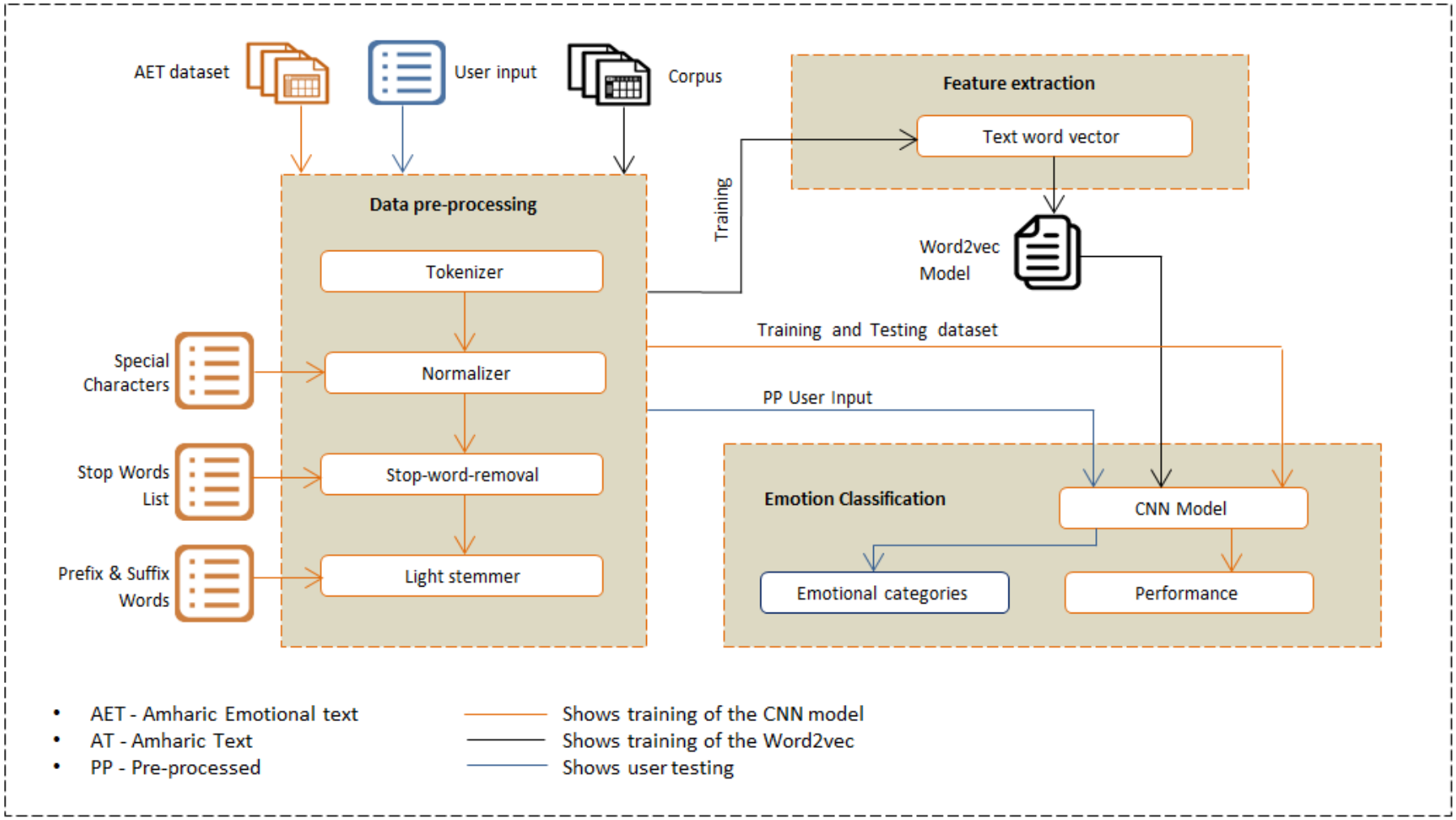

Figure 4-1 Emotion detection model for Amharic texts

4.3. Components of emotion detection model

4.3.1 Data Pre-processing

In this study, a crawler program web scraper was used to obtain emotional text from Facebook comments, the primary data source for our research FBC official Facebook page because this page was legal under the Facebook company terms and condition and people express his/her idea freely on social media. The pre-processing of collected data or text from social media involves the conversion of raw data into a clean data set. Therefore, certain steps are executed to convert the data into a small clean data set. At first, all the words in the data are tokenized to their individual words and as the next step, all the stop words are removed from the training data, next we remove stem words, finally remove special characters. Stop words are the commonly occurring words in the language that do not contribute any meaning that would benefit the text classification process. Removing numbers, accent marks, whitespaces and other diacritics are main preprocessing steps that are included during the design and development of our model.

Normalization and stop words removal

The data preprocessing included removing stop word and special characters and stopping words. First, special characters and stop words were removed to reduce the influence of experimental results. There are common stop words in Amharic which are used for grammatical purposes like ነዉ, ነበር, ሆኖም, እና, ነገርግን, prefixes የ, በ, ዛሆ, ስዛሆ, ሊይ and suffixes ን, ም, ና, ዎች etc which are non-informative to identify documents. In addition to the common stop words, there are also news specific stop words like ገለፁ, ዘግበዋል, አስታውቀ, etc these words do not carry important meaning and are usually removed from texts. These words are commonly occurring words in the language that do not contribute any meaning that would benefit the text classification process. It is not possible to remove Amharic stop words using NLTK, we designed our own stop words and special character removal because NLTK doesn‘t recognize Amharic stop words. Identification of stop words and special character enables language users to retrieve information fast and makes the language more powerful for information processing.

Algorithm 1 illustrates pseudo-code of preprocessing steps, stop-word removal, punctuation, and other diacritics.

```
For i=1 to Number of words in the document
   For j=1 to number of words in stopwords, punctuations, characters, ascii and numbers
    If
        Words (i) == stopwords (j) || punctuation (j) || characters (j) ||, ascii (j) || numbers (j)
             THEN
              Eliminate Words (i)
    End If
  End For
End For
```

Algorithm 4-1 Algorithm to remove stop-words, punctuation, characters, ascii and numbers.

Tokenization

Tokenization is the process of splitting an input sequence into tokens. The tokens usually consist of either a single word or what is called an Ngram, meaning that N consecutive words are split into a single token. The idea is to preserve some of the information that is stored in the order of the words. We took Amharic text and break it up into its individual words. These tokens are then used as the input for normalization, feature selection, and emotion classification. Deeplearning4j Java library recently developed by apache, we use this library for tokenization process. It is a framework with wide support for deep learning algorithms includes implementations of the restricted Boltzmann machine, ANN, DBN, CNN and RNN, word2vec and doc2vec.

Light stemming

The process of reducing a word to its word stems that affixes to suffixes and prefixes or to the roots of words known as a lemma. Stemming is important in natural language understanding and natural language processing. But in our project we use light steaming mean removal of suffix and prefix because as we know Amharic is one of under resource language there is no efficient morphological analyzer NLTK tool for this language. We found only one analyzer called Horn morph which is developed using machine learning algorithm it's not much efficient and also our model is developed using deep learning algorithm and DL4J library tools.

1. Let x number of words/text
2. For all “x” words , repeat 3-5
3. Cheek by prefix rules
   if match found apply, // prefix matching
   else go to step 4
4. Cheek by suffix rules
   if match founds apply, // suffix matching
   else go to step 5
5. Display the stem of word

Algorithm 4-2 Amharic text light stemming.

4.3.2 Feature extraction

An appealing aspect of many deep learning approaches is the ability to automatically extract features from largely unprocessed data. Feature extraction involves reducing the number of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power. Also it may cause a classification algorithm to over fit to training samples and generalize poorly to new samples. The input is large labeled data, and this data is tokenized, preprocessed, stemmed and features are extracted before it‘s used for training. This stage is where the dataset is transformed into a vector of numbers. The output from this stage is a fixed-size vector representation for each word.Word2vec using skip-gram model is used for generating word vectors. In this study we use Word2vec Skip-gram model for feature extraction.

After having this vector representation of the text using word2vec, let‘s see how word2vec with CNN works. Using theWord2vec emotion tagging was transformed into a word vector. When deep learning was used for text classification task, it was necessary to convert semi structured or unstructured text into a vector representation that could be understood and processed by the machine. In the training process of the Word2vec word vector, we chose the best parameter result through multiple rounds of iterative tuning based on the size of the experimental data set. By taking this feature vectors it feeds CNN with data. After the training data is fed, the model

building process starts to form a model considering the features and classifies the sentence to different categories. We used Deeplearning4j Java library for implementation of word2vec.

4.3.3 Emotion classifier

The classification process uses different machine learning algorithms to build a model. The input for the Sentiment classifier is the output of feature extractor which is a file containing the word vector of words from a given Amharic text. There are many works using deep learning for natural language processing and most use variations of the simple neural network in order to achieve their respective goal within their relevant context. However, they share a common foundation in taking raw textual data and representing the individual words or word pairs as vectors, to be used for further processing. Word embedding is a distributed representation of a word often a one-hot representation which is suitable for the input of neural networks where a word corresponds to a one-hot vector indicating a binary value denoting its presence in a document.

A convolution neural network is similar to a normal neural network in its operation but its architecture is better suited for image processing problems as the input vectors are in 3D form and can perform better than regular neural networks on a larger scale. The efficiency of a CNN has inspired use cases for natural language tasks by transforming one dimensional textual data into a matrix as the input for the network. In the training step, we used our data to incrementally improve our model‘s ability to predict whether a given comment is categorized to which emotional state. In this process of training the CNN model, we used Deeplearning4j Java library for implementation of CNN model with training data to learn from. When the classifier is trained accurately, it can be used to detect an unknown Amharic text. In our CNN model we use 6 input layer, fully connected layer and output layer (Softmax classifier).

This multi-channel convolutional neural network for emotion classification involves using multiple versions of the standard model with different sized kernels we use 6 kernels (2, 3, 4, 5, 6 and 7 kernel size used for 6 layers respectively). This allows the sentences to be processed at different n-grams (groups of words) at a time, whilst the model learns how to best integrate these interpretations.

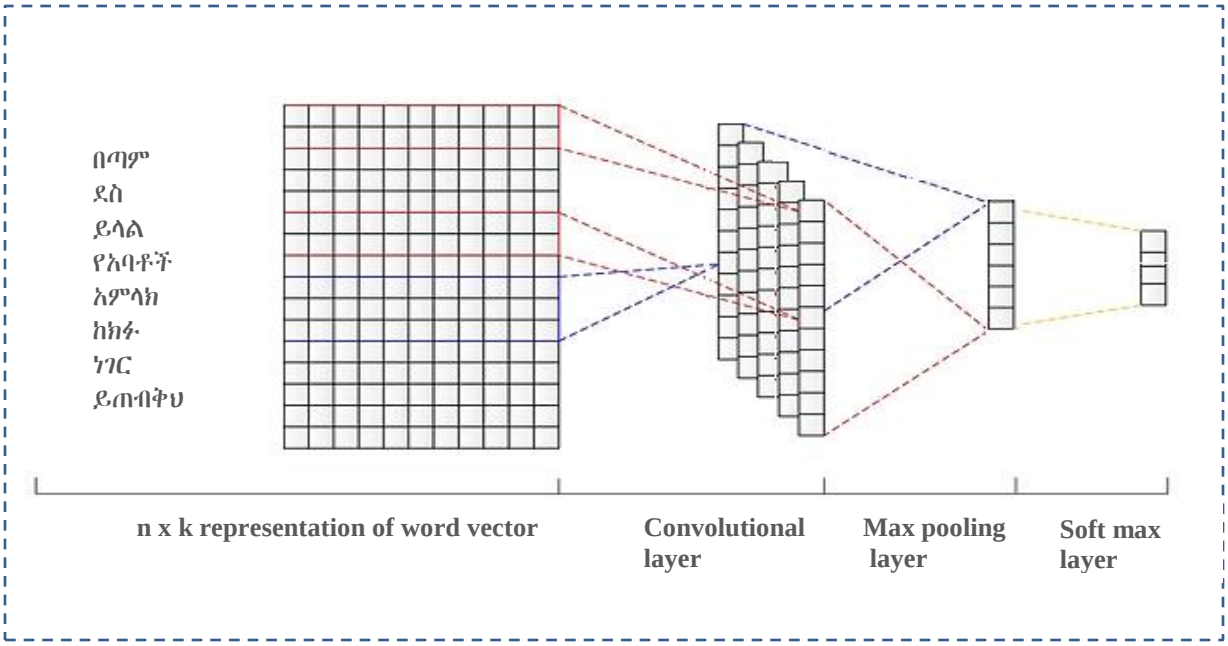

*Figure 4-2 Amharic-text convolutional neural network (*CNN*) model.*

The CNN's text classification results were divided into the four dimensions of emotion categories anger, disgust, happiness and sadness. As we described earlier in previous studies, CNNs have not only made great achievements in image processing but have also been proved to be effectively applied to natural language processing tasks, such as text classification and sentiment analysis. The basic CNN model is mainly composed of a convolutional layer, a pooling layer, and a fully connected layer. The network structure of the CNN (Figure 4-2) includes a convolutional layer input layers, output layer, a fully connected layer, and a softmax classification layer.

In our model we use 6 input layers with different kernel size (2-7 kernel size), activation layers, vertex layer that acts as a node in a computation graph, and output layer. After preprocessing the comment text the word vector training tool was used in the sentence feature layer to convert the sentence into a word vector representation. Sentence features are composed of a single word in the sentence, such as sentence S1 = [“ፈጣሪ ሆይ ግብፅን ሰላም ንሳት ለኢትዮጵይ ህዝብ የጨለማ እሾህ የሆነችብን”] (fetari hoye gebetsion selam nesat lethiopia hezebe yechelema eshohe yehonachehuben), sentences after word tokenization S1 = [“ፈጣሪ ”, “ሆይ ”, “ግብፅን ”,“ ሰላም ”, “ንሳት ”, “ለኢትዮጵይ”, “ህዝብ”, “የጨለማ ”, “እሾህ”, “የሆነችብን”]; the feature of S1 is represented by the set of each lexical feature after sentence tokenization. In the sentence feature

layer, by combining the features of the sentence, the number of features of the sentence increases greatly, and the feature composition of each sentence is inconsistent. In the training process of the Word2vec word vector, we chose the best parameter result through multiple rounds of iterative tuning based on the size of the experimental data set. The specific parameters in the word vector training process were set as follows vector length vector size = 100, window size window = 5, model was Negative Sampling model, neural network learning rate alpha = 0.025, word frequency min count was set to 5, and number of training iterations was 5.

The first layers embed words into low-dimensional vectors. The next layer performs convolutions over the embedded word vectors using multiple filter sizes. Next, we max-pool the result of the convolutional layer into a long feature vector, add dropout regularization, and classify the result using a softmax layer. Using six layers of CNN convolution operation, each CNN network layer contained multiple convolution kernels, each of which had a fixed size. Then the convolution operation and the output result used to the MAX pooling layer to further reduce the dimension of features. We applied a fully connected layer, obtained weights after all locations shared training, and output the results to the softmax layer. The softmax layer mapped the input to the interval (0, 1) through the softmax function to obtain the probability that the data to be classified belonged to four categories and finally obtained the prediction result of the emotion category.

The model architecture, shown in figure 3, has slight deferent of the CNN architecture of Kim [18]. Her we use 6 convolutional layers with 6 filter size with k-dimensional word vector corresponding to the n number of word in the sentence. We have described the process by which one feature is extracted from one filter. The model uses multiple filters (with varying window sizes from 2 -7) to obtain multiple features. These features form the convolutional layer and are passed to a fully connected softmax layer whose output is the probability distribution over four labels. For regularization we employ dropout on the vertex layer with a constraint on l2-norms of the weight vectors (0.5). Dropout prevents co-adaptation of hidden units by randomly dropping out.

Chapter 5: Evaluation and implementation of the proposed model

In this chapter, the experimental results of the developed prototype system, setups/procedures, the evaluation parameters, the tools used for implementing the prototype, the procedures to integrate the different components, the proposed algorithm, the input review, output result and other related issues are described. Since there is no readymade Amharic dataset, the experimentation is beginning with preparing of Amharic dataset that is used for the training, testing and word embedding. The dataset that we collect is a text file written (encoded) in Amharic alphabet. The collection of data is done from FBC officiate Facebook page. Those data collected from online sources are unlabeled texts. The collected dataset consists of 72,800 review sentence (comments) and more than 140,000 sentences for word embedding. The data are in four different emotional categories namely happiness, sadness, anger and disgust.

5.1 Experimental setup

In this project we use open-source programming languages Java specially DL4J library which has its own capability to accommodate NLP tasks. The experimented results are obtained using a PC with Intel(R) Core (TM) i5-5500U CPU @2.40 GHz with 8.0 GB RAM. The experiments have been developed using DL4J library with Java backend, CSS, Bootstrap and with tomcat server.

```
Number of threads used for OpenMP BLAS: 2
DefaultOpExecutioner - Backend used: [CPU]; OS: [Windows 10]
DefaultOpExecutioner - Cores: [4]; Memory: [2.0GB];
DefaultOpExecutioner - Blas vendor: [OPENBLAS]
```

Figure 5-1 Experimental setups

5.1.1 Emotional dataset

The main challenge we faced during data collection from the FBC is due to the lack of readily available data written in Amharic language, datasets are built manually with help of an IT expert from the our company. We use web scraper that is free web scraping web data extraction tool to scrape FBC data from Facebook. All the 72,800 sentence review from FBC page are manually categorized and labeled by an independent individual from our staff in to happiness, sadness, anger and disgust. There is no scientific rule or method for dividing of the dataset we get a best accuracy in 88% of training and 10 % of testing. So her we divided the training and testing sets from 72,800 of the total review sentence we use 64,000 (88%) and 8,800 (12%) for training and

testing our model respectively. Abbreviations and spelling errors were not considered and not added to the trained dataset after labeling. A sample of the dataset is summarized in Table 1.

Categories	Sample comment
Happiness	በጣም ደስ ይላል! የአባቶች አምላክ ከክፉ ነገር ይጠብቅህ
Sadness	መከላከያው መጥፎ አመራር ውስጥ በመውደቁ አዝናለሁ
Anger	የአማራ ህዝብ ቁጣ የመግለጫ ጋጋታን አያስተናግድም ድል ብሩህ አዕምሮ ለተቸው የአማራ ህዝብ
Disgust	የዜጎችን ደንነት ሳታረጋግጡ የደህንነት መስርያ ቤት ቱልቱላዎች

Table 5-1 Sample Amharic comment dataset

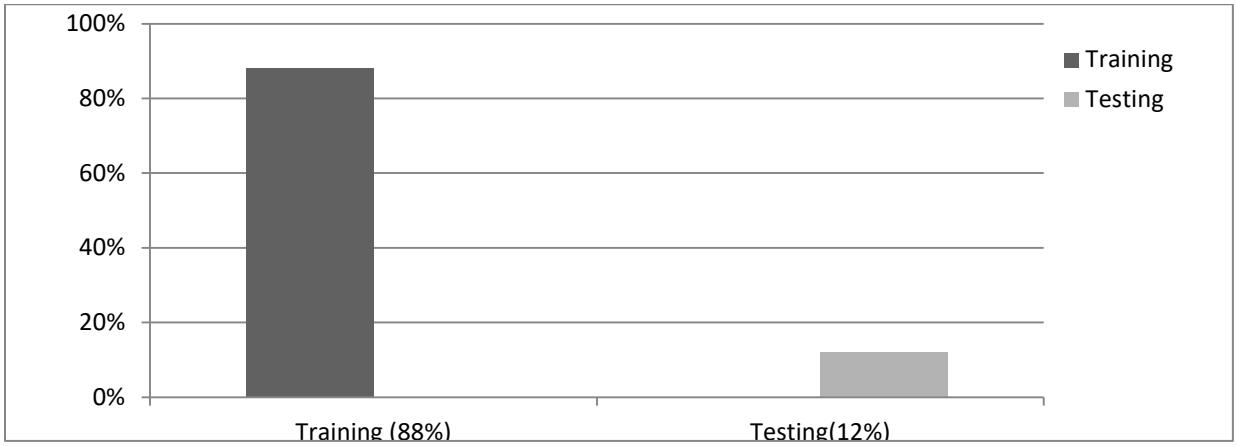

Figure 5-2 Amharic-text training and testing datasets

Similarly, from 72,800 of the total review sentence 18,200 (25%) are labeled as happiness, 18,200 (25%) sadness, anger 18,200 (25%) and disgust 18,200 (25%).

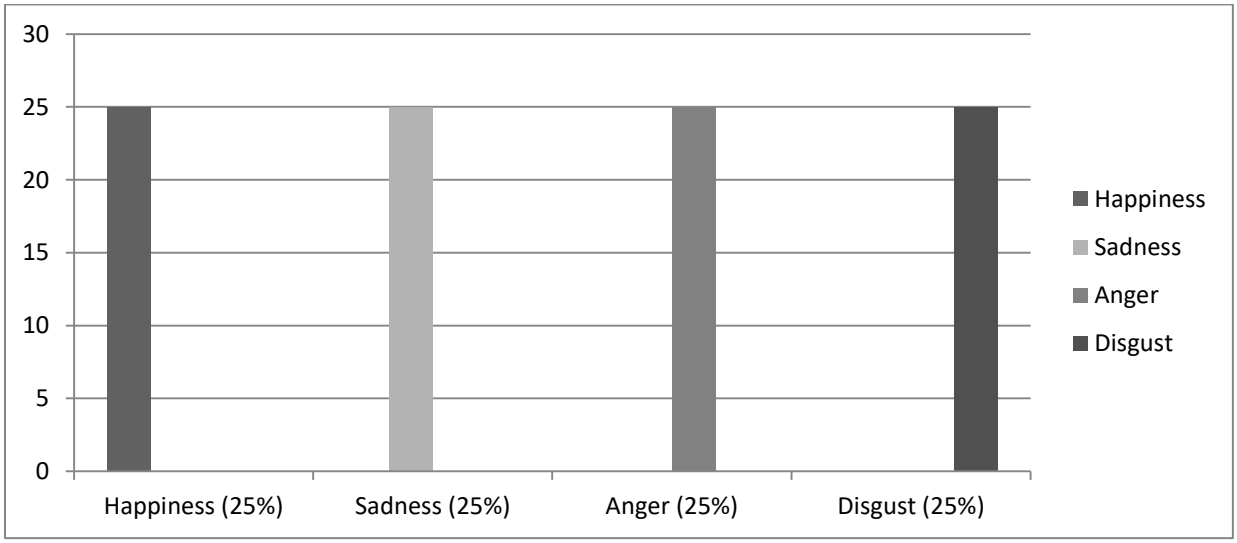

Figure 5-3 Amharic-text splinted as happiness, sadness, anger and disgust

5.1.2 Word embedding

After preprocessing tasks, the unlabeled free-text document is transformed into numeric values dense vector representation of words that capture something about their meaning. The main reason is to make a machine learning model to understand and process the natural language. Word embedding captured the contextually related meaning of a word in a document, syntactic, and semantic similarity, and relation with other words. We trained the data using one of word embedding model which is word2vec. And obtained a word2vec using the skip-gram method. Skip-gram is better for capturing infrequent words. We develop an unsupervised corpus that contains more than 140,000 words to support CNN embedding.

The implementation is using Java DL4J library and our pre-processing (stop word removal, light stemming and special character removal) algorithms and others methods. There are different word2vec parameters that we specify in our training. We reconfigured some of the most important parameters for our word2vec skip-gram model and use the rest default parameters. The parameters that are reconfigured in our model are listed below.

Word2Vec Parameter	**Value**
minWordFrequency	(5)
iterations	(5)
layerSize	(100)
seed	(42)
windowSize	(5)
iterate	(iter) setTokenPreProcessor
learningRate	(0.001)
tokenizerFactory	(t)
minWordFrequency	(5)
iterations	(5)

Table 5-2 Amharic-text word2vec parameters

Finally, the training result is returned in the file type that we want. It may be text file, or .model or .dat. This file consists of the number of unique vocabulary and dimension number 21,118 sizes of vocabulary, 100 vector size and 88,912 words.

```
o.d.e.a.m.t.c.PrepareWordVector - Fitting Word2Vec model for DLEDMAT....
o.d.m.s.SequenceVectors - Starting vocabulary building...
o.d.m.w.w.VocabConstructor - Sequences checked: [88912], Current vocabulary size:
[21118]; Sequences/sec: [16868.15];
o.d.m.e.l.WordVectorSerializer - Projected memory use for model: [16.11 MB]
o.d.m.e.i.InMemoryLookupTable - Initializing syn1...
o.d.m.s.SequenceVectors - Building learning algorithms:
o.d.m.s.SequenceVectors - building ElementsLearningAlgorithm: [SkipGram]
```

Figure 5-4 Amharic-text word2vec number of vocabulary and vector size

In addition, there is the vocabulary with its context in real number representation. For instance, a vocabulary "ቅራቅንቦ" represented as follow:

ቅራቅንቦ 0.0016111298464238644 -0.0035935817286372185 -0.00824932660907507 3.17157246172428l3E-4 0.0022003345657140017 0.0039100972935557365 -0.004496249370276928 -8.1561942351982E-4 -0.0015113254776224494 --0.005422754213213921 6.701859529130161E-4 0.0012756618671119213 -0.0020730302203446627 0.0029821591451764107 -0.0022298749536275864 -0.007514002732932568 0.0037293846253305674 -0.0016620977548882365 0.002279126551002264 -0.0023304461501538754 -2.041321567958221E-5 -0.004682666156440973 6.376933306455612E-4 -0.0034329514019191265 0.0011051499750465155 0.0029876413755118847 … Up to its dimension.

Figure 5-5 Amharic-text word2vec in real number representation

This real number result using word2vec built in function. To get top-N most similar words for a given query word we use most wordsNearest () function. In Figure below shows top-10 most similar words for a given word. And the real number value shows the cosine similarity between the query word and context word.

```
10 Words closest to 'ህዝብ': [ህብረተሰብ, ማህበረሰብ, ሴራ, ክልል, እያለቀ, ክልል, ካለቀ, ታሪክ, ጎን, መቸም]
10 Words closest to 'ኦሮሚያ': [ቤንሻንጉል, ዞን, ኦሮምያ, ክልል, ቤኒሻንጉል, ጉምዝ, ሻሸመኔ,ህዝብ, ጉሙዝ, አጣየ]
10 Words closest to 'ድል': [ማሸነፍ, ብርታት, መከላከያ, ሰራዊታችን, ትኖራለች, ኢትዮጵያ, ብሎናል, ኢትዮጵያዊያን,
ጁንታዎች, ሠራዊታችን]
10 Words closest to 'ጀግና': [ብርቱ, ግበዝ, ጀግኒት, ቆራጥ, ጀግኖች, ይገባታል, ህዝብ, ዚህች, ኑርልን]
10 Words closest to 'እግዚአብሔር': [አምላክ, ፈጣሪ, ይስጣቸው, ይባርክ, ይስጣችሁ, ይስጥህ, እግዚያብሔር, አሜን,
እውአምላክ, ይመስገን]
ህዝብ - ማህበረሰብ = 0.9650861620903015
ኦሮሚያ - ወለጋ = 0.6979387998580933
ድል - ሰራዊት = 0.74365009665489197
ጀግና - ብርቱ = 0.994432806968689
እግዚአብሔር - ፈጣሪ = 0.8858354091644287
```

Figure 5-6 Amharic-text word2vec similar words

5.1.3 Evaluation metrics

Evaluation of our model is made with the evaluation parameter that compares the number of reviews which are categorized correctly and incorrectly. The Amharic text /comments were pre-processed. First, the characters other than Amharic in a text were removed, and then multiple consecutive spaces were converted into a space. The space before and after the sentence were removed. Each word was separated by a space. Also stop word, numbers, light stemming and special characters were removed. Second, each character in a sentence was trained for word embedding by the DL4J library word2vec tool. When a word did not appear in a vocabulary for the first time, it would be added to the vocabulary. Finally, a dictionary containing 21,118 vocabulary sizes was obtained.

Finally, we use CNN model for emotion classification. There are different CNN parameters that we specify in our training. The use of 32 batch size, and the convolution filter weights RELUE and softmax weights were taken uniformly from the interval [0, 1]. The maximum pooling size was 4, and the pooled values were concatenated to produce a single vector at the fully connected dense layer which was used to calculate class probabilities. The model was trained using the "Adam" learning rate method with 25 epochs to avoid over fitting using a dropout. We used the dropout parameter P with p = 0.5 and the L2 regularization parameter value of 0.001 for the convolutional layer. The parameters that are reconfigured are listed below.

Parameter	**Value**
batchSize	32
vectorSize	100
nEpochs	20
cnnLayerFeatureMaps	100
activation()	Activation.LEAKYRELU
weightInit()	WeightInit.RELU
updater()	Updater.ADAM
convolutionMode()	ConvolutionMode.Same
l2()	0.000001
activation()	Activation.SOFTMAX

Table 5-3 Amharic-text CNN parameters

We designed a prototype that accepts raw Amharic texts/comments and detect the emotion of each comment then display the classification of the texts if it is sadness, happiness, anger and disgust and it displays the number of category each class of text. The application developed shows the user interface and output of the classifier. It allows the user to insert Amharic text and checks the emotional classification of the input text. After the inputs text first the text will be normalized (stop words removed, special character and light stemming) then feed into the classifier then it classifies give the output. We develop the prototype with java Backend, Bootstrap Frontend, DL4J libraries for our Word2vec and CNN classifier, we run over Tomcat 9.0 web server and we use Eclipse IDE (Eclipse-java-2020-12-R-win32-x86_64) development tool. Simple web application is shown in fig 10, 11 and 12.

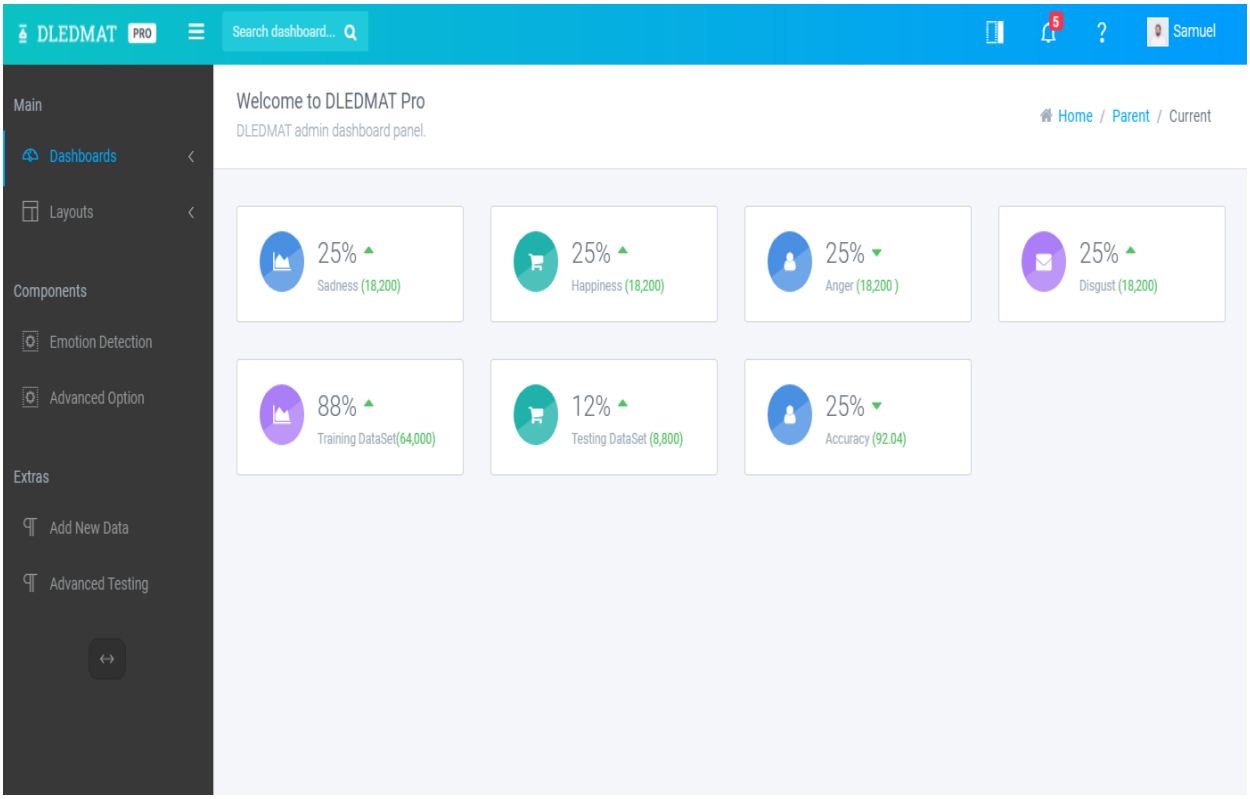

Figure 5-7 DLEDMAT model emotion detection page

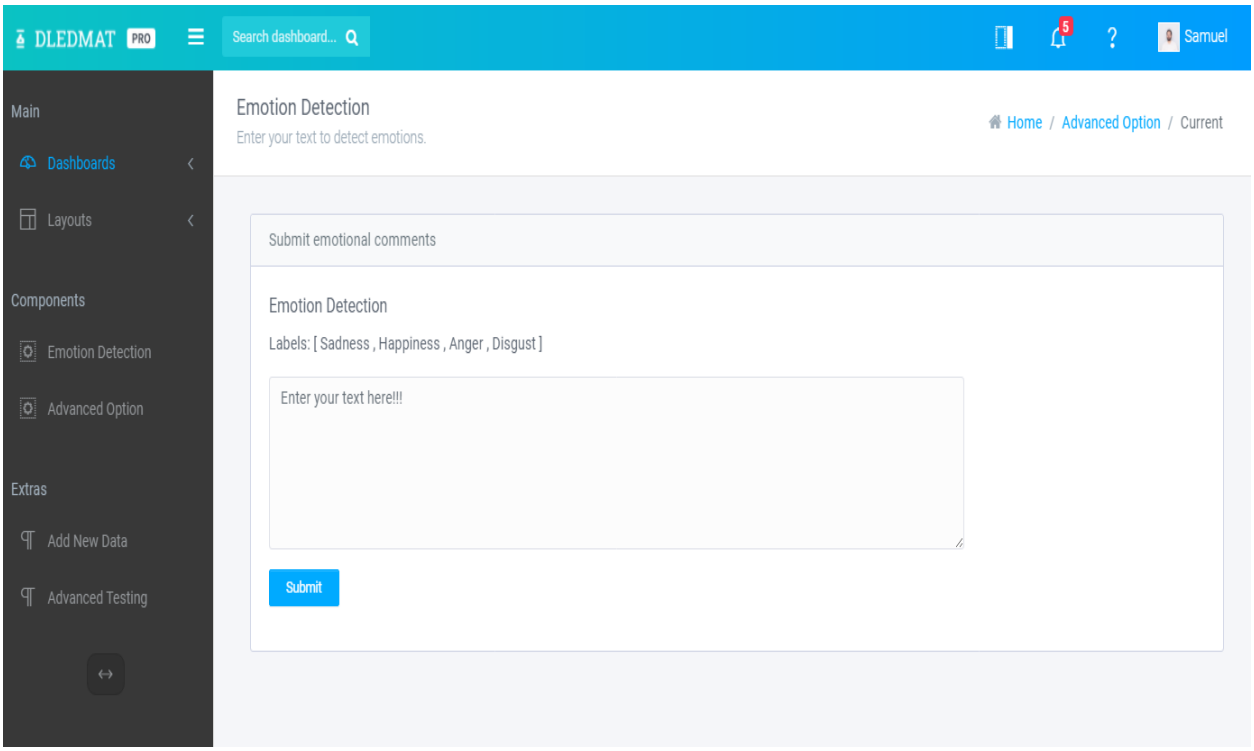

Figure 5-8 DLEDMAT model after detecting emotion from user input

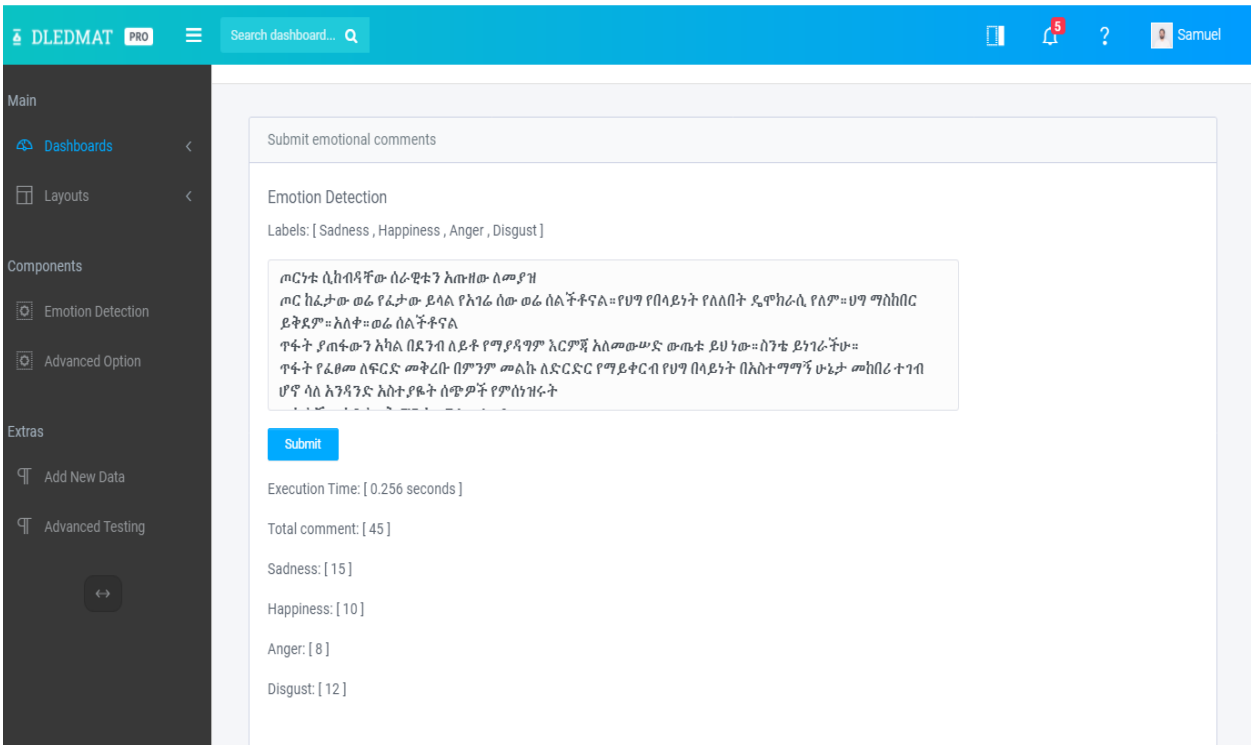

Figure 5-9 DLEDMAT model after detecting emotion result

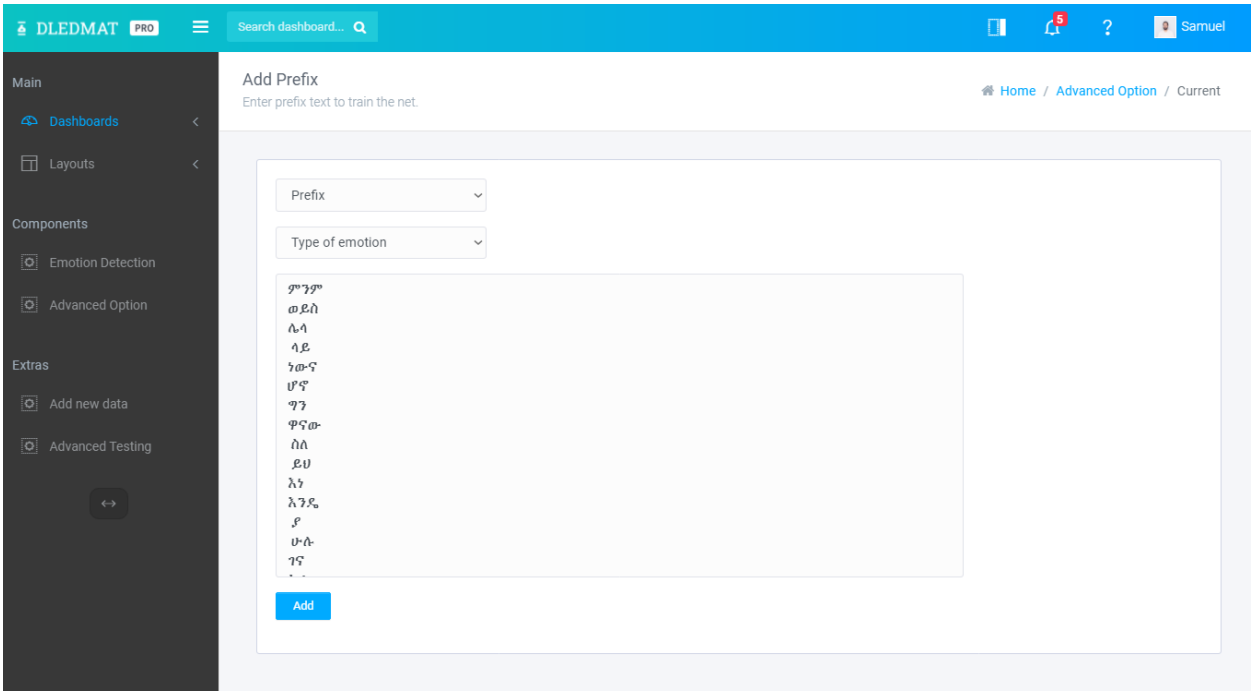

Figure 5-9 DLEDMAT for inputting prefix or others.

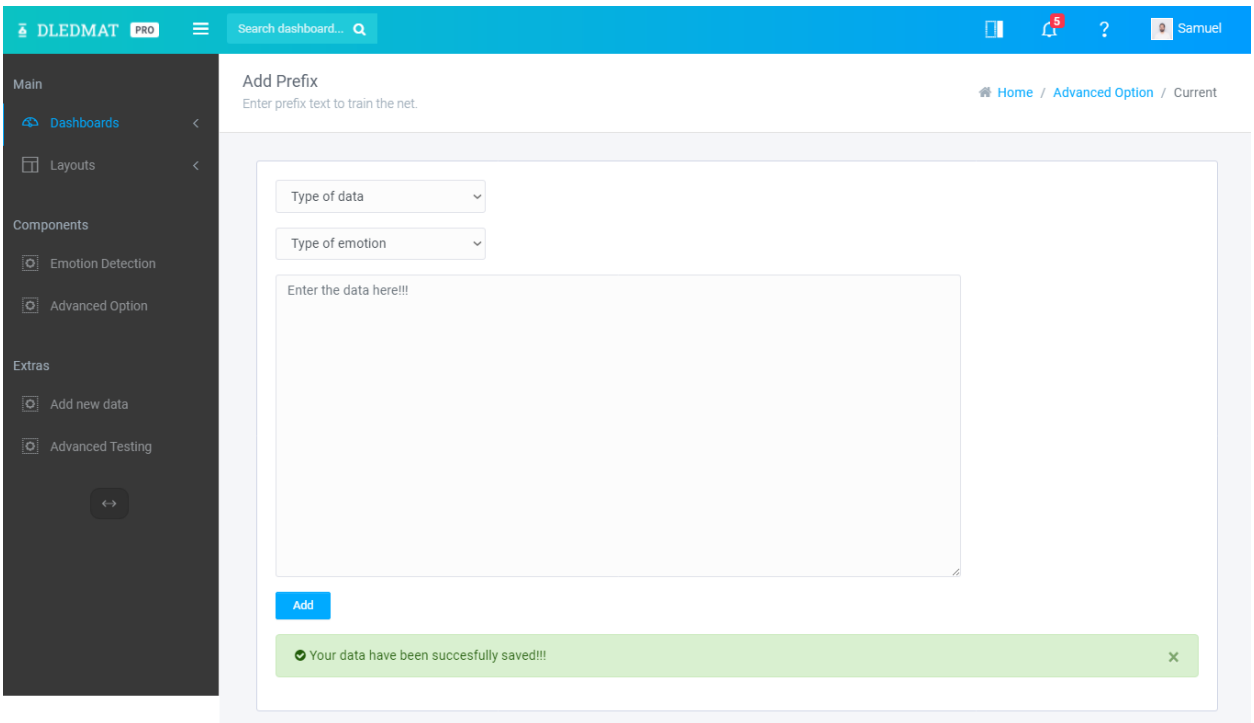

Figure 5-10 DLEDMAT after inserting prefix.

5.2 Experimentation result

In this section, we present the experimental results of the three different experiments. The first experiment is done using a CNN with word embedding using word2vec. The second experiment is conducted using CNN without word embedding. And finally, the result of the experiment conducted using the RNN with word embedding using weod2vec. Comparison of all the different experimental results is also presented in this section. All the dataset of emotional comments we created from social media are used for conducting the experiments. Each comment was classified by the system according to the procedures described earlier and all the results were recorded. Then the results of the two deep learning algorithms were compared with and without word embedding.

5.2.1. Experiment one

This experiment used the CNN with wrod2vec.The experiment is conducted on the emotional comments that divided as testing 12% of the total dataset. The results measured by accuracy, precision, recall and F-measure for each classification/output class is presented as follows.

```
========================Evaluation Metrics=========================
 # of classes:    2
 Accuracy:        0.8749
 Precision:       0.9040
 Recall:          0.8380
 F1 Score:        0.8697
Precision, recall & F1: reported for positive class (class 1 - "Happiness")

=========================Confusion Matrix==========================
  0    1
-----------
 1784 173 | 0 = Positive

 315 1629 | 1 = Negative
```

Figure 5-11 Experimental result of CNN with wrod2vec two outputs

```
=======================Evaluation Metrics=======================
 # of classes:    4
 Accuracy:        0.7101
 Precision:       0.7106
 Recall:          0.7105
 F1 Score:        0.7096
Precision, recall & F1: macro-averaged (equally weighted avg. of 4 classes)


=========================Confusion Matrix=========================
    0    1    2    3
---------------------
 1298  172  184  343 | 0 = Anger
  236 1315  180  226 | 1 = Disgust
  133  118 1565  128 | 2 = Happiness
  218  177  166 1409 | 3 = Sadness
```

Figure 5-12 Experimental result of CNN with wrod2vec four outputs

Number of Output / Classification	Accuracy	Precision	Recall	F1 Score
Two (positive, negative)	87.46%	90.71%	83.38%	86.89%
Four (happiness, sadness, anger and disgust)	71.01%	71.06%	71.05%	70.96%

Table 4 Experimental result of CNN with wod2vec compared to number of output

5.2.2. Experiment two

This experiment is conducted mainly to see the effect of using CNN without word2vec and need to observe did word2vec improved the classification model performance by learning the relationships among words. Also, this experiment is conducted on the emotional comments that divided as testing 12% of the total dataset. The results measured by accuracy, precision, recall and F-measure for each classification/output classes are presented as follows.

```
=======================Evaluation Metrics=======================
 # of classes:    4
 Accuracy:        0.5122
 Precision:       0.5390
 Recall:          0.5172
 F1 Score:        0.4859
Precision, recall & F1: macro-averaged (equally weighted avg. of 4 classes)

=========================Confusion Matrix=========================
 0 1 2 3
---------
 1298  172  184  343 | 0 = Anger
  236 1315  180  226 | 1 = Disgust
  133  118 1565  128 | 2 = Happiness
  218  177  166 1409 | 3 = Sadness
```

Figure 5-13 Experimental result of CNN without wod2vec

5.2.3. Experiment three

This is the last experiment conducted considering the CNN with word2vec and to compare with the RNN with word2vec.Also in this experiment is conducted to compare the number of output and accuracy using the emotional comments that divided as testing 12% of the total dataset. The results measured by accuracy, precision, recall and F-measure for each classification/output classes are presented in table 5.1 as follows.

```
=========================Confusion Matrix=========================
    0    1    2    3
----------------------
 1298  172  184  343 | 0 = Anger
  236 1315  180  226 | 1 = Disgust
  133  118 1565  128 | 2 = Happiness
  218  177  166 1409 | 3 = Sadness
Confusion matrix format: Actual (rowClass) predicted as (columnClass) N times
==================================================================

=======================Evaluation Metrics=======================
 # of classes:    4
 Accuracy:        0.6807
 Precision:       0.6822
 Recall:          0.6810
 F1 Score:        0.6806
```

Figure 5-14 Experimental result of RNN with wrod2vec

Number of Output / Classification	Accuracy	Precision	Recall	F1 Score
CNN with Word2vec	71.01%	71.06%	71.05%	70.96%
CNN without Word2vec	51.22%	53.90%	51.72%	48.59%
RNN	68.07%	68.22%	68.10%	68.06%

Table 5-4 Compression of Experimental result

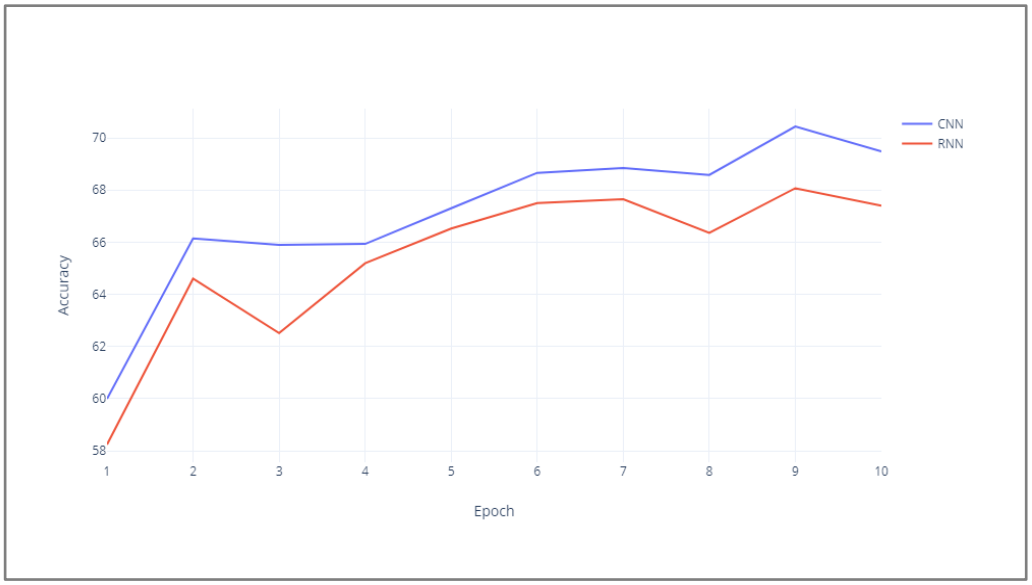

Figure 5-14 Compressions of CNN and RNN.

In the above figure 17, we evaluate the two deep leering algorithms for our model as component of classifier. In this two experiment we use 10 epochs for detecting the accuracy of the model in the case of CNN with word2vec we get 71.01% of accuracy at epoch 9 and RNN 68.07% accuracy also at epoch 9. In general, CNN can abele to handle the longer dependency of the words through different convolutional filters. When the context of the word is used to determine the polarity of the text rather than the probability of the occurrence of the word both CNN and RNN are the best approaches. The RNN by its nature has the capability to hold relevant information to the task at hand. This makes it better for text classification and sentiment analysis tasks but relatively slower computational time than CNN.

Related work using different approach	Classification	Accuracy
Wondwossen Philemon and Wondwossen Mulugeta "multi-scale approach to sentiment analysis"	Positive , Negative	44.3%
Maha Heikal, Marwan Toki and Nagwa EI-Makky "Arabic sentiment analysis "	Positive , Negative	64.46%
Zewdie Mossie and Jenq-Haur Wang "apache spark based model to classify Amharic Facebook posts and comments"	Hate , Hate	79.83%
Hiwot Wonago Kululo , "Information filtering of social media Amharic texts based on Sentiment Analysis"	Offensive, Non-Offensive	72%
Dongliang Xua,b, Zhihong Tiana, Rufeng Laic, Xiangtao Kongb, Zhiyuan (Thomas) T and Wei Shie "Microblog emotion classification model"	Positive, Negative	97.6%
Eman Hamdi(&) , Sherine Rady, and Mostafa Aref, "CNN model for Emotion Detection from Tweets "	Positive , Negative	80.6%
Deep Learning Based Emotion Detection Model for Amharic Text (CNN with word2vec)	Positive, Negative	87.46%
Deep Learning Based Emotion Detection Model for Amharic Text (CNN with word2vec)	Happiness, Sadness, Anger, Disgust	71.01%

Table 5-5 Compression of DLEDMAT with related work

5.3 Discussion of the results

As shown above, the experiment is done with different deep learning algorithms, experimental setups and has shown us promising results. The experimental setups are the reasons for the variations of experimental results. Also we evaluate the system using metrics like accuracy, F1 score, recall, and precision. As shown in Table 4, the result of experiment 1 uses CNN-Word2vec with two and four number of output category or classification. Our model CNN-Word2vec for two number of output or category (positive and negative) perform accuracy 87.49%, precision 90.40%, recall 0.8380% and F1 Score 86.97% respectively. The model was tested to classify whether the post and comments are happiness, sadness, anger and disgust and able to detect and

classify achieved average of accuracy 71.01%, precision 71.06%, recall 71.05% and F1 Score 70.96% respectively. The result form this experiment shows very promising and this experiment implies directly the strength and efficiency of the proposed approach is depend on the number of output category used, this shows the performance decreased with regard to the number of classification increases and this is due to both languages the model complexity.

On the other hand, from the result of experiment 2 that is shown in Table 5, the performance of the approach CNN without Word2vec is decreases accuracy 51.22%, precision 53.90%, recall 51.72% and F1 Score 48.59% respectively. Furthermore, from the result of the experiment 3 shown in figure 16, the performance of the model accuracy 68.07%, precision 68.22%, recall 68.10% and F1 Score 68.06% respectively.

Finally, we illustrated the compression of the evaluation results of experiment and related work in Table 5 and 6. For the purpose comparison of our model with related work, we use our model results in two ways first with two numbers of outputs, second with four numbers of outputs and we use to measures accuracy of our model for compression with related work. Our model has shown to work well with word2vec and with two number of output. Also, it has been shown that CNN can work for Amharic emotion detection and showed a considerable performance on CNN than RNN.

Chapter 6: Conclusion and recommendation

6.1. Conclusion

The Social media has dramatically changed the way that people express their emotions and fleeing. They can now comment on products at web page and express their view on almost everything in Internet forums, discussion groups and blogs. This online word of mouth behavior represents new and measurable sources of information with many practical applications. Now if one wants to buy a product, he/she is no longer limited to asking his/her friends and families because there are many product reviews on the web which gives emotion of existing users of the product. For a company, it may no longer be necessary to conduct surveys, organize forum groups or employ external consultant to find consumer emotion. However, it is difficult for a human reader to find relevant sources, extract related sentences with emotion, read them summarize them and organize them into useful forms. As a result, automated emotion detection systems are needed. Emotion analysis from a text mining grows out of this need. Due to its tremendous value for practical applications, there has been an explosive growth of both research in academics and applications in the industry.

An effort was made to develop sentiment analysis for Amharic. Considering the morphological complexity of the language in this research light stemming was applied in the preprocessing stage. Character normalization and stop words removal was used while implementing and designing the system. Evaluation of the system was done by comparing the two most known classifiers CNN and RNN also with selected related works. The experimental result in this study showed a considerable performance on CNN with Word2vec classification. An accuracy 87.49%, precision 90.40%, recall 0.83% and F1 Score 86.97% achieved for two number of output respectively and the proposed model achieved average of accuracy 71.01%, precision 71.06%, recall 71.05% and F1 score 70.96% the high performance of our model in sentence level emotion classification in four number of output recorded during testing. We conclude that to the best of our knowledge, this is a first research work on Amharic emotion detection on sentence level with four different classifications, and achieve better result both in the completeness as well as the correctness of the implemented model.

Overall, we conclude that there is no single technique in the text classification or sentiment analysis domain that would contribute to the accuracy dramatically. Each component in our model contributes to better accuracy and must be adjusted to fit the problem domain.

6.2 Contribution of the study

Some of the major contributions of this study include:

- Designing a model for emotion detection of Amharic texts.
- We consider designing our own Amharic emotional dataset that contains above 70,000 Amharic emotional comments as a contribution to this research area.
- We develop an unsupervised corpus that contains more than 140,000 words to support CNN embedding.
- This work allows an immediate comparison of current state-of-the-art text classification techniques in the context of the Amharic language.
- We develop sentence level with four category of emotion analysis for Amharic texts so that we created a general understanding of the subject matter.
- Applying word2vec and CNN algorithm using DL4J library to optimize the accuracy of our model and compare the performance with RNN.

6.3. Recommendations and future work

The system designed in this study is deep learning based emotion detection for Amharic text using word embedding and CNN. In this study we include only four emotional states (Anger, Sadness, Happiness and Disgust) but there are many types. So we recommend further research in this study to fully explore the problem of emotion detection of posts/comments as well as the improvement of our experiment results. We have identified the following future and the results we found in this research showed that emotion detection can be done for under-resourced language. However, to enhance more the quality and performance of the emotion detection, the following ideas are recommended for further research work.

- Expanding this research by adding additional emotion detection ways like emoji.
- In this research, due to lack of efficient morphological analyzer, we just use light stemming for prefix and suffix removal, Amharic text emotion detection system can be implemented using with morphological analyzer for preprocessing and the performance and accuracy of the model can be improved.
- Enhances the model with different sense of word and abbreviations.
- Enhances the model with additional emotional states, in this model we includes only four emotional states (Anger, Sadness, Happiness and Disgust).

References

1. A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of twitter data," in Proceedings of the workshop on languages in social media. Association for Computational Linguistics, 2011, pp. 30–38.
2. Hiwot Wonago Kululo, Information filtering of social media Amharic texts based on Sentiment Analysis, 2020.
3. Pang, B., Lee, L.: Opinion Mining and Sentiment Analysis, Foundations Trends Inf. Retrieval 2(1-2), 1–135 (2008) Geez to Amharic Machine Translation.
4. META-NET White Paper Series, Retrieved from Multilingual Europe Technology Alliance, 2018.
5. "*META-NET White Paper Series*", (2018, January 23). [Online]. Available: http://www.meta-net.eu/whitepapers/overview
6. "*Internet World Stats*", http://www.internetworldstats.com/africa.htm#et.11-18-2020. [Online]. Available: (2018, January 21 4 pm)
7. M Haenlein and A. M, Kaplan, An empirical analysis of attitudinal and behavioral reactions toward the abandonment of unprofitable customer relationships. Relatsh, Mark, vol. 9, no. 4, pp. 200228, 2010.
8. J Singh, G Singh, and R Singh, A review of sentiment analysis techniques for opinionated web text, CSI Trans ICT, 2016.
9. M. Day and C. Lee, Deep Learning for Financial Sentiment Analysis on Finance News Providers, no. 1, pp. 11271134, 2016.
10. Ethnologue, languages of the world Wikipedia, the free encyclopedia W. John Hutchins, Machine translation: A brief history, 1995.
11. "*Amharic*"Wikipedia, [Online]. Avhttps://en.wikipedia.org/wiki/Languages_of_Ethiopia, 11-19-20202. (2018, January 23).
12. Dongliang Xu, Zhihong Tian, Rufeng Lai, Xiangtao Kong, Zhiyuan (Thomas) Tan, Wei Shi PII: Deep Learning Based Emotion Analysis of Microblog Texts: S1566-2535(20)30302-X DOI
13. Zewdie Mossie1 and Jenq-Haur, SOCIAL NETWORK HATE SPEECH DETECTION FOR AMHARIC LANGUAGE Wang2 1Department of International Graduate Program in Electrical Engineering and Computer Science.2018.
14. Sumayh S. Aljameel *, Dina A. Alabbad, Norah A. Alzahrani, Shouq M. Alqarni, Fatimah A. Alamoudi, Lana M. Babili, Somiah K. Aljaafary and Fatima M. Alshamrani ,A Sentiment Analysis Approach to Predict an Individual's Awareness of the Precautionary Procedures to Prevent COVID-19 Outbreaks in Saudi Arabia.
15. Wondwossen Philemon and Wondwossen Mulugeta, A Machine Learning Approach to Multi-Scale Sentiment Analysis of Amharic Online Posts HiLCoE, Computer Science Programme, and Ethiopia wondaphil@yahoo.com.
16. Maha Heikal, Marwan Torki and Nagwa El-Makky ,The 4th International Conference on Arabic Computational Linguistics (ACLing 2018), November 17-19 2018, Dubai, United Arab Emirates Sentiment Analysis of Arabic Tweets using Deep Learning .
17. Yeshiwas Getachew and Abebe Alemu, DEEP LEARNING APPROACH FOR AMHARIC SENTIMENT ANALYSIS UNIVERSITY OF GONDAR Article · November 2018,
18. Yoon Kim, Convolutional Neural Networks for Sentence Classification, New York University, 2014.

19. "*Design science (methodology*)", From Wikipedia, the free encyclopedia [Online]. Available: https://en.wikipedia.org/wiki/Design_science, (2021, Jun 21 6 pm).
20. https://www.forbes.com/sites/forbestechcouncil/2017/06/05/the-big-unstructured-data-problem/?sh=7418cf07493a, (2021, August 5, 6 pm).
21. "*The 6 Types of Basic Emotions and Their Effect on Human Behavior*" [Online]. https://www.verywellmind.com/an-overview-of-the-types-of-emotions-4163976, (2021, January Tuesday, 4 pm).
22. Teshome Kassie, WORD SENSE DISAMBIGUATION FOR AMHARICTEXT RETRIEVAL: A CASE STUDY FOR LEGAL DOCUMENTS. 2019.
23. "*Amharic*" From Wikipedia [Online]. Available: https://en.wikipedia.org/wiki/Amharic, (2021, January Tuesday, 4 pm).
24. "*Africa Internet Usage, 2020 Population Stats and Facebook Subscribers*", [Online]. Available. Https://www.internetworldstats.com/stats1.htm.(2021, Jun 21 6 pm).
25. Maria D. Lozano , Emotion Detection: A Technology review1 Jose Maria Garcia-Garcia Computer Science Research Institute University of Castilla-La Mancha Albacete, Spain josemaria.garcia14@alu.uclm.es Victor M. R. Penichet University of Castilla-La Mancha Albacete, Spain Victor.Penichet@uclm.es University of Castilla-La Mancha Albacete, Spain.
26. A. Kleinsmith and N. Bianchi-Berthouze. Affective Body Expression Perception and Recognition: A Survey. In IEEE Transactions on Affective Computing, vol. 4, no. 1. Jan.-March 2013. 15-33. DOI: 10.1109/T-AFFC.2012.16.
27. "*What Are Word Embedding for Text*" by Jason Brownlee on October 11, 2017 [Online]. Available. https://machinelearningmastery.com/what-are-word-embeddings, (2021, July 21 6 pm).
28. "*Fast Text Amharic Embedding Vectors*" [Online]. Available: https://github.com/Abe2G/FastText-Amharic-Embedding-Vectors, (2021, July 21 6 pm).
29. Edward Chao-Chun Kao Chun-Chieh Liu , Towards Text-based Emotion Detection A Survey and Possible Improvements , Institute of Information Systems and Applications National, 2016.
30. D. B. Bracewell, "Semi-Automatic Creation of an Emotion Dictionary Using WordNet and its Evaluation," Proc. IEEE conference on Cybernetics and Intelligent Systems, IEEE Press, Sep. 2008, pp. 21-24, doi: 10.1109/ICCIS.2008.4670735.
31. Abdullah et.al, S. S. (2018). SEDAT: Sentiment and Emotion Detection in Arabic Text Using CNN-LSTM Deep Learning 2018 17th IEEE International Conference on Machine Learning and Applications IEEE. 2019.
32. RULE-BASED SENTIMENT ANALYSIS OF UKRAINIAN REVIEWS Mariana Romanyshyn Department of Computer-Aided Design, Lviv Polytechnic National University, Lviv, Ukraine mariana.scorp@gmail.com.
33. "*Different Types of Learning in Machine Learning*" [Online]. Available https://machinelearningmastery.com/types-of-learning-in-machine-learning, (2021, July 21, 5 pm).
34. "Emotion recognition", [Online] https://en.wikipedia.org/wiki/Emotion_recognition. (2021, July 21, 7 pm).
35. Eman Hamdi(&) , Sherine Rady, and Mostafa Aref, sentiment analysis using convolutional neural network model for Emotion Detection from Tweets, 2015.

36. DeepanshuVijay and Aditya Bohra Corpus Creation and Emotion Prediction for Hindi-English Code-MixedSocial Media Text, 2018.
37. Sheeba Grover and Amandeep Verma, Design for emotion detection of punjabi text using hybrid approach, 2016.
38. Johanes Effendi, Alfan Farizki Wicaksono and Mirna Adriani, two-stage emotion detection on Indonesian tweets, 2015.
39. Omneya Rabie, and Christian Sturm, Feel the Heat: Emotion Detection in Arabic Social Media Content, Proceedings of the International Conference on Data Mining, Internet Computing, and Big Data, Kuala Lumpur, Malaysia, 2014.
40. Malak A. Abdullah, Mirsad Hadzikadic and Samira Shaikh, SEDAT: Sentiment and Emotion Detection in Arabic Text Using CNN-LSTM Deep Learning, Conference: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA).
41. Gebremeskel, Selama, Sentiment Mining Model for Opinionated Amharic Texts, 2010.
42. Melese Mihret and Muluneh Atinaf, Feature Level Sentiment Analysis Model for Awngi Text, 2014.
43. Abreham Getachew, Opinion Mining from Amharic Entertainment, October-2014.
44. Tulu Tilahun Hailu, Linguistic Localization of Opinion Mining From Amharic Blogs, 2014.
45. P. Vateekul and T. Koomsubha, A Study of Sentiment Analysis Using Deep Learning Techniques on Thai Twitter Data, 2016.
46. Abreham, G. (2014). OPINION MINING FROM AMHARIC ENTERTAINMENT TEXTS OPINION MINING FROM AMHARIC ENTERTAINMENT TEXTS, Addis Abeba Universty, Addis Abeba, 2014. Addis Abeba, Ethiopia: Addis Abeba Universty.
47. Cheng, B., & Titterington, D. (1994). Neural networks: a review from a statistical perspective. Statistical Science 9, 2–54.
48. Effendi et.al, A. F. (2015). Two-Stage Emotion Detection on Indonesian Tweets, ICACSIS, 143-146.
49. Gaind et.al, V. S. (2019) Emotion Detection and Analysis on Social Media. Global Journal of Engineering Science and Researches (ICRTCET-18) (2019) 78-89.
50. Dongliang Xua,b, Zhihong Tiana, Rufeng Laic, Xiangtao Kongb, Zhiyuan (Thomas) Tand and Wei Shie , (13 June 2020) Deep Learning Based Emotion Analysis of Microblog Texts.
51. Ashish Katrekar AVP, an Introduction to Sentiment Analysis, Big Data Analytics, 2015.
52. Mind Tools. 2017. Body Language Understanding Non-Verbal Communication. Retrieves from: https://www.mindtools.com/pages/article/Body_Language.htm.
53. Emotion Detection State of the Art Haji Binali School of Information Systems Curtin University Perth, Australia h.binali@postgrad.curtin.edu.au Vidyasagar Potdar School of Information Systems Curtin University Perth, Australia.
54. J. Yang, D. B. Bracewell, F. Ren, and S. Kuroiwa, “The Creation of a Chinese Emotion Ontology Based on HowNet”, Engineering Letters, vol. 16(1), Feb. 2008, pp. 166-171.

Appendix 1

1.1. List of stop word

ንእ	ለዚያው	ስለዚሁ	በእነዚህ	እረ	ከአንዳንድ	የአንዱ
አለ	ለይ	ስም	በእንደዚህ	እራሱ	ከአንድ	የአንዱን
አሉ	ሊሆን	ስምንት	በእያንዳንዱ	እር	ከእነዚህ	የአንድ
አስራ	ላይ	ስንት	በዋና	እቺ	ከዚሁ	የዚሁ
አንቀጽ	ላይም	ስድስት	በዚህ	እነዚሁ	ከዚህ	ያህል
አንዱን	ሌላ	ስጥ	በዚህም	እነዚህ	ከዚህም	ያለ
አደር	ሌላው	ሶስተኛ	በዚህና	እነዚህን	ከዚያ	ያለበት
እነዚህም	ሌላውን	ሶስት	ቢሆን	እና	ካልሆነ	ያለን
እነዚህኑ	ሌሎች	ራሱ	ቢሆንም	እናተ	ወር	ያለው
እነደሆነ	ሌሎች	ሸለቀ	ቢስ	እንኳን	ወይ	ያለውን
እና	ልክ	ሺህ	ቢያንስ	እንደ	ወይም	ያለዉ
እንኳ	ልዩ	ቁጥር	ባለ	እንደ	ወይስ	ያሉ
እንደነዚህ	መ	በሁለቱም	ባሉ	እንደሆነ	ወይዘሮ	ያላቸው
እንዲሁም	መሆኑ	በሁለት	ብሎ	እንደሆነና	ወደ	ያላቸውን
እንጂ	መሆኑን	በሁኔታው	ብቻ	እንደሆኑ	ውስጥ	ያልሆነ
ከሀምሳ	መሆናቸው	በኡላ	ተራ	እንደዚህ	ውስጥ	ያንቺ
ወይም	መሆናቸውን	በሆነ	ተብሎ	እንደገና	ውስጥ	ያው
የት	መሆን	በሆነው	ተው	እንዲህ	ውይ	ይህ
ሀያ	መሰረት	በሆኑ	ነበር	እንዲሆኑ	ዚ	ይህን
ሁለቱ	መቶ	በሉ	ነች	እንዲሆን	የ	ይሁን
ሁለቱንም	መች	በላይ	ነችን	እንዲቆይ	የሆነ	ይሄ
ሁለት	ማለት	በሌላ	ነው	እንዴት	የሆነችን	ይህ
ሁሉ	ማናቸ	በሌሎች	ናቸው	እንጂ	የሆነው	ይህም
ሁነው	ማናቸውም	በመሆን	ናቸውን	እኛ	የሆነውን	ይህችው
ሁኔታ	ማናቸውንም	በሙሉ	አሁን	እኮ	የሆኑ	ይህን
ሆነ	ማንም	በሙሉም	አለ	እዚሁ	የሆኑበታል	ይህንኑ
ሆነው	ማንኛውም	በሚል	አለበት	እጅግ	የሆኑት	ይህንን
ሆና	ማንኛውንም	በሚችሉ	አለው	ከሀያ	የሆኑትን	ይህንንም
ሆን	ማዘዝ	በሚገባ	አለው	ከሁለት	የሆናል	ይሆናሉ
ሆን½	ምኑ	በማለት	አሉ	ከሆነ	የለውም	ይሆናል
ሆኖ	ምኑ	በማቀድ	አላት	ከሆነና	የሌለው	ይሆን
ሆኖም	ምናለ	በማናቸውም	አምስት	ከሆነው	የሌላ	ይላል
ሆይ	ምን	በማይበልጥ	አስር	ከላይ	የሌላቸውን	ይም
ነገሩ	ምንም	በስተቀር	አስከ	ከሌላ	የሌላውን	ይሻላል
ለ	ምንጊዜም	በሶስተኛ	አረ	ከሌሎች	የሌሎች	ይኸኛው
ለሆነ	ሰላሳ	በራሱ	አንቀጽ	ከመቶ	የሚሆነው	ይኸው
ለሆነው	ሰባት	በቀር	አንቀጽሀ	ከማናቸውም	የሚሆኑ	ይፋ
ለሆኑ	ሰኣት	በተለይም	አንደሆነ	ከሰላሳ	የሚሆኑት	ደስ
ለሆኑት	ሲሆኑ	በታች	አንደኛ	ከሰባት	ሆነ	ጀምሮ
ለሌላ	ሲሆን	በነሱ	አንዱ	ከስምንት	የሚችለውን	ጊዜ
ለሌሎች	ሲሆንና	በነዚሁ	አንዲት	ከስድስት	የማናቸውም	ጊዜም
ለማናቸውም	ሲል	በነዚህ	አንዳንድ	ከሶስት	የማያንስ	ጋር
ለምን	ሲባል	በኔ	አንዴ	ከርሱ	የማይበልጥ	ግን
ለነዚህ	ሲኖር	በአምስት	አንድ	ከብር	የት	ግድ

ለነገሩ	ሳለ	በአንቀጽ	አይ	ከነበረው	የትም	የኔ
ለአንድ	ሳይሆን	በአንዱ	አይሆንም	ከነዚሁ	የነበረውን	አየ
ለእነርሱም	ስለ	በአንዳንድ	አይደል	ከነዚህ	የነገሩ	ናት
ለእያንዳንዱ	ስለሆነ	በአንድ	እ	ከአምስት	የኛ	እንካን
ለዚሁ	ስለማናቸውም	በእስር	እስ	ከአስራ	የአንቀጹ	እንዴ
ለዚሁ	ስለዚ	በእነዚሁ	እስከ	ከአስር	ከአንዳንድ	

1.2. List of prefix

ለ	ስለ	አ	እስኪ	እንደ	ከእነ	ቅድመ	ሳት	ቢ	አል
ለእነ	በ	አል	እስኪ	እንደየ	ከየ	በ	ላት	ባይ	እንዲ
ሊ	በእነ	አን	እስኪ	እንዱ	የ	ለ	ሲ	ስለ	ሳት
ሌየ	በየ	አንድ	እስክት	እንዲ	የሚ	የ	ከ	አ	ባለ
መ	ቢ	አየ	እስክን	እንድን	የእነ	ባ	ሳይ	አይ	
ሲ	ቤየ	እ	እነ	እየ	ይ	እየ	ላይ	ካል	
ሲ	ት	እስከ	እን	ከ	ካለ	እንደ	በነ	ሊ	
ስ	ን	እስከን	እንደ	ከነ	ሳለ	ይ	ለ	ል	

1.3. List of suffix

ሁ	ምቹ	ምውም	ባት	ትም	ናምን	ናውን	ኡት	ዎቹ	ዎችም	ውም
ህ	ምና	ምውና	ባቸው	ቸው	ናን	ን	ኢት	ዎቹም	ዎችምና	ውምና
ለት	ምናም	ምውን	ባችሁ	ቹ	ናንም	ንና	ኤ	ዎቹምና	ዎችና	ውና
ለችው	ምናን	ሰ	ባችው	ችን	ናንና	ንኦ	ኦች	ዎቹምን	ዎችናም	ውናም
ላት	ምን	ሳ	ተው	ነት	ናንን	ኛ	ክ	ዎቹና	ዎችናን	ውን
ላቸው	ምንም	ሽ	ቱ	ነና	ናው	ኚ	ክኡ	ዎቹናም	ዎችን	ዉናን
ላችሁ	ምንን	ሽን	ባት	ናም	ናውም	ኞች	ዋ	ዎቹን	ዎችንም	ዉንዉ
ማ	ምቹ	በት	ባቸው	ናምና	ናውና	አዊ	ው	ዎች	ዎችንና	ዉኦ

1.4. List of synonyms

ኃይል	"ብርታት", "አቅም","ጥንካሬ"	ጀግና	ግበዝ	ፍርድ	ፍትህ
ተግባር	"ክንውን"," ድርጊት"	በርቱ	ጠንክሩ	ድንቅ	የሚገርም
ክልክል	"ግዝት", "እርም", "ሀራም"	ይሻላል	ይበጃል	ያስከፋል	ያሳዝናል
ህመም	"በሽታ", "ደዌ", "ስቃይ"	መልካም	ጥሩ	ውጊያ	ጦርነት
መንግስት	"ህዝብ", "ማህበረሰብ", "ህብረተሰብ","ወገን"	መጀመሪያ	ቅድሚያ	ቀድሞ	ድሮ
ህዝብ	"ማህበረሰብ", "ህብረተሰብ", "ወገን", "ሀገር"	ልክ	ትክክል	ይደንቃል	ይገርማል
ድብቅ	"ስውር" , "የማይታይ" "ሚስጥር" "ህቡእ"	ወሬ	አሉባልታ	ሰላም	ፀጥታ
ሀሩር	"ቃጠሎ" "በርሃ" "ሙቀት"	ሀሰት	ውሸት	ሹመት	ስልጣን

1.5. List of Special character

[a-zA-Z 0-9 ።?!.¡:፣#-+@*-=~)(é`;'፡_/፣…፤,\"¿፩፪፫፬፭፮፯፰፱፲፳፴፵፶፲፳፴፵፶፷፸፹፺] >>$%!^&*()><":';.,]

Appendix 2

Deep learning CNN emotion classification using DL4J

```
11:00:22.583 [main] INFO org.deeplearning4j.nn.graph.ComputationGraph - Starting
ComputationGraph with WorkspaceModes set to [training: ENABLED; inference: ENABLED],
cacheMode set to [NONE]
11:00:22.685 [main] INFO org.nd4j.linalg.factory.Nd4jBackend - Loaded [CpuBackend]
backend
11:00:25.084 [main] INFO org.nd4j.nativeblas.NativeOpsHolder - Number of threads used
for linear algebra: 2
11:00:25.087 [main] WARN org.nd4j.linalg.cpu.nativecpu.CpuNDArrayFactory -
********************************** CPU Feature Check Warning ***********************
11:00:25.087 [main] WARN org.nd4j.linalg.cpu.nativecpu.CpuNDArrayFactory - Warning:
Initializing ND4J with Generic x86 binary on a CPU with AVX/AVX2 support
11:00:25.087 [main] WARN org.nd4j.linalg.cpu.nativecpu.CpuNDArrayFactory - Using ND4J
with AVX/AVX2 will improve performance. See deeplearning4j.org/cpu for more details
11:00:25.087 [main] WARN org.nd4j.linalg.cpu.nativecpu.CpuNDArrayFactory - Or set
environment variable ND4J_IGNORE_AVX=true to suppress this warning
11:00:25.087 [main] WARN org.nd4j.linalg.cpu.nativecpu.CpuNDArrayFactory -
************************************************************************************
11:00:25.312 [main] INFO org.nd4j.nativeblas.Nd4jBlas - Number of threads used for
OpenMP BLAS: 2
11:00:25.359 [main] INFO org.nd4j.linalg.api.ops.executioner.DefaultOpExecutioner -
Backend used: [CPU]; OS: [Windows 10]
11:00:25.359 [main] INFO org.nd4j.linalg.api.ops.executioner.DefaultOpExecutioner -
Cores: [4]; Memory: [2.0GB];
11:00:25.359 [main] INFO org.nd4j.linalg.api.ops.executioner.DefaultOpExecutioner -
Blas vendor: [OPENBLAS]
Number of parameters by layer:
      Cnn1   60100
      Cnn2   90100
      Cnn3   120100
      Cnn4   150100
      Cnn5   180100
      Cnn6   210100
      globalPool   0
      out    2404
Loading word vectors and creating DataSetIterators
org.deeplearning4j.models.embeddings.loader.WordVectorSerializer - Trying DL4j
format...
org.deeplearning4j.models.embeddings.loader.WordVectorSerializer - Trying
CSVReader...
Starting CNN training and evaluation:
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration 0
is 1.370726466178894
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration 100
is 1.3137003183364868
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration 200
is 0.8227765560150146
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration 300
is 0.9878020882606506
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration 400
is 1.005570411682129
```

```
 org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
500 is 0.9626182913780212
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration 600
is 1.0076035261154175
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration 700
is 1.2129766941070557
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration 800
is 1.251576542854309
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration 900
is 0.8099284768104553
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
1000 is 1.2393940687179565
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
1100 is 1.2156049013137817
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
1200 is 1.1151947975158691
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
1300 is 0.8742799162864685
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
1400 is 1.177909255027771
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
1500 is 0.7419812679290771
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
1600 is 0.8141526579856873
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
1700 is 0.5976892113685608
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
1800 is 1.0373454093933105
org.deeplearning4j.optimize.listeners.ScoreIterationListener - Score at iteration
1900 is 0.9873589873313904
org.nd4j.linalg.dataset.AsyncMultiDataSetIterator - Manually destroying AMDSI
workspace
Epoch 21 complete. Starting evaluation:
11:06:09.934 [main] DEBUG org.nd4j.linalg.dataset.AsyncMultiDataSetIterator -
Manually destroying AMDSI workspace

========================Evaluation Metrics========================
 # of classes:    4
 Accuracy:        0.7101
 Precision:       0.7106
 Recall:          0.7105
 F1 Score:        0.7096
Precision, recall & F1: macro-averaged (equally weighted avg. of 4 classes)

=========================Confusion Matrix=========================
    0    1    2    3
---------------------
  803  272   64  858 | 0 = Anger
  221 1162  127  447 | 1 = Disgust
  128  251 1288  277 | 2 = Happiness
  186  303  109 1372 | 3 = Sadness

Confusion matrix format: Actual (rowClass) predicted as (columnClass) N times
==================================================================
```

Example of users comment : ወሬ ሠለችን ተግባር ዜሮ
Example of users comment : እንደዚህ አድርገናል እንደዚህ አድርገናል የሚሉት ማድረገኮ ግዴታቼው ነው ሲጂመር ለውጡን ማን አምጥቶት ነው አያፍሩም እንዴ አድርገናል ሲሉ
Example of users comment : ቀዳዳህን ሣትሰፋ የሰዉ ቀዳዳ ኣትይ
Example of users comment : አንተን ብሎ ጠቅላይሚንስተር ደሞ አንተ ከርሰምነህ
Example of users comment : ይገርማል ውሸት አመላላሽ
Example of users comment : ፋና ውሸታም ሰለ ትግራይ ምት ለምን አትናገርም ውሸት አመላላሽ
Example of users comment : ቆሻሻ ምንያህል ከፈሏት እንኳን ያሰቃዩሽ በዚህ አይነት መጀመሪያም ሲያስሩሽ ልክ ነበሩ በስህተት ነው የተፈታሽ ሁለት ምላስ
Example of users comment : ገደል ግቡ ለባጠቃላይ ጊዜ አለው
Example of users comment : ቆሻሻ ዘራፊ ሁላ በንፁህን ዴም ቁማር እንዴተጫወታችሁ ፈጣሪ በባለሥልጣን ቤተሠብ ላይ ላአና ያውርድባችሁ
Example of users comment : የእበት ስብስብ
Example of users comment : ገደል ግቡ
Example of users comment : ገደል ግቡ በማርያም ታማችሁ አታሳምሙን
Example of users comment : ህዝብ እያለቀ ለማን ነው የምታለሙት ምድረ ነብሰ ገዳይም መረዳዳት ድንቄም መዋደድ
Example of users comment : ቀደዳ ሠለችን ተግባር ዜሮ
Example of users comment : የምንፈልገው ሰላም ሰማችን ያዉርድልን
Example of users comment : እግዝአብሄር በዘመንህን ይባርከው ልብ ሚነካ መሰለህ መልክት
Example of users comment : ይደመሰሳል ድል ሰፊው ህብረተሰብ
Example of users comment : ኢትዮጵያ ሀገሬ አንቱታሽ ከፍ ከፍ ይበል
Example of users comment : የሀገራችን ውበት በአብይ ተገልጦልናል አንቱታና ምስጋና ይገባዋል
Example of users comment : ግበዝ ሀገር ወዳድ አንቱታና ሞገስ ይገባሻል
Example of users comment : እኔኮ የግድቡን ሙሌት እንደ ሰበር ዜና ይነገረናል ብዬ ስጠብቅ ነበረ
Example of users comment : ምርጥ ኢትዮጵያዊ እውነተኛ የሀይማኖት አባት አንቱታ ይስጥልን ሀገር ማዳን ከእርሶ ነው መማር ያለብም
Example of users comment : በጣም ደስ ይላል የአባቶች አምላክ ከክፉ ነገር ይጠብቅህ
Example of users comment : አልሀምዱሊላህ አሁን በደንብ አመንኩ ከክፉ ነገር አላህ ሀገራችንን ይጠብቅልን
Example of users comment : ጎችን የማፈናቀል የመንግስት ግዴታ ነዉ በሉት
Example of users comment : የመተከል አማራወች በደላቸው ምንድን
Example of users comment : ጥሩ ሽልማቱ ገና ብዙ እጠብቃለን ገና ብዙ ይሰራል ብየ አምናለሁ
Example of users comment : ማሻ አላህ መጨረሻውን ያሳምረው
Example of users comment : ባንተ ኢትዮጵያ እንደምትበለፅግ አምናለሁ
Example of users comment : ሚኒሊክ ብዙ አመት ሀላ እንዲ ትከበራለህ ጀግና
Example of users comment : ድል መከላከያ ሰራዊታችን
Example of users comment : ውጭ ጉዳይ ሚኒስትር ሁሉንም ፈተናዎች ተጋፍጣችሁ ታላቅ ድል አብቅታችሁናል
Example of users comment : መልካም እድል ዋልያዎቹ
Example of users comment : ዋልያዎቹ መልስ ማሸነፍ ነዉ
Total number of comment is = 34
1, Anger
12, Disgust
18, Happiness
3, Sadness

Appendix 3
Word Embedding Word2vec Skipgraph Using DL4J

```
org.deeplearning4j.examples.advanced.modelling.textclassification.customcorpusword2ve
c.PrepareWordVector - Load & Vectorize DLEDMAT Sentences....
org.deeplearning4j.examples.advanced.modelling.textclassification.customcorpusword2ve
c.PrepareWordVector - Building DLEDMAT Word2Vec model....
08:51:50.629 [main] INFO org.nd4j.linalg.factory.Nd4jBackend - Loaded [CpuBackend]
backend
08:51:52.976 [main] INFO org.nd4j.nativeblas.NativeOpsHolder - Number of threads used
for linear algebra: 2
08:51:52.978 [main] WARN org.nd4j.linalg.cpu.nativecpu.CpuNDArrayFactory -
********************************* CPU Feature Check Warn ************************
 org.nd4j.linalg.cpu.nativecpu.CpuNDArrayFactory - Warning: Initializing ND4J with
Generic x86 binary on a CPU with AVX/AVX2 support
 org.nd4j.linalg.cpu.nativecpu.CpuNDArrayFactory - Using ND4J with AVX/AVX2 will
improve performance. See deeplearning4j.org/cpu for more details
 org.nd4j.linalg.cpu.nativecpu.CpuNDArrayFactory - Or set environment variable
ND4J_IGNORE_AVX=true to suppress this warning
 org.nd4j.linalg.cpu.nativecpu.CpuNDArrayFactory -
************************************************************************************
org.nd4j.nativeblas.Nd4jBlas - Number of threads used for OpenMP BLAS: 2
org.nd4j.linalg.api.ops.executioner.DefaultOpExecutioner - Backend used: [CPU]; OS:
[Windows 10]
org.nd4j.linalg.api.ops.executioner.DefaultOpExecutioner - Cores: [4]; Memory:
[2.0GB];
org.nd4j.linalg.api.ops.executioner.DefaultOpExecutioner - Blas vendor: [OPENBLAS]
org.deeplearning4j.examples.advanced.modelling.textclassification.customcorpusword2ve
c.PrepareWordVector - Fitting Word2Vec model for DLEDMAT....
org.deeplearning4j.models.sequencevectors.SequenceVectors - Starting vocabulary
building...
org.deeplearning4j.models.word2vec.wordstore.VocabConstructor - Target vocab size
before building: [0]
org.deeplearning4j.models.word2vec.wordstore.VocabConstructor - Trying source
iterator: [0]
org.deeplearning4j.models.word2vec.wordstore.VocabConstructor - Target vocab size
before building: [0]
org.deeplearning4j.models.word2vec.wordstore.VocabConstructor - Sequences checked:
[100000]; Current vocabulary size: [89587]; Sequences/sec: 8360.50; Words/sec:
70088.37;
org.deeplearning4j.models.word2vec.wordstore.VocabConstructor - Waiting till all
processes stop...
org.deeplearning4j.models.word2vec.wordstore.VocabConstructor - Vocab size before
truncation: [111980], NumWords: [1191393], sequences parsed: [143714], counter:
[1183721]
org.deeplearning4j.models.word2vec.wordstore.VocabConstructor - Scavenger: Words
before: 111980; Words after: 32002;
org.deeplearning4j.models.word2vec.wordstore.VocabConstructor - Vocab size after
truncation: [32002], NumWords: [1026752], sequences parsed: [143714], counter:
[1183721]
org.deeplearning4j.models.word2vec.wordstore.VocabConstructor - Sequences checked:
[143714], Current vocabulary size: [32002]; Sequences/sec: [7068.71];
```

```
org.deeplearning4j.models.embeddings.loader.WordVectorSerializer - Projected memory
use for model: [24.42 MB]
org.deeplearning4j.models.embeddings.inmemory.InMemoryLookupTable - Initializing
syn1...
08:52:13.793 [main] INFO org.deeplearning4j.models.sequencevectors.SequenceVectors -
Building learning algorithms:
08:52:13.794 [main] INFO org.deeplearning4j.models.sequencevectors.SequenceVectors -
building ElementsLearningAlgorithm: [SkipGram]
08:52:13.802 [main] INFO org.deeplearning4j.models.sequencevectors.SequenceVectors -
Starting learning process...
org.deeplearning4j.models.sequencevectors.SequenceVectors - Epoch: [25]; Words
vectorized so far: [123042121]; Lines vectorized so far: [100000]; Seq/sec:
[9563.89]; Words/sec: [11767609.12]; learningRate: [1.0E-4]
09:23:42.584 [main] INFO org.deeplearning4j.models.sequencevectors.SequenceVectors -
Epoch [25] finished; Elements processed so far: [127385000]; Sequences processed:
[702405]
09:23:42.584 [main] INFO org.deeplearning4j.models.sequencevectors.SequenceVectors -
org.deeplearning4j.examples.advanced.modelling.textclassification.customcorpusword2ve
c.PrepareWordVector - Writing DLEDMAT word vectors to text file....
org.deeplearning4j.models.embeddings.loader.WordVectorSerializer - Saving header:
32002 100 143714 org.deeplearning4j.models.embeddings.loader.WordVectorSerializer -
Wrote 32002 with size of 100
org.deeplearning4j.examples.advanced.modelling.textclassification.customcorpusword2ve
c.PrepareWordVector ***** Writing word vectors to text file finished **********
10 Words closest to 'ህዝብ': [ህብረተሰብ, ማህበረሰብ, መቸም, እያለቀ, ህዝቡ, ቀር, ባዶ, ብላቹ,
                            አጀንዳ, ሾርት]
10 Words closest to 'ኦሮሚያ': [ክልል, ዞን, ኦሮምያ, ቤንሻንጉል, አማራ, ቤኒሻንጉል, ተወላጆች,
                            አመራሮች, ኦሮሞ, ጉምዝ]
10 Words closest to 'ድል': [ማሸነፍ, ብርታት, መከላከያችን, ኢትዮጵያ, ትኖራለች, ጀግኖቻችን,
                           ሰራዊታችን, ዘላለም, ዛሬም, ብሎናል]
10 Words closest to 'ጀግና': [ብርቱ, ግበዝ, ጀግኒት, ምርጥ, ቆራጥ, ሁለም, ይገባታል, እንዳንች,
                            ርታልን, ታሪክ]
10 Words closest to 'እግዚአብሔር': [አምላክ, ፈጣሪ, ይባርክ, አሜን, ይስጥህ, ይስጣቸው, ይስጣችሁ,
                            ይመስገን, ጌታ, ምህረቱን]

ህዝብ - ማህበረሰብ = 0.955482542514801
ኦሮሚያ - ወለጋ = 0.6557158827781677
ድል - ሰራዊት = 0.7970857858657837
ጀግና - ብርቱ = 0.9932829737663269
እግዚአብሔር - ፈጣሪ = 0.8914170861244202
```

Declaration

I, the undersigned, declare that this thesis is my original work and has not been presented for a degree in any other university, and that all source of materials used for the thesis have been duly acknowledged.

Declared by:

Name: Eyob Tesfu Merine

Signature: _____

Date: _____

Confirmed by advisor:

Name: Ayalew. B

Signature: _____

Date: _____