DEBRE BERHAN UNIVERSITY

SCHOOL OF COMPUTING

DEPARTMENT OF INFORMATION SYSTEMS

SENTIMENT MINING AND ASPECT BASED SUMMARIZATION OF OPINIONATED AFAAN OROMOO NEWS TEXT

BY

WEGDERES TARIKU

NOVEMBER 2017

DEBRE BERHAN, ETHIOPIA

DEBRE BERHAN UNIVERSITY

COLLEGE OF COMPUTING

DEPARTMENT OF INFORMATION SYSTEMS

SENTIMENT MINING AND ASPECT BASED SUMMARIZATION OF OPINIONATED AFAAN OROMOO NEWS TEXT

A THESIS SUBMITTED TO THE COLLEGE OF GRADUATE STUDIES OF DEBRE BIRHAN UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN INFORMATION SYSTEMS.

BY

WEGDERES TARIKU

NOVEMBER 2017

DEBRE BERHAN ETHIOPIA

DEBRE BERHAN UNIVERSITY

COLLEGE OF COMPUTING

DEPARTMENT OF INFORMATION SYSTEMS

SENTIMENT MINING AND ASPECT BASED SUMMARIZATION OF OPINIONATED AFAAN OROMOO NEWS TEXT

BY:

WEGDERES TARIKU

NAME AND SIGNATURE OF MEMBERS OF THE EXAMINING BOARD

TITLE	NAME	SIGNATURE	DATE
1. ADVISOR	DR. MILLION MESHESHA	_____	_____
2. CHAIRPERSON	_____	_____	_____
3. EXTERNAL EXAMINER	_____	_____	_____
4. INTERNAL EXAMINER	_____	_____	_____

DEDICATION

This thesis is dedicated to my family and to all those who gave me all kinds of support_next to God

ACKNOWLEDGEMENT

First thing first Praise is to God, the Almighty for having guided me at every stage of my life. Many thanks and sincere gratefulness goes to my advisor Dr. Million Meshesha, without his valuable guide, support, and continuous follow-up; this research would never have been possible. My special thanks go to Damena Desta, Ashabir Hunegnaw, Fitsum Andersom,Yilma Goshime, Abdulkadir kedir, Muhidin Kedir, Abdi Shimalis and my class mate information system students who helped me throughout my works. In addition, I want to thank, staff of Oromia Radio and Television Organization for providing me the raw data from which I could construct a corpus. I feel happy to thank Mizan-teppi University for sponsoring me; I would like to extend my thanks to the Debre Berhan university college of computing science for their support in research fund and opening master's program in information systems. I also extend my thanks to my parents for their support every time and everywhere. Finally, I would like to thank all who helped me directly or indirectly to finish this study.

Contents

List of tables

List of figures

List of Equation

List of Acronyms

API	Application Program Interface
IR	Information Retrieval
ME	Maximum Entropy
NABSS	Non-aspect based sentiment summarization.
NB	Naive Bayes
NLP	Natural language processing
OM	Opinion mining
ORTO	Oromia Radio and Television Organization.
POS	Part of speech tagging.
RDBMS	Relational database managing system
SVM	Support vector machine
TF-IDF	Term Frequency and inverted document frequency
UGC	User generated Content
WWW	World Wide Web

ABSTRACT

The inventions of web 2.0 have paved the way for the rapid growth of user-generated content. Now we are in the world in which masses of user-generated content (i.e. word of mouth) are easily unrestricted online on the web in different domain. This massive number of user-generated content can be used as source to mine the sentiment orientation of the opinion holder as positive or negative. This massive number of user-generated content can be difficult manually to collect, understand, summarize, analyses for decision making. For this reason, studying on this particular topic of opinion mining has attracted an attention of many researchers to tackle this obstacle. Opinion mining can be done in one of the three different levels: sentence, document and feature/aspect level. In the middle of these three levels, feature level opinion mining is the detail and complex one but has a better benefit to meet customers and organizations need.

While there are many feature level opinion mining models that have been undertaken for international and local languages, to the level of the researcher knowledge opinion mining and summarization at aspect level in Afaan Oromo language is never done until now. Therefore, this study investigates and aims to develop sentiment mining and aspect based opinion summarization of service review in Afaan Oromo language for Oromia Radio and Television Organization (ORTO). A total of 400 reviews collected from Oromia Radio and Television Organization in news domain is used for the study.

Developed feature level opinion mining model mainly consists of five components: ORTO document review, pre-processing, aspect extraction, polarity detection, and aspect based sentiment summarization and draw bar chart to represent aspect based opinion polarity graphically.

Experimental result shows that for positive class precision of 90% and recall of 87%, whereas for negative class precision 87% and 89.7% recall is achieved. The major challenge observed in this study is polarity of opinions. This is because user mostly give their opinion in context based or/and indirect manner. They may use positive words to provide their negative feeling or vice versa. Thus it needs further research to make the system consider context based or semantic opinion mining.

Keywords*: Sentiments; Opinionated Afaan Oromo News Texts; Feature level opinion mining; Opinion summarization; Lexical database; Oromia Radio and Television Organization Text News*

CHAPTER ONE

INTRODUCTION

1.1 Background

The inventions of web 2.0 have concreted the way for the rapid growth of user-generated content [**1**]. Now we are in the world in which masses of user-generated content (i.e. word of mouth) are easily unrestricted online on the web in different domain. This user-generated content (UGC) can be posted online by an opinion holder from different review sites, blogs, discussion forums, social media review, and different organizations and non-organizations site, etc. This massive number of user-generated content can be used as source to mine the sentiment orientation of the opinion holder as positive, negative or neutral. For this reason, studying on this particular topic of sentiment mining have attracted an attention of many researchers [**2**].

Currently people often uses opinion, feeling ,attitudes or emotion that can be expressed by users of the web online against some particular entity for making decision in order to get "what is the opinion of the opinion holder on some particular entity (i.e. object)?" [**1**].

In addition to this an opinion that can be supplied by users of the web online can help organizations in improving their product and services. For example, most of us ask our friends for which department we have to join, which movie we have to watch, which hotels can give us better service, what kind of phone we have to buy and which book we have to read. For this reason, an answer expected from friends toward these questions is called an opinion. The process of extracting the attitude of the holder of an opinion toward something is called opinion mining or sentiment mining [**1**]. Therefore, sentiment miner tackles the problem of waiting an answer from friends [**3**].

The object in which an opinion is expressed can be product, service, organization, events, issues etc. Different researchers' call sentiment mining as opinion mining, sentiment analysis, opinion extraction, emotion analysis, subjectivity analysis review mining etc. Although the name is different they all are under one field of study that is called opinion mining with small task dissimilarity. This opinion mining is hot research topic in the current trend in NLP although it is broadly considered in field of data mining [**4**]. These topics have inspired many researchers also

due to availability of its diverse application areas.

Sentiment mining can be classified in to three levels of granularity [**2**]: sentence level, document level and aspect level. Sentence level sentiment mining task is undertaken to classify reviews sentence in to two concepts subjective and objective. Subjective is one that can be classified in to either positive or negative sentiment orientation. This can answer the question is this review sentence is positive or negative? While the objective ones are facts therefore, it can be categorized as neutral (i.e. no opinion) [**5**]. For Example, today i heard good sport news, shows subjective concept since it contains opinionated words 'good' which is positive. Therefore, the sentence has positive sentiment orientation.

Document level sentiment mining task is used to know the overall sentiment given over particular entity by an opinion holder. This can answer the question, is this review document positive or negative? In document level there is no way to know opinion orientation on each and every individual aspect. It is simply providing the general sentiment orientation for review document [**1**]. Consider the following review document; the sport news of ORTO is amazing, the political news of ORTO is awful, the business news of ORTO is bad and the science and technology news of ORTO is great.

In document level there is no way to know in which features the reviews have positive and negative opinion but the only thing it does is giving aggregate value negative since the negative review is greater than the positive one in this particular example. Thus, in document even though review has an amazing network features it doesn't care about the component of the object but object itself. In both sentence and document, level it is not possible to know which features exactly is loved and hated by an opinion holder [**1**]. In aspect level, this thing is possible since the aspect level opinion mining deals with the components or attributes of an object (e.g. Product, Movie etc.) at finer grained level of analysis [**2**]. See an example followed; I heard science and technology news yesterday from ORTO and it is amazing news. From this example, we understand that an opinion is expressed on ORTO component science and technology news that it is amazing rather than ORTO itself. Therefore, in aspect level one is able to know an opinion of an object somehow at detail level. This can answer the question is this review sentence feature is positive or negative. Thus, this research focus is on aspect summarization of opinionated Afaan Oromo news text reviews.

Aspect level opinion mining is done using different methods and techniques. In this study, we used rule-based approach to detect the aspect of the opinions and to predict the polarity of the sentiments and then summarize the polarity of opinions based on each aspects using bar chat.

Using this method first, we take a opinions text as input then we detect the aspect of the opinions then we predict the polarity of the opinions by cross checking the comment with root words available in the lexicon database and checking negation before and after the opinions root words to detect the polarity of opinions given from Oromia Radio and Television organizations in news domain then we count the total polarity of the opinions which are available on each aspect then summarization of aspect based opinion is done and it is displayed using bar chart which is easily understandable by user or organization who use the opinions.

The major challenge of Afaan Oromo language in sentiment analysis and summarization is lack of available resource such as preparation of lexicon words database, part of speech tagging used to identify the noun, noun phrase, adjective, verb and adverb, to detect aspect of the opinions and polarity of opinions. From these challenges we developed lexicon of the opinions words database.

1.2 Motivation

After, the last few years the application of opinion mining in different business activities with natural language processing and information retrieval has let tremendous inspiration for various researchers. Therefore, no exception for us. Opinion mining is directly connected with our daily life since it is utilized in making improvement in human computer-interaction, business intelligence based on user interest understanding. This time we are in the world in which online business activities are rapidly increasing in a swift manner. This business activity which is performed online has unrestricted huge amount of information which has useful incent for decision making but not managed correctly. Automating this information is useful for online users even though the operations were manually performed.

Therefore, automatic sentiment mining and summarization system can help any users of the web to get access throughout the reviews by reducing time lost in reading an information and analyzing it. This information could be unstructured therefore, we need to structure it in concise summary which is easily understandable than the manual task. Non-aspect based opinion

summarization considers sentiment classification scores for presenting general summary. This could be performed by counting the predicted polarity value of a given input review.

An aspect identification step is not included in these types of non-aspect based opinion summarization. Furthermore, the result generated cannot show finer grained level of sentiment. This kind of summary do not help customer to know about particular product at finer grained level. In addition, opinion mining and summarization can be undertaken in different languages but none of them can work in Afaan Oromo languages Therefore, this could be also one of the main inspiration of this research work.

1.3 Statement of the problem

In this busy world, no one is ready to give time to read, understand, summarize and organize features of opinions given on products or services for faster decision-making [**1**]. Providing concise and quick summary from thousands of review in hundreds of documents is overwhelming. It is very complex and challenging tasks especially when it has sub-features and when there are comments and replies showing relationship between individual reviews. Therefore, presenting overall short summary of particular review documents at finer grained level with visualization is very important.

In this study, we consider service review for Oromia Radio and Television Organization (ORTO) because it is one of the most popular organization broadcasting news, drama, music, sport programs and other educational services. On these, aspect services users provide their opinion through different mechanisms like phone, web, SMS, manually or face to face by going to the office of Oromia Radio and Television Organization.

Currently Oromia Radio and Television Organization read, understand and summarize those opinions opinionated on a service one by one. This way of reading, analyzing, understanding and summarizing thousands and millions of aspect based opinion mining and summarization is a difficult task. So that there is a need for automatically mining and summarization of sentiments at aspect level given on product, or service reviews. For this reason, many researches on sentiment mining and summarization have been done and are being undertaken for English [**6**], French, Turkey [**7**] and Arabic [**8**].

To the extent of the knowledge of the researcher on local language opinion mining done for Amharic language by different authors such as Tulu Tilahun [**9**] Proposes opinion mining from Amharic blog, Ephrem Alamirew [**10**] has conduct automatic acquisition and annotation of Opinionated Amharic text, Selama Gebremeskel [**11**] has studied sentiment-mining model for opinionated Amharic texts but opinion mining and summarization at aspect level in Afaan Oromo language that show the product or service features that are liked or disliked by the users or the customer's never attempted until now, while the amount of Afaan Oromo text on the web is increasing from time to time. Therefore, this study investigates and aims to develop sentiment mining and aspect based opinion summarization of service review for Oromia Radio and Television Organization (ORTO).

To this end, this study attempts to address and answer the followings research questions.

- ✓ How Data (corpus) can be prepared for service review in Afaan Oromo language?
- ✓ How to develop suitable prototype for polarity classification?
- ✓ How to summarize aspect level opinion in service review domain?
- ✓ To what extent the prototype performs opinion prediction and aspect level opinion summarization?

1.4. Objective of the study

1.4.1. General Objective

The general objective of this study is to design and develop Sentiment Mining and Aspect Based Opinion Summarization Model of service Review in Afaan Oromo for Oromoia Radio and Television Organization (ORTO) in news domain.

1.4.2. Specific Objectives

The specific objectives of this research works are discussed as below:

- Extensive study of available literature on the opinion mining and summarization approaches and techniques.
- Data cleaning and corpus preparation for product or service reviews.
- Understanding and analyzing the complex nature of customer reviews in product domain.
- Constructing a model for opinion mining and summarization at aspect level in service review domain.
- Evaluating the performance of the prototype.

1.5. Scope and Limitation of the study

The main aim of this study is to develop sentiment mining and aspect based opinion summarization of service review in Afaan Oromo language. This research is conducted based on data acquired from Oromia radio and television organization through the website of ORTO. Specifically, we focus on news domain, news has different aspects like sport news, political news, business news, metrology news, science and technology news, health and education news, art and culture news, international news. Opinion holders like individuals or organizations can express their opinion on a given service in different mechanisms like text, voice, picture, videos from these mechanisms we process text based user generated content.

The approaches we used for opinion mining include feature extraction, classification and summarization using lexicon based and rule based approach. This research gives an attention for regular opinions-specifically for direct opinion. Regular opinion can be expressed in two ways direct and in direct. Direct expression is for example *"safu drama which is presented in ORTO is amazing!"* and indirect expression is for example *"by closing the camera I boost the battery performance of my iPhone 5 phone".* This research can handle mixed sentence. Mixed sentences are sentence that can be positive in one sense and negative in another.

This research also covers feature detection, polarity identification or sentiment classification (i.e. positive and negative polarity) at aspect level and summary of aspect with its corresponding polarity. Every text formation available online can be classified into facts and opinions, Facts are an objective information about the given object or neutral opinion means like no opinions which is not used for decision-making. Opinion is about reflecting one's perception or views on particular entities, product or service. Due to time limitations, this study does not cover all root words.

1.6. Methodology of the study

Research methodology is the way of systematically solving research problem. This research was conducted to undertake Sentiment Mining and Aspect Based Opinion Summarization for constructing a model of Afaan Oromo text concerning Oromia Radio and Television Organization (ORTO)**.** Towards achieving the main objective, the following step-by-step procedures are followed.

1.6.1. Research design

This study follows empirical research. Empirical research is researching using empirical evidence. It is a way of gaining knowledge by means of direct and indirect observation and experience. Such research is often conducted to answer a specific question or to test a hypothesis. Empirical evidence (the record of one's direct observation or experience) can be analyzed qualitatively or quantitatively. Through quantifying the evidence of making sense of it in qualitative form, researcher can answer research questions that should be clearly defined and answerable with the evidence collected [**12**].

Empirical research can be classified in to two, which is experiment and non-experiment. In experiment first we conducted experimental setup like we decide about how many experiments needs to be conducted, with what parameters, weights, algorithms, datasets, in this research we used java programming language to build the prototype, we conducts one experiment which takes secondary data as input from Oromia Radio and Television Organization at news text domain to process automatically feature level opinion mining and summarization using bar chart.

Non-experiment is a fieldwork, which about to gather data from ORTO at news domain from the web, user acceptance testing and domain experiment testing; to what extent the developed prototype is, achieve a good result in aspect level opinion mining and summarization using bar chart for visualization

Several steps are used in automatic feature level sentiment mining and summarization systems. Generally, we can classify them in to five phases of development task. The first phase is the data cleaning and corpus preparations since data that we have collected were raw data. The second phase understands the nature of review in context with an Afaan Oromo language structures. The third phase of development is constructing appropriate model. The fourth step is designing an algorithm and implementing with programming language. The final fifth step is evaluation.

1.6.2. Data source and data collection methods

Most of the datasets used for conducting experiment was collected from Oromia Radio and Television Organization website (http://www.orto.gov.et), at news domain. Opinion mining and aspect-based summarization techniques are evaluated on 400 review dataset collected from Oromia Radio and Television Organization at news domain through web. The reason behind choosing these domains is the availability of user generated content in Afaan Oromo language

electronically such as in web, blogs and online forums in others domain. As a result, it is relatively more easy and manageable to collect Afaan Oromo news reviews than any other domains. Secondary data sources are used for this research, which are data that already exists in ORTO at news domain.

1.6.3. Design Approach

To model sentiment mining and feature level opinion summarization from Afaan Oromo user generated content, we used rule based approach to detect the polarity of opinion and summarization. The rule-based approach relies on handcrafted rules. The handcrafted rules can be made based the nature of Afaan Oromo language, to detect the aspect of the opinions, then we build lexicon words database of Afaan Oromo language that express opinion words like good, bad, amazing, awful, great etc. To detect the polarity of the opinions and count the number of positive and negative opinions available in each aspect and displays the summary of opinions in bar chart, that can be easy for visualization of opinions for customers and ORTO.

1.6.4. Implementation tools

In order to achieve our objective, we use java-programming language for implementing the algorithm to creating user interface to facilitate aspect based opinion summarization and MySQL for creating database for data repository such as keyword or lexicon of the opinion to detect the polarity of the opinion.

We use java-programming language because of its object oriented**,** in java everything is an Object. Unlike several other programming languages like C and C++ when Java is compiled, it is not compiled into specific machine, quite into platform independent byte code. This byte code is distributed over the web and interpreted by virtual Machine (JVM) on whichever platform it is being run. Java is intended easy to learn. With Java's secure feature, it allows to develop virus-free, tamper-free systems. Java creates an exertion to eliminate error prone situations by emphasizing mainly on compile time error checking and runtime checking. Java byte code is translated on the fly to native machine instructions and is not stored anywhere. The development process is more rapid and analytical since the linking is an incremental and lightweight process. With the use of Just-In-Time compilers Java enables high performance. Java is considered more dynamic than C or C++ since it is designed to familiarize to an evolving environment. Java

programs can carry extensive amount of run-time information that can be used to verify and resolve accesses to objects on run-time.

We use MySQL because of many good reasons; it is a fast, easy-to-use relational database managing system (RDBMS). MySQL is free under an open-source license. Therefore, we have nothing to pay to use it. It handles a large subset of the most expensive and powerful database packages and Uses a standard form of the well-known SQL data language. MySQL works on many operating systems and with many languages including PHP, PERL, C, C++, JAVA etc. MySQL supports huge databases, up to 50 million rows or more in a table. MySQL is customizable.

1.6.5. Evaluation techniques

This research focuses on designing and developing the system using different approaches, techniques and tools then evaluation of the designed system is the important tasks. Therefore, precision, recall, F-Measure, and user acceptance testing evaluate the performance of the system. Precision is the fraction of retrieved instances that are relevant/correctly classified, while recall is the fraction of relevant/correct instances that are retrieved [**13**].F-measure is one of a performance measure that combines Recall and precision into a single measure of performance, this is just by taking into account the product of Precision, Recall divided by their average. User acceptance testing is evaluating usability of the developed prototype like is system user-centered design and easy to use in feature/aspect level opinion mining, and aspect based summarization.

1.7. Significance of the study

These days we are in the world in which online business activities are rapidly increasing. This online business can generate huge amount of review. In this situation, opinion mining has its importance for people to make effective decision. Therefore, developing automatic sentiment mining and aspect based opinion summarization model can be used with the following importance.

The study has a huge value for the Oromia Radio and Television Organizations by taking users review on the service they gave for the organization to make efficient and effective decision. Also, have great significance for ORTO for service quality improvement by understand what the Afaan Oromo language speaker's customer like and dislike in the service. By taking user

generated content ORTO can be competitive radio and television broadcast in Ethiopia and in Africa. Aspect based opinions summary can be saving efforts and time by helping the organization to find which features/aspect will be improved in the service that customer dislike it ,if they developed it to a working system.

Customer or user those use ORTO service can get better news service, while the news is the key to get update information for our daily life to cope up with this dynamic and complex world.

It also helps other researcher those working in the similar area they use as reference and it is used as input for those who are working on recommender system. Mining and summarization of Afaan Oromo opinions from these vast amounts of reviews from the web becomes vital, and has many function for user, organization and for other researcher who are working on Afaan Oromo language they use as reference.

1.8. Operational definition of terms

Sentiment/opinion: -An opinion is an individual's or organization private state; it exemplifies the personality's assessments, evaluations, emotions, beliefs, judgments and ideas concerning a particular item/subject/topic/service.

Feature/aspect of opinion: - A feature is an attribute/part of an Oromia Radio and Television Organization News, which is being commented in some way in the review by users. Feature of opinion is simply the target of the opinion.

Opinion mining: -the process of finding opinion, attitude, emotion, appraisals of opinion holders about a given service, product or topic from huge amount of created information is known as opinion mining.

Opinion polarity/orientation: -The polarity or orientation of a sentiment can be statedeither positive, negative or neutral, based on encouraging and discouraging opinion given by user.

Opinion summarization: -One opinion from a single person is usually not sufficient for giving effective decision on a given issue or service. This directs that some form of summary of opinions is necessary; the method of summary are based on aspects given on a service or product is called aspect-based opinion summary or feature-based opinion summary. Positive and negative scores of aspects are separately aggregated; hence, we get an aggregate positive score and negative score of aspect. Aspect based opinion summary algorithm take the polarity of positive

and negative and aspects as input and then count the total polarity as positive and negative of opinions and aspects then summarize the polarity of opinions based on aspects by using bar char for visualization .

Lexicon database: -is a database, which have opinion words that is used for polarity of sentiment classification in the review texts

Bar chart: - is a chart or graph that provides gathered opinions with rectangular bars with lengths relative to the values that they denote. The bars can be designed vertically or horizontally. A bar chart or graph is a chart that uses horizontal or vertical bars to show comparisons among opinion aspects or categories. One axis of the chart displays the specific aspect or categories of opinion being compared, and the other axis represents a distinct value of opinions here under the aspect, which can be separate color of aspect for visualization purpose.

1.9. Thesis Organization

Brief discussion on the organization of document structure is given as follows. Chapter one introduce statement of the problems, objective, scope, limitation, and significance of the study. In chapter two an overview of sentiment mining and summarization at different levels is given. Detail explanation is given for some methods and techniques used by other researchers in the same or related area and reviews of related research work on opinion mining and aspect-based summarization are discussed. Moreover, brief explanation of sentiment mining and aspect summarization works at different levels and on different languages and domains as well as different approaches and techniques used are presented. In chapter three introductions to Afaan Oromo language and its writing, Afaan Oromo Morphology and Afaan Oromo Alphabets, is discussed detail. In chapter four the general architecture of the proposed model for Afaan Oromo automatic opinion mining and summarization at aspect level is developed, techniques and algorithm selection process is discussed in detail. Also general and domain based corpus or data preparation process can be done used for experiment validation. In chapter five the experimental result of the proposed models in general with different algorithms and evaluation of proposed model is discussed. In chapter six conclusions of the overall research and future work/recommendation is discussed.

CHAPTER TWO

LITERATURE REVIEW

2.1. Overview

An opinion is an individual's private state; it denotes the personality's assessments, evaluations, emotions, beliefs, judgments and ideas concerning a particular item/subject/topic. Opinions of others can have great impact on and offer management for governments, social communities, individuals and organizations in the process of making effective decision-making [**14**].Every text based information exist online can be classified into two major classes, facts and opinions [**1**]. Facts are an objective information about the given object or event while opinions are subjective information about a given object or entities. Subjective is about reflecting one's perception or views on particular entities. Most of the work done previously were focused on facts such as web search, information retrieval etc. In contrary to this little attention were given for opinions compared to facts since, there was a few user generated content before the coming of web 2.0. Previously, people used to make decision based on the opinion of friends or families.

For an organization to decide about a given product or service conducting survey or target group is common. But, after coming of World Wide Web especially web 2.0 the world has transformed. Today, nobody is tired of asking friends for opinions toward particular product or services to make decision since huge volume of information about a given entity are available online. In addition to this there is no organizations are participating in conducting survey or focus group in order to know what others say about themselves. This is because of huge volume

of review information that are produced online on different social networking sites, discussion forums, blogs etc. Since the size of information is very huge, it's very tough for ones to read, understand and produce useful summary of every review information provided on different aspect of the entity.

Therefore, to tackle this problem an automated opinion mining and aspect based summarization system is required [**8**]. This chapter discusses area of research on the general concept of opinion mining, sentiment classification, aspect based and non-aspect based opinion summarization with their techniques that were utilized.

2.2. Opinion Mining

Opinion mining is an emerging area of study under the umbrella of text mining and natural language processing. After the coming of the web, users can share their experience on particular product or services by reflecting reviews. Reviewers toward objects such as individuals, organizations, product, services events etc. can post the reviews online through blogs, movie, and products or service [**4**]. The review posted online by web users cannot only limited toward the entity it can also be given for component or attribute set of an entity. These reviews are increasing in an alarming rate these days. Hence, large volume of user generated content available which needs to be mined to analyses. Therefore, the process of finding opinion, attitude, emotion, appraisals of opinion holders about a given topic from huge amount of created information is known as opinion mining [**15**].

Opinion mining is also known as sentiment mining, sentiment analysis, emotion detection and review mining. Different researches are undertaken under this field of study such as subjectivity detection, opinion summarization, comparative opinion mining, opinion spam detection and opinion question answering, entity and feature detection [**16**].

2.3. Component of Opinion Mining

As suggested by Vinodhini and Chandrasekaran [**17**].there are five major components of opinion mining.

Opinion orientation: opinion orientation defines the opinion given on particular object as positive, negative or neutral. For example: *According to Tola, ORTO sport news he heard yesterday at 4:20 AM is amazing.* This statement states that the ORTO sport news is amazing. Thus, the word "amazing" in this statement shows sentiment orientation, which is positive.

Opinion Holder: is writer or an author of the opinion that expresses feeling on some particular object or provide product or service review. From the above statement, Tola is an opinion holders or writer of the opinion.

Opinion Object: is an object on which writers express an opinion, opinion object may be product or service. On the above example, Sport news is a service on which an opinion is expressed.

Opinion time: an opinion holder will hold an opinion over some object feature at some point in time. On the above example, opinion can be given at 4:20 AM.

Other customers: are other individuals or organization that uses these opinions, which are expressed by opinion holders to make decision. From the above example, ORTO is an organization that cares about opinions.

2.4. Levels of opinion mining

Sentiment mining can be classified in to three levels of granularity [**1**]: sentence level, document level and aspect level.

2.4.1 Sentence level opinion mining

Sentence level sentiment mining task is undertaken to classify reviews sentence in to two concepts subjective and objective [**1**]. Subjective is one that can be classified in to either positive or negative sentiment orientation. This can answer the question is this review sentence is positive or negative? While the objective ones are facts therefore, it can be categorized as neutral (i.e. no opinion)**.**For example, *"I heard ORTO political news yesterday and it is awful!"* This sentence shows subjective concept since it contains opinionated words 'awful 'which is negative. Therefore, the sentence has negative sentiment orientation. Consider another example *"I hear sport news today and it's great!"* This sentence shows subjective concept since it contains opinionated words 'great 'which is positive**.** Therefore, this sentence has positive sentiment orientation.

2.4.2 Document level opinion mining

Document level sentiment mining task helps us to know the overall sentiment given over particular entity by an opinion holder. This can answer the question whether review document is positive or negative. In document level there is no way to know opinion orientation on each and every individual aspect. It is simply providing the general sentiment orientation for review document. Let us see an example provided by an author on one-product reviews. Suppose this is the review document, *("ORTO sport news is amazing", "ORTO political news is awful", and "ORTO metrology news is great!").* In document level it is challenging to know in which features the reviews have positive and negative opinion but the only thing it does is giving aggregate value negative since the negative review is greater than the positive one in this

particular example. Thus, in document even though review has amazing sport news features it does not care about the component of the organization instead service itself.

2.4.3 Aspect/feature level opinion mining

In both sentence and document levels it is not possible to know which features exactly is liked and disliked by an opinion holder. In aspect level, however is possible since the aspect level opinion mining deals with the components or attributes of an object (e.g. Product, service, Movie and etc.) at finer grained level of analysis. See an example followed. *("I bought an IPhone three days and battery is still working!").*From this example, we understand that an opinion is expressed on IPhone component battery that it is good rather than IPhone itself. Therefore, in aspect level one is able to know an opinion of an object somehow at detail level. This can answer the question whether a given review sentence feature is positive or negative**.**

Opinion mining also consists of sub component such as sentiment classification, Opinion Question Answering, comparative mining, opinion trustworthiness and opinion summarization.

Opinion question answering: - Most of the state-of-the-art Question Answering (QA) systems aid the needs of answering truthful queries such as "When was Abebe Tola born?" and "Who won the Nobel Music Prize in 2017?" Though, in addition to truths, people would also like to know about others' sentiments, opinions, and feelings toward a particular topics, services, products, and events [**18**]. Opinion questions reveal responses about people's sentiments (*for example* "What will make someone successful in life?" and "What is the public opinion on human cloning?") which tend to distribute across diverse documents. Traditional question answering approaches for factual questions are not effective sufficient to retrieve answers for opinion questions so an opinion question answering system is essential that can be considered.

Comparative opinion mining: A comparative opinion define a relation of similarities or dissimilarities between two or more objects or service and/or a preference of the opinion holder based on some common features of the service or object [**19**]. For example, "*malt tastes better than Pepsi*" and "*malt tastes the best*" express two comparative sentiments. A comparative sentiment is frequently defined using the *comparative* or *superlative* form of an adjective or adverb, even though not permanently.

2.5. Opinion Summarization

Opinion summarization is one of the discipline comes under sentiment analysis. The goal of an opinion summarization is to make quick and concise summary of review which can be brought from large size of review corpus of entities [**2**]. This allows any users of the system to make better decision on a given product through getting concise summary in fast manner. Opinion summarization has sub task that need to be accomplished such as aspect detection, opinion detection and actual summary generation.

Most of existing works classify opinion summarization into two broad categories such as abstractive summarization and extractive summarization [**20**]. In abstractive summarization, understanding the concepts that are expressed by authors in a review document in well manner is needed. In these, understanding the reviews in short format is required. In order to detect crucial idea abstractive summarization employs many natural language processing techniques on review source. In extractive summary, concatenation of important sentence together is carried out first by choosing crucial sentence from reviews source.

As compared to aspect based opinion summarization, Non-Aspect Based Sentiment Summarization comprises various cluster such as advanced text summarization, basic sentiment summarization [**20**].Text summarization focus on identifying the most important points of a text and expresses them in a shorter document [**21**].Basic sentiment summarization is types of non-aspect based opinion summarization considers sentiment classification scores for presenting summary. This could be performed by counting the predicted polarity value of a given input review. An aspect identification step is not included in these types of non-aspect based opinion summarization. Furthermore, the result generated cannot show finer grained level of sentiment. Therefore, this kind of summary do not help customer to know about particular product at finer grained level. Both sentence and document levels are using non-aspect based opinion mining

Sentence level or document level opinion mining were not providing a notion on what the customer exactly likes or what is not [**4**]. Besides, having positive in polarity of the overall opinion of document does not mean that customer likes everything (all features or aspect) of a given object or entity. Similarly having negative opinion polarity for a given document does not also mean that the customers hate every aspect or feature of an entity, or object. In a common

sense opinion, holder can writes a review on a given object or its features. This review may have optimistic or pessimistic polarity. To get an opinion on product service at finer grained level opinion mining at aspect level is a must. For aspect-based summarization, there are three commonly used tasks, such as aspect detection, opinion detection and opinion summarization [**4**].

2.5.1 Aspect detection

This task will determine aspect on a given objects of reviews. For example, in a review statement that says, *"The script writer of the movie is amazing".* The opinion *"amazing"* is given on feature *"script writer".* Aspect extraction is under an umbrella of information extraction since information extraction is used for extracting structured information from unstructured sources in automatic manner [**4**]. Traditional information extraction techniques are frequently developed for official type for example in news, scientific papers, which have some problems to be functional well to sentiment mining applications. We target to extract fine-grained information from sentiment text or documents for example blogs, reviews, and forum, since; it has complex nature due to noisy and complex characteristics of reviews that can be exploited for extraction.

Thus, it is important to develop and design extraction techniques that are particular to an opinion mining. This circumstance can be used also for entity extraction. Since most of applications are focused on online reviews, ours are not exceptional. Basically, two common format of reviews are most commonly used online under this circumstance [**4**].

Format 1: pros, cons and detailed description: authors are required to provide some discussion on the advantages and disadvantages of a given topics or objects by separating them both.

Format 2: Free Format: here authors are expected to write freely no more separation of advantages and disadvantages of a given topics are required. Since format 1 consists of one aspect in a segment it's simple. Therefore, our attention is toward format 2 that can be freely expressed by an author's [**4**].

Aspect based opinion summary is based on opinion from various opinion holder since opinion from single opinion holder has practically less importance in most of the opinion mining

application [**4**].Consequently, we need some form of summary, which comprises both quantitative and qualitative summary.

To detect the aspect/feature of the opinions there are different methods is here such as NLP, and frequent item. In natural language processing (NLP) techniques/heuristic, nouns and noun phrases describe features and adjectives and adverbs express opinions

Frequent item set:-Frequent subsets of words as candidate features, Pruning by compactness/redundancy/mutual information.

While there is not part of speech tagger available for Afaan Oromo language we create a rule that detect the aspect of the opinions from a given opinions, first we feed the ORTO opinions reviews to the developed prototype then the system splitting the opinions document line by line by using punctuation marks, then find category/aspect of opinions by titles of opinions finally count the number of news categories which is available in the reviews

2.5.2 Sentiment detection

The next step is sentiment detection, to determine the sentiment polarity on each aspect [**2**]. This method depends on a sentiment word dictionary that holds a list of positive and negative words for example great, amazing, stylish, awful, best, bad, awesome etc. These were used to match terms in the opinionated text. Also, since other special words might also change the orientation, special linguistic rules are proposed. These rules consider negation**,** the presence of negation terms can change the polarity, and there is a need to handle it. For example, the review statement that says, "I don't like watching sport news on television "has negative polarity since the positive opinion terms are negated by term **"**don't". Therefore**,** negation should be handled in a review sentence very carefully. On the one hand, negation terms may come positive and in the other hand, it comes as positive. For example, the review statement that says, **"**Todays ORTO culture and tourism news is not bad**"** has positive polarity since the positive opinion terms are shifted by term "not". Opinion detection is the most commonly studied area under natural language processing field of study [**2**]

In fact, it is a hot research area in opinion mining we are going to discuss sentiment detection that, were carried out under sentiment summarization [**20**]. Under this circumstance sentiment detection is not only limited to detecting the sentiment orientation of review instead the corresponding aspect of the review. In here, different people could have different opinion on the same aspect. This happens since opinion reflected by different people on the same aspect can be

different. For example, *"The ORTO business news quality is amazing" and "The ORTO science and technology news quality is awful".* This nature differs sentiment detection under opinion summarization from others [**20**]. In this study Lexicon and/or rule-based method is used for Sentiment Prediction/detection. This method has been used in opinion summarization [**1**]. The method requires dictionary of words or positive and negative seed list to predict opinion. This seed list/dictionary of words can be matched with corpus to determine the sentiment orientation of a given review. In some cases, this lexicon can be utilized with pre-defined set of rules and POS taggers or parser [**22**]

2.5.3 Summary generation

This step is followed after sentiment prediction and aspect detection step has carried out. This summarization can help users to get brief and quick summary of the reviews supplied. For summary generation described review sentence are required. This review sentence can comprise an aspect on which an opinion is given and its corresponding opinion. In previous section, the task of classifying an opinion extracted into positive and negative granularity was accomplished. Here, clustering can be made for aspects that are synonyms. Then, after the positive and negative opinions with corresponding aspect are clustered the graphical and text summarization of an aspect on a given object with its corresponding opinion can be generated and displayed [**23**]. Summary with timelines needs to be given since opinions can be changed in accordance with time. Providing overall summary on the recent review is very crucial to have latest decision since reviews that can be provided by an online user before one day and after one day can be different [**20**]. There are various methods for summary generation such as clustering, sentence phrase selection and visualization. Clustering is group similar topics together, sentence Phrase Selection is used to select some part of original data and show, but cannot show all details and bad selection may miss information, Coreference resolution is focus on 'who' says what about 'whom', this method is used to recognize the opinion holder and target more clearly but sometimes opinion holder is not important for example in service or product review and relatively its complex computation and Visualization is presentation of reviews graphically.

In this study, we use visualization method for ORTO text news aspect based opinion summarization. Visualization helps to make an opinion presented easily readable and understandable format. This visualization does not only benefit customers who use an opinion to make decision but it can also benefit other researcher to show the result on the polarity of an

opinion generated. To visualize the polarity and aspect based summarization of the opinions we use bar chart, first we gave number of aspects and polarity of opinions on each aspect then the bar chat algorithm can draw the graph which is easily understood by organizations and user of the opinions.

2.6. Evaluation methods

Evaluation of the prototype system is made with the evaluation parameter that matches the number of user-generated reviews, which are categorized correctly and incorrectly. We use the metrics such as accuracy, precision P, recall R, F-measure and accuracy.

Precision: - measures the exactness of a classifier. Precision is the ratio of the number of reviews classified correctly to the total number of reviews in a given category. A high precision means less false positive, while a lower precision means more positives that are false.

$$P(precision) = \frac{TP}{TP+FP} \text{----------} \quad \textbf{Equation 2. 1 precision formula}$$

Where TP (True positive) represent the number of user-generated content, which are correctly classified and FP (False positive) represent the number of user-generated content, which are incorrectly classified.

Recall: -measures the completeness or sensitivity of a classifier. It is the ratio of TP and the whole reviews belonging to the category. A high recall means less false negative, while lower recall means more negatives that are false

$$R(recall) = \frac{TP}{TP+FN} \text{----------} \quad \textbf{Equation 2. 2 Recall formula}$$

Where, FN resent the number of reviews, which are missed by the classifier, i.e. neither Classified correctly nor incorrectly or unclassified category.

There is trade-off between precision and recall. Greater precision decreases recall and greater recall leads to decreased precision.

F-measure is one of a performance measure that combines Recall and precision into a single measure of performance, this is just by taking into account the product of Precision and Recall divided by their average. Therefore, F-measure is expressed by formula as follows:

$$F\ measure = \frac{2 * precision * recall}{precision + recall} \text{ - - - - - - - - - - - - - -}$$ Equation 2. 3 F measure

2.7. Challenges in opinion mining

The ways of expressing an opinion is seen as one of the challenges that are left to be addressed in opinion mining since opinions that are expressed by the users of the web online varies from individuals to individuals. This difference is due to difference in the way every individuals thinks and lives are different. For instance, an opinion that is forwarded by an expert is different from an ordinary user this is due to variation in skills or knowhow of individuals on particular objects (entities) [**5**]. Another challenge in opinion mining that remains to be solved is domain deference. Domain deference has complex nature in its inconsistency; corpus used is different from domain to domain.

The review aspect detection can register lower performance while traversing through different domain. Detection of spam or fake reviews can also be a challenge in opinion mining. The review can be authentic or spam there is a need of detection before processing. This could be possible through identification of redundant reviews, outliers etc. [**5**]. Mixed sentence can also be another challenge in opinion mining since an opinion term (for example, long) which is positive in one sense can be negative in another. Therefore, it needs notions to detect such kinds of terms. Natural language processing overhead is another challenge since it contains ambiguity, co-references and implicitness.

Another challenge is parsing the sentence in order to know the subject and object the adjectives or a verb refers to. Most of the research work has been given consideration for English and Chinese languages. The consideration should also be given to other languages. Another challenge is opinion changes gradually over time. It shows an opinion of people over time, people might be convinced on particular product because an organization might have made change over their product over time. Selection of opinion based text from other kinds of textual information like facts are very challenging task. Identification of an opinion holder/author is also one of the challenges in an opinion mining recently. There is also an open challenge in opinion mining, this could be implicit way of aspect detection, extracting an aspect resides in different corpora [**1**].

2.8. Related Works

There are many research works that were done internationally and locally on sentiment mining, aspect based summarization in different languages like English, Arabic, Turkish, French, Germany and Amharic using different techniques and approaches. Most approach to sentiment analysis follows either machine learning approach or lexicon based approach. Machine learning approach can be classified into two such as Supervised (labeled data) and Unsupervised (no labeled data) learning. The Supervised approach can be rule based classification, decision tree, linear classification and Probabilistic classifier. Lexicon based approach can be classified into two such as dictionary based and corpus based approach (semantic and statistical) [**4**].In the following sections we introduce some of these researches which relate to our research.

International works done on sentiment mining

Shweta and Jain [**24**] conduct research on Feature based sentiment mining and opinion Orientation, which extracts the aspect and opinions from sentences and decides whether the particular sentences are positive, negative or neutral for each aspect. Negation is also controlled by the system. To govern the semantic polarity of the sentences a dictionary-constructed technique of the unsupervised approach is adopted. To determine the sentiment words and their synonyms and antonyms WordNet is used as a dictionary that has positive and negative words. The evaluation result shows that 'Aspect based opinion Orientation System' is achieved well in phone domain. The experiments have been implemented by using 50 sentences of phone analyses. The results show that the 'Feature based Sentiment Orientation System' achieves well with respect to the phone domain and shows the accuracy of 67%.

Velasquez and Felipe bravo-Marquez [**25**] proposes an extension of Bing Lius [**1**] aspect based opinion mining approach in order to apply it to the tourism domain. The extension Concerns with the fact that users refer differently to deferent kinds of products when writing reviews on the Web. Since Liu's approach is concentrated on physical product reviews, it might not be straight functional to the tourism domain, which presents aspect that is not measured by developed model. Through an exhaustive study of on-line tourism product reviews, they found these aspects and then model them in their extension, proposing the use of new and more difficult NLP-based rules for the activities of subjective and sentiment classification at the feature-level. Their work also involved the development of a general architecture for an aspect-based opinion-mining tool,

which they used to build a prototype and they analyze appraisal from Trip Advisor in the perspective of the tourism industry in Los Lagos, a Chilean administrative region also known as the Lake District. Results they prove that their extension is able to attain better than Liu's model in the tourism domain, improving both Accuracy and Recall for the tasks of subjective and sentiment polarity classification. Mostly, the approach is very effective in predicting polarity of sentiments, performing an F-measure of 92% for the task. However, on average, the algorithms were only proficient of extracting 35% of the explicit aspect expressions, using a non-extended approach for this task. Finally, results also showed the effectiveness of their design when applied to solving the industry's particular issues in the Lake District, since almost 80% of the users that used the tool considered that it adds respected information to their business.

Esra Akba [**7**]conduct research on aspect based opinion mining on Turkish tweets. He explores the problem of sentiment mining by extracting feature/aspects of entities/topics on assembly of short texts. They emphasis on Turkish tweets that hold informal short messages. Utmost of the existing resources such as lexicons and labeled corpus in the literature of sentiment mining are for the English language. Their approach would enhance the sentiment examines to other languages where such rich sources do not exist. After a set of preprocessing steps, they extract the aspects/feature of the product(s) from the available data and cluster or aggregate the tweets centered on the extracted aspects. In addition to their manually built Turkish judgment word list, an automated generation of the words with their opinion strengths is suggested using a word choice algorithm. Then, they recommend a new exemplification of the text according to opinion strength of the words, which they discuss to as opinion based text representation. The Aspect vectors of the text are created according to this novel representation. They adapt machine learning methods to generate classifiers based on the multi-variant scale aspect vectors to determine mixture of positive and negative sentiment and to check their performance on Turkish tweets.

Saensuk, Songram and Chomphuwiset [**26**]explore Feature-Based Opinion Mining on Smart-Phone Reviews. In smart-phone market, the overall review may not be a practical recommendation for customers to choose their phones. It is essential to summarize opinions of users based on each features of smart-phones. This will be an advantage for customers who want to buy smart-phones and smart-phone companies use this information to improve features of their smart-phone. In this study, authors proposed a method for mining opinions on smart-phone

reviews written in Thai. The method summarizes positive and negative polarity of each feature of smart-phones. In this research, first, smart-phone reviews are collected from smart-phone pages on Facebook using Facebook Graph API. Second, the review dataset is cleaned and then perform word segmentation using a Thai segmentation technique. Then finding similarity of words based on their feature is performed. Finally, each feature is used to decide the polarity of words as either negative or positive. From the experimental result, it was shown that the proposed method gives 70.17% of accuracy.

Liu, Zhixin and Wang [**27**]Extract Product Features in Chinese Web for Opinion Mining by analyzing the characteristics of Chinese product reviews on the Internet, a novel method based on feature clustering algorithm is compared with the aspect-level sentiment mining problems. Particularly, 1) features considered in their study include not only the explicit features but also the implicit features. 2) Opinion words are divided into two categories, vague opinions and clear opinions, to deal with the task. Aspect clustering depends on three features: the corresponding sentiment words, the similarities of the aspect in text and the structures of the features in comment. Moreover, the context information is used to enhance the clustering in the procedure. Experimental evaluation shows the outperformance of feature-based.

Another study done by Alasmar [**8**] presents a method using ontology that work at aspect level classification to categorize Arabic opinion on a given product or service by identifying the essential features from the review based on level of these features on the ontology tree and to produce an opinion summary for each aspect in the entire corpus by identifying the sentient of its sub-feature terms in the ontology. To evaluate their work, they use public datasets, which are hotels and books datasets. They use accuracy, recall, precision, and f-measure metrics to evaluate the performance and compare the results with other supervised or unsupervised techniques. In addition, subjective assessment is used in their method to show the effectiveness of aspect and opinion extraction method and summarization. They show that their method improves the performance compared with other work that can be done on sentiment mining classification techniques, attaining 78.83% f-measure in hotel domain and 79.18% in book domain. Furthermore, subjective evaluation shows the effectiveness of their method by earning an average f-measure of 84.62% in hotel domain and 86.31% in book domain.

Local works done on sentiment mining

Tullu Tilahun [**9**] proposes feature level opinion mining model for Amharic language in the domain by employing manually crafted rules and lexicon. The proposed model consists of five major components that can extract features, determine opinion words regarding identified features with their semantic orientation, aggregate multiple opinions and generate structured summary. They conduct two experiments that have been for features extraction and sentiment words determination by using 484 reviews from three different domains. The first experiment they did showed that an average precision of 95.2% and recall of 26.1% were attained in the features extraction and an average precision of 78.1% and recall of 66.8% were achieved in the determination of opinion words. The precision of the second experiment in features extraction gets lower by 15.4% whereas the precision of opinion words determination gets higher by 1.9% and the recall of both features extraction and opinion words determination gets greater by 7.8% and 25.9% respectively when they compared to the first experiment. Diverse study stated that there are two types of features: explicit and implicit. They have devoted to the extraction of explicit features. Extracting implicit features is very important that will be considered in the future work.

Selama Gebremeskel [**11**] has studied document level sentiment mining for opinionated Amharic text in movies and newspaper domain using general and domain specific opinion terms. The system they designed is based on the proposed model discovers positive and negative opinion terms including contextual valence shifters i.e. as negations and assigns an initial polarity load to all detected sentiment terms in order to decide the polarity classification of the opinionated text. The lexical of Amharic opinion terms they construct are used to discover and assign initial polarity value to the opinion terms detected. A prototype system they developed is used to authenticate the proposed model and the algorithms designed. Evaluation on the prototype are done using movie and newspaper reviews where the result obtained with these test data is very much encouraging. They present experimental results of the three different experiments. The first experiment is done using a particular general purpose dictionary without considering the contextual valence shifter terms. The second experiment is achieved using two opinions lexica: the general purpose lexicon and the domain specific lexicon. Finally, the result of the third experiment conducted using the two lexica and considering the contextual valence shifter terms.

They use 254 movie reviews and 49 newspaper reviews for conducting all the experiments. Each review was classified by the system prototype according to the procedures and all the results were recorded as showed in table blow.

System	Class	Precision	Recall	F-measure
General purpose Amharic sentiment terms	**Positive**	0.929	0.823	0.867
	Negative	0.6	0.573	0.589
Basic + domain lexicon	**Positive**	0.937	0.943	0.939
	Negative	0.62	0.78	0.69
Both lexica + contextual valance shifter terms	**Positive**	0.943	0.949	0.945
	Negative	0.666	0.842	0.743

Table 2-1 Experimental results achieved by selama [11**]**

Feature level sentiment mining can be future research work direction, which is concerned with identification and extraction of commented features and determining the sentiments towards these features. For example, movie features such as director, actor, lighting … etc, can be identified and their corresponding opinions can be determined. This is a more detailed area of study in the sentiment mining research works.

Abreham Getachew [**28**] conducts research constructing opinion-mining model that classify Amharic opinionated text into positive or negative. Experimental results indicate conducted are using 616 Amharic opinionated texts gathered from Ethiopia Broadcasting Corporation, diretube.com and habesha.com sites. The Experiment they conduct indicates that Information Gain feature selection methods perform the best through all algorithms (Naïve Bayes, Decision Tree and Maximum Entropy). Based on their relative performance of classification, Navies bays with 90.9% accuracy performs Decision Tree with 83.1% and Maximum Entropy with 89.6%. However, they did not control negation because of the use of unigram as aspect for classification. They recommend for further research to consider bi-gram and tri-gram to come up with a better feature set for opinionated Amharic texts.

Ephrem Alamirew [**10**] further proposed a system that can automatically extract the comment and label it in to positive, negative and neutral. It also extracts the targets to whom the comment is given on, which improves the quality of opinion mining and enhance the opinion mining process by saving time and costs for annotating using experts. They made a prototype using a java programming language that can integrate with Facebook API for extracting comments and replies for a particular post by inserting post ID then annotate for further study. For evaluation of annotation task they use Cohen's K, they get 0.5 for positive, 0.58 for Negative, 0.4 for Neutral, and for target 0.7 the result they achieve is low. This is because of the nature of comments they use; most user give comment by using English letter for writing Amharic comment. This can be resolved using special normalization for translating this kind of comment. They get0.83 accuracy results of the automatic extraction of opinionated Amharic comment. They recommend for further research to consider subjective and objective classification to enhance the performance of the model created in their research.

A study conducted by Mohammed Tune [**29**] a graph-based opinion summarizing system whose vertices contain message objects or topic under discussion and its reply nodes that are labeled with opinion polarities. The opinions used to undertake this study were obtained from social media (Facebook) and annotated by experts. The proposed model extracts the summary of these opinions polarity from the corpus of opinion-oriented graph. Hence, it is possible to achieve enhanced decisions by knowledge derived from this graph. The experimental results provided information that cannot be provided by solely straightforward text mining or Computational linguistics. Because, the methods of computational linguistics cannot notice the semantic relationships between more than two sentence. The experimental results showed that with the help of such model they able to achieve promising results in decision making that can be easily adapted to any topic with similar objective. However, the graph-based opinion mining by itself does not automatically identify the orientation of a text written in more than one language. For this, developing multilingual system is compulsory for this identification

System	**Reviews**	**Class**	**Precision**	**Recall**	**F-Measure**
Without discussion	English	Positive	0.789	0.873	0.828
		Negative	0.88	0.673	0.762

chain + In discussion chain	Amharic	Positive	0.5	0.573	0.536
		Negative	0.65	0.68	0.67
	Afaan Oromo	Positive	0.636	0.72	0.675
		Negative	0.6	0.75	0.67

Table 2-2 Experimental results achieved by Mohammed Tune [29**]**

From the above review of related works, one can understand that all of the local studies are done for Amharic language product or service opinionated texts. However, there are a lot of opinionated texts in Afaan Oromo language; which needs to be analyzed, mined and summarized. Accordingly, this study is initiated to undertake aspect based summarization of opinionated Afaan Oromo news text.

CHAPTER THREE

AFAAN OROMO LANGUAGE AND ITS WRITING

The Oromo people establish the single largest ethnic group in Ethiopia, where the Oromia region contains a third of Ethiopia's land area and population. The Oromo language, also well known as Afaan Oromo, is spoken as a first language in Ethiopia by 87% of Oromia's 27 million people. In further regions of Ethiopia, 1.4 million people consider it their mother tongue, meaning native Afaan Oromo speakers overall about 25 million in Ethiopia [**30**]. About 500,000 more live in Kenya and Somalia. Many others as yet uncounted speak Afaan Oromo as a second language. It is the most spoken language in the Cushitic family, which also includes Somali, Sidamo, and Afar. Afaan Oromo language uses Latin based script called Qubee. Currently, it is an official language of Oromia sate (which is the largest Regional State among the current Federal states in Ethiopia. Being the official language, it has been used as medium of instruction for primary and secondary schools of the region. Moreover, Afaan Oromo language is offered as a subject in some universities with Bachelor and Master's Degree levels.

Even though Afaan Oromo is Latin based script language but its structure of writing and grammar like normative, case, genitive, dative, locative, and definiteness are different from other Latin language like English, and European languages such as France, German ,Turkish. The grammatical system of Afaan Oromo language is quite difficult and exhibits many structures common to other languages, i.e., it is modulated language that uses postpositions more than prepositions.

3.1. Afaan Oromo Writing System

Afaan Oromo language uses Latin based script called Qubee, has been adopted and become the official script of Afaan Oromo since 1991 [**30**]. The Afaan Oromo writing system is a modification to Latin writing system. Thus, the language shares a lot of features with English writing with some modification. The Afaan Oromo writing system of the language is known as "*Qubee Afaan Oromoo*" is straightforward which is designed based on the Latin script. Thus letters in English language is also in Oromo except the way it is written. According to [**31**], Afaan Oromo is a phonetic language, which means that it is spoken in the way it is written. Unlike English or other

Latin based languages there are no skipped or unpronounced sounds/alphabets in the language.

The Qubee Afaan Oromoo writing system has 33 letters that consists of all the 26 English letterswith an addition of 7 combined consonant letters which are known as "*Qubee Dachaa*".These include *ch, dh, sh, ny, ts, ph* and *zy*. All the vowels in English are also vowels in "*Qubee*".These are *a, e, o, u* and *"i"*. Vowels have two natures in the language and they can result indifferent meaning. The natures are short and long vowels. A vowel is said to be short if it is one. If it is two, which is the maximum, then it is called long vowel. Consider these words: *lafa* (ground), *laafaa* (soft). In a word where consonant is doubled the sounds are more emphasized. For example, *dammee* (sweety), *damee* (branch).

3.2. Afaan Oromo Punctuation marks

Punctuation is placed in text to make meaning clear and reading easier. Analysis of Afaan Oromo texts reveals that different punctuation marks follow the same punctuation pattern used in English and other languages that follow Latin Writing System [**32**]. Like English, the following are some of the most usually used punctuation marks in Afaan Oromo language [**32**].

- "***Tuqaa***", Full stop (.): Like English full stop is used at the end of a sentence and also in abbreviations.
- *"**Mallattoo Gaafii**"*, Question mark (?): is used in interrogative or at the end of a direct question.
- *"**Rajeffannoo"***, Exclamation mark (!): is used at the end of command and exclamatory sentences.
- *"**Qooduu**"*, Comma (,): is used to separate listing in a sentence or to separate the elements in a series.
- *"**Tuqlamee colon**"*, Colon (:): the function of the colon is to separate and introduce lists, clauses, and quotations, along with several conventional uses, and etc. Unlike English language apostrophe (***'***) is not punctuation mark in Afaan Oromo, rather it is part of words. For example, *har'a* (today), *re'ee* (goat) etc.

3.3. Afaan Oromo Morphology

Like in a number of other African and Ethiopian languages, Afaan Oromo has a very complex and rich morphology [60]. It has the basic features of agglutinative languages involving very extensive inflectional and derivational morphological processes. In agglutinative languages like Afaan Oromo, most of the grammatical information is conveyed through affixes, such as, prefixes, infix and suffixes attached to the root or stem of words. Although Afaan Oromo words

have some prefixes and infixes, suffixes are the predominant morphological features in the language.

Almost all Afaan Oromo nouns in a given text have person, number, gender and possession markers, which are concatenated and affixed to a stem or singular noun form. In addition, Afaan Oromo noun plural markers or forms can have several alternatives. For instance, in comparison to the English noun plural marker, *s (-es)*, there are more than ten major and very common plural markers in Afaan Oromo including: *-oota, -oolii, -wwan, -lee, an,een, -eeyyii, -oo,* etc.). As an example, the Afaan Oromo singular noun *mana* (house) can take the following different plural forms: *manoota (mana +oota), manneen (mana + een), manawwan (mana + wwan)*. The construction and usages of such alternative affixes and attachments are governed by the morphological and syntactic rules of the language [**33**]. Afaan Oromo nouns have also a number of different cases and gender suffixes depending on the grammatical level and classification system used to analyze them. Frequent gender markers in Afaan Oromo include *-eessa/-eettii, -a/-ttii* or *–aa/tuu.*Consider the following example.

Afaan Oromo	*Construction*	*Gender*	*English*
Sangoota	*Sangaa +oota*	*Male*	*Ox*
Jaarsolii	*Jaarsa +olii*	*Male*	*Elder*
Obboleessa	*Obol +essa*	*Male*	*Brother*
Beektuu	*Beek +tuu*	*Female*	*knowledgeable*

Table 3-1Afaan oromo morphology

Afaan Oromo Nouns

Most of Afaan Oromo nouns and adjectives are marked for masculine or feminine gender. Nouns have an essential masculine or feminine gender that cannot be determined by the form of the noun, with a few exceptions when biological gender is associated with a particular suffix, such as eessa for masculine and -eetti for feminine nouns, e.g., obboleessa ‘brother’ and obboleetti ‘sister. ‘Adjectives agree with the nouns they modify in gender. All nouns and adjectives are marked for number: singular and plural, e.g., for masculine nouns nama ‘man’ – namicha ‘the man’; for feminine nouns haroo ‘lake’ – harittii ‘the lake’. All nouns are marked for case. Nouns

can be used attributively to express modification. There is also a difference between proximal and distal demonstrative pronouns, e.g., kana ‘this’ and san ‘that’.

Afaan Oromo Adjectives

Adjectives are very important in Afaan Oromo because its structure is used in every day conversation. Oromo Adjectives are words that describe or modify another person or thing in the sentence [**34**]. Unlike English adjectives are usually placed after the noun in Afan Oromo. For instance, in ***Caalaan mana guddaa bite*** “Chala bought a big house” the adjective ***guddaa*** comes after the noun ***mana***. Moreover, in Afan Oromo sometimes it is difficult to differentiate adjective from noun. Consider the following example

Afaan Oromo	*English*
Dhugaa	*Truth ,reality ,right*
Dhugaa kootii	*I am right,*
Obboleettin hiriyaa dhugaati	*Sister is the friend for truth/sister is a true friend* *True served as adjective.*

Table 3-2 Afaan Oromo adjective example

Adjectives in Afaan Oromo can be classified based on their function like English such as those express color, size, shape, quality and taste.

Afaan Oromo Verbs

Afaan Oromo verbs consist of a stem plus suffixes representing person, gender, number, tense-aspect, mood, and voice. Verbs agree with their subjects in person and number. Verbs, with the exception of the verb ‘be’, agree with their subjects in gender, when the subject is a 3rd person singular pronoun ‘he’ or ‘she’. There are basically two tense/aspect divisions: complete (perfective/past) and incomplete (present or future progressive). Compound tenses are formed with auxiliary verbs. There are four moods: indicative, interrogative, imperative, and jussive. The latter is used to express commands, permission, and agreement. There are three voices: active, passive, and the so-called autobenefactive (semi-passive/middle).

Afaan Oromo Conjunctions

A word that can be used to join or connect two phrases, clauses and sentences is known as a conjunction [**35**]. Conjunctions can be divided into coordinating and subordinating conjunctions. Coordinating conjunctions are used to connect two independent clauses. Mostly these conjunctions are used when the speaker needs to lay emphasis on the two sentences equally. Some of these conjunctions in Afaan Oromo include: *garuu* „but", moo„ or", *kanaafuu*„ therefore", *haata'u malee*„however/so", *tu'ullee*„even though" etc. Consider the following example: *Hawiin siriitti hinqo'atu garuu qormaata dabartee jirti*. This means „Hawi is not studying hard but she passed the exam "*garuu*" in this sentence is coordinating conjunction. It is used to join the two sentences *Hawiin siriitti hinqo'atu'*and *qormaata dabartee jirti* which are independent. Subordinating conjunctions are those conjunctions that are used to join main clause withsubordinate clause. A subordinating conjunction is always followed by a clause [**36**]. Afaan Oromo subordinating conjunctions include *yoo* as if, *akka waan* as if, wayta*/yeenna* as when, *hamma* as until, *erga* as after", *dursa* as before" etc. The following example illustrates the above case: *Akka waan nabeeku fakkeesse*. This means "He acts as he knows me". *Akka waan*in this sentence is used as subordinating conjunction. It joins one subordinating clause that is *waan nabeeku* 'As he knows me" and *fakkeesse* 'He acts".

Afaan Oromo Prepositions

Prepositions in Afaan Oromo, links nouns, pronouns and phrases to other words in a sentence. The word or phrase that the preposition introduces is called the object of the preposition [**37**]. Examples

Afaan Oromo	*English*
Inni gara Jimmaa deeme	*He went to Jimma*
Hangan dhufutti na eegi	*Wait for me until I came back*
Namni akka harkaan waa hojjechuuf fayyadamu *Arbi maalitti fayyadamaa?*	*As people use hands to work something, what does elephant use?*

Table 3-3 Afaan oromo preposition example

Afaan Oromo Word order and boundaries

The unique word order in afaan Oromo sentences is Subject – Object – Verb. Modifiers, pronouns, articles, and case markers follow the nouns they modify [**38**]. Most Latin languages a word is separated from other words by white space character [**39**]. Afaan Oromo also uses white space to separate words from each other. Moreover, parenthesis, brackets, quotes, etc., are being used to show a word boundary. Furthermore, sentence boundaries punctuations are almost similar to English language i.e. a sentence may end with a period (.), a question mark (?), or an exclamation point (!) [**36**].

3.4. Challenge of Afaan Oromo in sentiment analyses and summarization

Understanding, analyzing and summarizing Afaan Oromo opinionated news text in automated manner have some challenges. The major challenge of Afaan oromo language in sentiment analysis and summarization is lack of available resource such as preparation lexicon database, part of speech tagging used to identify the noun, noun phrase, adjective, verb and adverb, used to detect aspect of the opinions and polarity of opinions. The other challenge is Afaan Oromo is morphologically very rich, so it's difficult to detect the entire stem or root of words

The other challenges are there is lack of a system that identify subjective and objective information or fact information, which is not used for news opinion mining for encouragement or discouragement on Oromia Radio and Organization news service.. The other one is users give their opinion in context based or/and indirect manner this is another challenging task to identify the polarity of opinions. In addition identifying the aspect of the opinion, determining the polarity of opinions and aspect based opinion summarizing is a challenging task.

CHAPTER FOUR

METHODS AND IMPLEMENTATION

This chapter describes the general system architecture, techniques used and algorithm selection of proposed sentiment mining and aspect based opinion summarization of service review in Afaan Oromo language that is discussed below in detail. The architecture describes the procedure or a flow of work how the developed model is solving problem of sentiment mining and aspect based summarization by drawing bar chart. Techniques are the tools or methods used to aspect detection, polarity detection and summarization of each polarity on aspect of user-generated content.

4.1. General system architecture

The process of sentiment mining and aspect based opinion summarization in Afaan Oromo language from Oromia Radio and Television Organization in news domain takes a review as input via web, and extracts the features that identifies a product aspect which are being commented by customer or user, determines the opinions polarity, and produces a summary or aggregates the results obtained from opinion polarity as an output and then represent by bar chart.

The general system architecture of automatic sentiment mining and aspect based summarization of opinionated news text in Afaan Oromo review is described below in figure 4.1. The system encloses different components based on the activities required to make the system efficient and effective. These components are Oromia Radio and Television Organization (ORTO) news review document, aspect extraction, pre-processing a given document, polarity detection, and aspects based sentiment summarization and finally draw bar char to represent the information for visualization

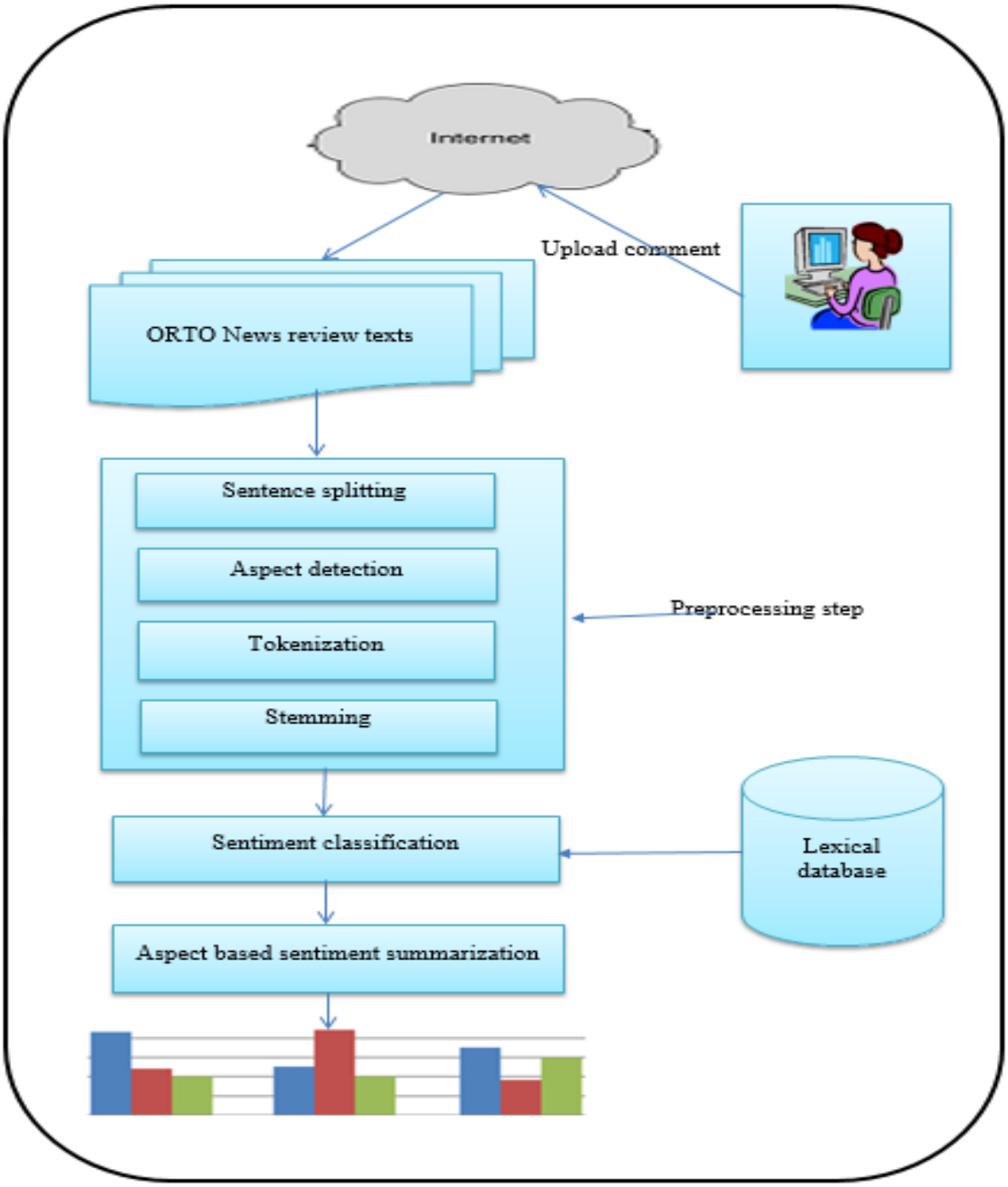

Figure 4.1Architecture of proposed system.

Given Oromia Radio and Television news review text as an input, the opinionated texts are preprocessed to prepare data set for training and testing. For text preprocessing, tasks such as sentence splitting, tokenization and stemming are performed. During text preprocessing ,first the review document is splited into sentence levels, from which words/tokens are identified in the

tokenization step, these list of words are stemmed to changing them in to base words. Then the rule-based algorithm are applied to build the model used to detect aspect, predict sentiment polarity by cross checking the tokens with the lexical database and count the polarity of opinions available under each aspects to summarize aspect-based sentiment. Finally draw the bar chart summary of the opinions by taking aspects of opinions and polarity of the opinions here under each step. The next duty is evaluation of the model and performance status of the model using different evaluation metrics such as precision, recall, F- Measure and user acceptance testing.

4.2. Pre-processing

In order to perform sentiment mining and summarization the text or document collected from Oromia Radio and Television Organization in news are pre-processed to acceptable representation before using it for constructing a model for opinion classification. Preprocessing task includes sentence splitter, tokenization, and stemming.

4.6.1 Sentence splitter

Sentence splitter is one of the pre-processing tasks that can be undertaken when we are working in sentiment mining and aspect summarization to extract the aspect of opinion and detect polarity of the sentiment in easy way by splitting the document of text up in to separate sentences. This was supported out by considering punctuation marks that are used at the end of sentence; like full stop (**.**) and exclamation marks (!). Sentence splitters, sometimes called sentence segmenters, divide text up into separate sentences per definite delimiters [**40**].

Sentence Split source code

//split the whole opinion based on the keyword "Title: " and save on "titles" array.

String[] titles = txtDisplay.getText().split("Title: ");

//Then split each index of "titles" array based on the keyword "\n " and save on "Opinion" array.

String[] opinion = titles[t].split("\n");

4.6.2 Feature or Aspect detection

Features or aspect detection is vital to know the main topic of a sentence to infer the emotion on such a topic rather than detecting sentiments of unimportant topics. We call aspect expressions for a product or service review in a sentence that are nouns and noun phrases explicit aspect expressions. For example, “sound” in “The sound of this phone is clear” is an explicit aspect expression. We call aspect expressions of the other types, implicit aspect expressions, as they often imply some aspects. For example, “large” is an implicit aspect expression in “This phone is too large”. After identifying aspect as noun and noun phrase then we find its frequency, most frequent aspect ones are chosen and the infrequent ones are unwanted. Since Afaan oromo part of speech tagger (POS) is not available to detect aspect of the opinion from the text review we use rule based methods. For feature or aspect extraction, the developed prototype takes review from ORTO news text as input then split the reviews line by line by using delimiters then find the category of news from a given opinion, if the titles of opinions is here on the review then take as the aspect of the opinions and iterate up to the end of review then count the aspect of the opinions available on the given reviews. Algorithm depicted in figure 4.2 is used for aspect detection.

Aspect Identification source code

```
//after sentence splited
//Then for all sentences saved on Opinion array identify the sentence saved at  index  0
//Then save the sentence on category array  as aspect or title.
for (int i = 0; i < opinion.length; i++) {
if(i==0){
category[t]=opinion[0];
            }
}
```

As we see in the above algorithm to detect the aspect of the opinions from the huge reviews using rule based algorithm, first the model accept the ORTO opinion reviews then by splitting the opinions sentence line by line by using punctuation marks, find category/aspect of opinions by titles of opinions then extract category of opinions and then count the number of news categories which is available in the reviews

4.6.3 Tokenization

Certain a character arrangement and a described document unit, tokenization is the task of breaking sentence into units, called tokens or words by discovering word boarders. Starting point of the word and end of the word is called word boundaries. Tokenization is also well-known as word segmentation.

4.6.4 Stemmer

Stemmer is a component that reduces morphological variants of words into base or root form. In morphologically complex languages like Afaan Oromo, a stemmer will lead to important improvements in opinion mining systems. Stemming is the process of eliminating affixes (prefixes and suffixes) fromwords. Stemming diminish variant word forms to a distinct "stem".For example in Afaan oromo, ‘afeeramuu’, ‘afeerraa’, ‘afeeramtani’, ‘afeeraman’, ‘afeeramaniiru’, all are rooted to‘afeer’.Stemming is done while building dictionary. For example deem stem is saved rather than saving deemi, deemte, deemuuf and etc. Only the root word and its polarity is saved in the dictionary and retrieved when needed. In this study we adopt Debala [**34**]stemmer.

4.3. Sentiment classification

This component identifies Orientation of an opinion on each aspect, i.e., polarity scores of opinion on aspect. Proposed method uses review corpus for assigning importance scores to opinion words, which is a dictionary or keyword of opinion words which contains positive and the negative words. Suppose if a word is found in the dictionary and if its corresponding value is positive, then this opinion term is positive. Similarly, if a term is found in the dictionary and if its corresponding value is negative, then this opinion term is negative. For negation, suppose if a word is found in the dictionary and if its corresponding value is positive and negated one times, then this opinion term is negative. Similarly, if a term is found in the dictionary and if its corresponding value is negative and negated one times, then this opinion term is positive. Polarity words are terms that can express opinions towards an object such as ‘gaarii’ (good) that expresses positive opinion, and‘gaarii miti’ (not good) or ‘gadhee’ (bad) that expresses negative opinion towards an object. Afaan Oromo keyword or lexicon database, which contains Afaan Oromo opinion words, the sentiment predictions help in summarizing the general sentiments polarity of the product’s aspect/feature.

Once the key word is saved in dictionary along with its polarity, the following line of code select the key word from the dictionary and save on “keyword” and its type or polarity on “type”.

To detect the polarity of opinions from huge number of opinions file based on aspect, first user accept the opinions review file then detect the polarity of opinions then initialize number of positive and negative polarity of opinions. Then Get keyword from lexical database to cross check with the words available in the reviews then check negation after and before the keyword then determine the polarity of opinions, if polarity is positive, count the positive polarity of opinions which here under each aspects and also if the polarity is negative, count the negative polarity which here under each aspects and finally summarize the total polarity available in the opinions based on aspects available by using bar chart for visualization.

4.4. Lexical database

Lexical database, which have root of opinion words like gaarii (good), bad (gadhe) etc that is used for polarity of sentiment classification in the review news texts

4.5. Aspect-Based Opinion Summary for visualization

Most opinion mining applications want to study opinions from a huge number of opinion holders. One opinion from a single person is usually not sufficient for giving effective decision on a given issue. This directs that some form of summary of opinions is necessary; the method of summary are based on aspects given of a service or product is called aspect-based opinion summary (or feature-based opinion summary).Positive and negative scores of aspects are separately aggregated; hence, we get an aggregate positive score and negative score of aspect. Aspect based opinion summary algorithm take the polarity of positive and negative and aspects as input and then count the total polarity as positive and negative of opinions and aspects then summarize the polarity of opinions based on aspects.

Theaggregated positive score and negative score of user-generated content has been used for generating summary using visualization tools, which is easy to understand users opinion in structured form. In this study, bar chart is drawn for generating summary of users news review about ORTO.

Once we design algorithm and procedure followed in sentiment and opinion summarization, in the next chapter experimentation and evaluation is undertaken.

4.6. Implementation

Every system is developed to undertake some functionality. These functionalities are evaluated to make sure that the systems are performing effectively. Effectiveness refers to the extent to which a system fulfills its objective. In the case of our prototype system, the exactness of extracting relevant features and the exactness of determining polarity of opinion words are evaluated. Implementation, testing environment, and evaluation metrics such as precision, recall and F-measure, experimental results and discussions are all subtopics that will be discussed in the following sections.

In order to achieve our objective, we used different environments and tools. Java 8.2 version programming language is used to develop the prototype with a graphical user interface (GUI) for the opinion mining and aspect based opinion summarization process. MySQL is used for creating database for data repository such as to store keyword of the opinion.

4.6.1 Building sentiment lexicon

The value of lexicon-based sentiment classification systems depends on the effectiveness of the sentiment lexicon. To construct the Afaan Oromo sentiment lexicon that is used to determine opinion polarity as positive or negative, there is no organized Afaan Oromo resource is available to use, because of this we must develop Afaan Oromo sentiment lexicon that have positive and negative to determine polarity of user generated content on given entity. To develop sentiment lexicon we use different sources like Amharic sentiment lexicon developed by Selama Gebre meskel [**11**]and English word net**,** which contains positive and negative opinion words by translating the word into Afaan Oromo language based on Afaan Oromo language rules**.**

Additionally, we use hard copy of Afaan oromo dictionary "Galmee Jechoota Afaan Oromo" made by Hinsene Mekuria to collect additional Afaan oromo sentiment terms. We used the dictionary to collect the opinion terms based on the following guidelines.

- A term that cans definite subjectivity (positive or negative) independent of any otherterm is carefully chosen into the sentiment lexicon.
- Terms that have a part of speech (POS) tagger value of Adjective are given priority.
- The most usually used contextual valence shifter terms are selected

Totally we collected around 1027 Afaan oromo opinion terms where 521of them are positive (+) terms and the rest of 506 are negative (-) terms. Finally, these Afaan oromo opinion terms are validated by a professional students from the Linguistics Department at Addis Ababa university.

4.6.2 Data Collection

Opinion mining and aspect-based summarization techniques are evaluated on 400 review dataset through web collected from Oromia radio and television organization news domain. The reason behind choosing these domains is the availability of user generated content in Afaan oromo language electronically such as in web, blogs and online forums in others domain. As a result, it is relatively more easy and manageable to collect Afaan Oromo news reviews than any other domains.

4.6.3 Oromia Radio and Television News Reviews

Oromia Radio and Television Organizations can give news service for the organization indifferent aspects such as sport news, political news, business news, and science and technology news, education and health news, international news and etc. On this news domain, service users have been giving their emotions or opinions.

4.6.4 Designed user opinion summarization

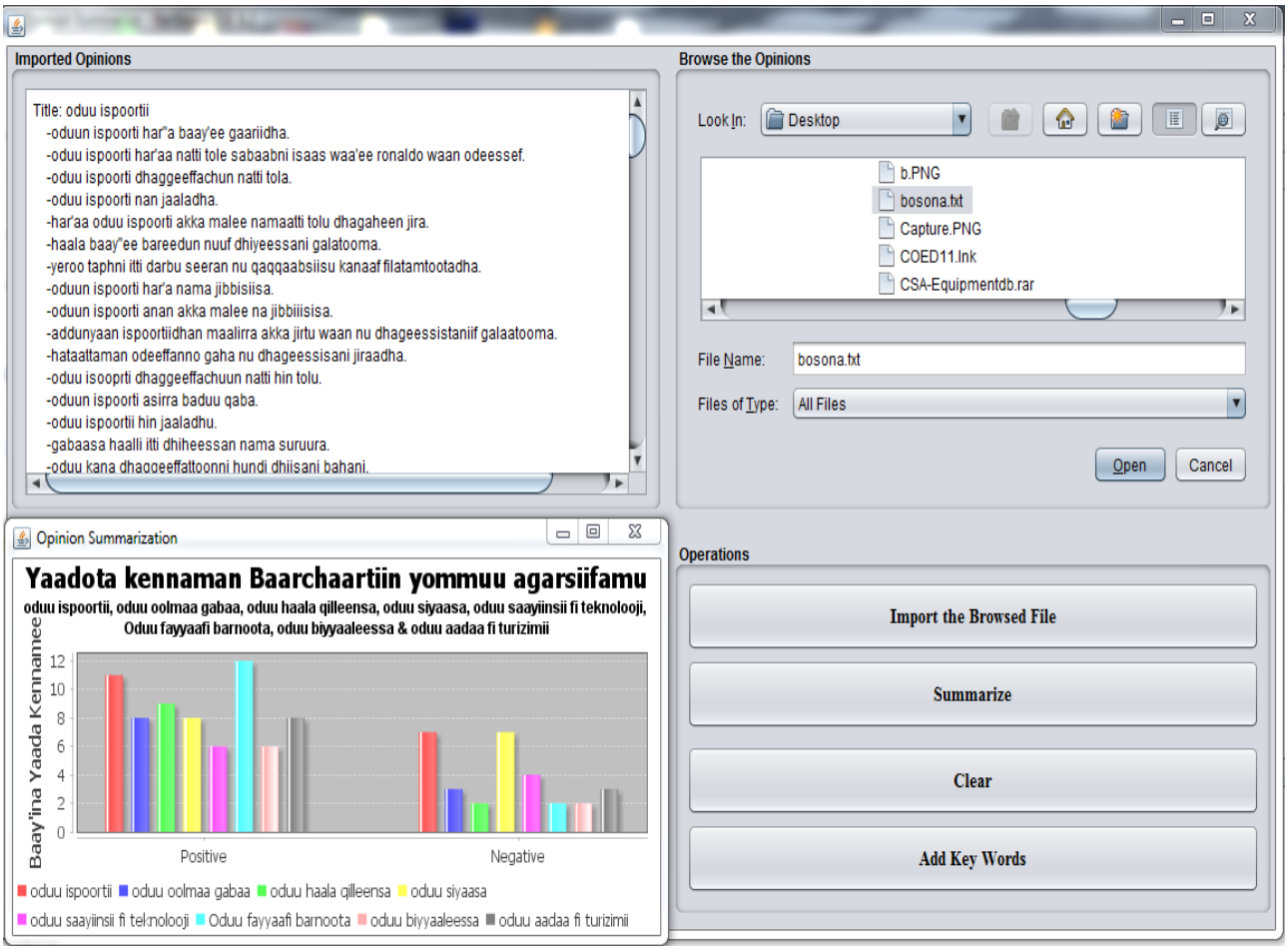

Figure 4-2: Feature level opinion mining and summarization with bar chart

As you see on the above figure, designed user opinion mining and summarization have the following functionality first user run the developed prototype and then the interface is displayed then users can start interacting with the system, user can feed the document in to the system using browser FileChooser options and then import the selected document into text display field by using import the browsed file button the press the summarize button then aspect summarization of bar chart based opinion is displayed that can be easy for visualization by user and organization to get concise summary of opinion for making decision.

Opinion mining and aspect-based summarization techniques are evaluated on 400 test data through web collected from Oromia radio and television organization at text news domain, each aspect have this much of data as described below table

Oromia Radio and Television Organization News Aspect/Feature	*Number of review Used*
Sport news/Oduu ispoortii	*70*
Business news/Oduu oolmaa gabaa	*56*
Metrology news/Oduu haala qilleensa	*48*
Political news/Oduu siyaasa	*18*
Science and technology news/Oduu saayiinsii fi technolooji	*60*
Health and Education news/Oduu fayyaa fi barnoota	*55*
Culture and tourism news/Oduu aadaa fi turizimii	*47*
International news/Oduu biyyaaleessa	*46*
Total	*400*

Table 4-1Number of data taken from Oromia Radio and Television Organization

As you see in the above table number of data taken from Oromia Radio and Television Organization in sport news aspect is larger than the other aspects.

CHAPTER FIVE

EVALUATION AND DISCUSSION

5.1 Evaluation Metrics

To measure the rate of opinion mining and aspect-based summarization of our system, we use the most common evaluation techniques, which select a small sample of words and compare the results of the system with a human judge. Evaluation of the prototype system is made with the evaluation parameter that matches the number of user-generated reviews, which are categorized correctly and incorrectly. We use the metrics such as accuracy, precision P, recall R, F-measure. In addition, user acceptance testing is performed to assess the usability of our proposed system.

5.2 System performance testing

System performance testing checks functionalities of developed prototype whether it meet level of accuracy required to achieve by researcher. In this study, we select three employees from Oromia Radio and Television Organization those who are working as customer opinion officer and opinion analyst, for purpose of evaluating or testing accuracy of developed prototype. Four hundred opinion reviews was given for each evaluators.

In testing the performance of the prototype, the result is classified into correct and incorrect classification of the polarity of the review cases by comparing the output of the system prototype with domain expert judgments of the review. The result is printed below in table 5.1.difference between human and system opinion classification.

Automatic opinion mining and aspect level opinion summarization	Class	Precision	Recall	F-measure
	Positive	0.90	0.871	0.885
	Negative	0.87	0.897	0.88

Table 5-1Experimental result for opinion mining and aspect based opinion Summarization

The result of the experiment shows the performance of the method. From the experiment, we observe performance of 90 % precision and 87.1 % recall in for positive class, for negative class

87% precision, 89.7% recall and is 88.3% system accuracy is achieved in the purpose of aspect based opinion summarization.

The other experiment was done for measuring the performance of the opinion mining and aspect based summarization process by comparing manpower/manually classifying opinions and aspect opinion summarization with developed system prototype. We choose three persons for manual classification and aspect based opinion summarization. We discuss it in the following table 5.2.

classifying andsummarizing aspect level opinion	Time	Man power	Review document
manpower classification and summarization	60 minutes	4	400 opinions
Automatic classification and summarization	2 minutes	1	400 opinions

Table 5-2Comparisons between manpower and automatic opinion mining and aspect based opinion summarization.

A comparison of the time consumption of the proposed system with manual opinion mining and aspect based opinion summarization task is also done to measure time consumption. The result shows automatic aspect based opinion summarization system is saving time and manpower costs. For example, within 2 minutes 400 opinionated review are classified by the proposed system; while it takes 60 minutes, we classify the polarity and aspect based opinion summarization using four person. This means that the proposed system saves more 96 % of the time wasted by individual to clarify manually.

5.3 User acceptance testing

User acceptance testing is the basic issue for the development of successful and effective prototype development system. To address the issue of user acceptance testing, the researcher uses visual interaction along with close-ended questionnaires. Visual interaction evaluation method allows the domain expert to make comments by interacting with the prototype. It is used to evaluate the performance of the prototype from the users' point of view. Similarly, the questionnaires are helped to assess and evaluate the acceptability and applicability of prototype

for sentiment mining and aspect based opinion summarization in service review domain also, we consider the attitude of the users about the prototype development system.

For the purpose of user acceptance testing process, ten (10) persons are selected; five from domain experts in Oromia radio and television organization staff and five are from different users who know Afaan Oromo language to evaluate the system. Before starting the evaluation process, the researcher gave explanation about the prototype system in detail for evaluators. This explanation helped the evaluators to avoid the variation of awareness among them about the system .After evaluators are interacting with the system by using test cases, which have similar parameters with the rules in the prototype; they give their feedbacks on the questionnaires. The questions depend on the user interface design aspects, which is basic for users interface satisfaction. These questions assessed whether the user interface of the system is easy to use, attractiveness, and the problem solving ability and contribution in opinion mining and aspect based opinion summarization.

All these seven closed ended questions answered as from Excellent, Very Good, Good, Fair, to Poor. Therefore, for the comfort of analyzing the relative performance of the system based on users evaluation, the researcher assigned numbers for each word as Excellent = 5, Very good = 4, Good = 3, Fair = 2 and Poor = 1. The system evaluators give the value for each closed ended questions. The following table indicates the results achieved.

No	Questions	Poor	Fair	Good	Very good	Excellent	Average
1	Is the prototype easy to use and interact with it?			2	5	3	4.1
2	Is the prototype system attractive?			1	5	4	4.3
3	Is the system efficient in time?				6	4	4.4
4	Is the system giving the right opinion polarity classification based on aspects?			1	7	2	4.1
5	What do you think the contribution of the system in the study area?			1	6	3	4.2

	Total Average	4.2

Table 5-3User acceptance testing

As it discussed based on above questions users gave their answer for the question. When we express by present on each value for open-ended questions, in the prototype easy to use and interact with it evaluation, 20% of the evaluators scored as good and 50% as very good and 30% as excellent. In the prototype system attractive, the evaluators gave 60% as very good and 40% as excellent. On the system efficiency, the evaluators gave 10% as good, 60% as very good and 30%as excellent. In the case of polarity classification based on aspects, evaluators gave their value 10% as good, 70% as very good and 20%as excellent. In the bar chart aspect based opinion summarization, the evaluators gave their value 50%as very good and 50%as excellent. In the contribution of the system in the study area criteria evaluators gave their attitude 10% as good, 60%as very good and 30%as excellent. Finally, the average performance of the prototype system according to the evaluation results filled by the domain experts is 4.2 out of 5 or around 84%, which is a very good achievement by user acceptance testing.

5.4 Discussions of results

The experimental results shows the effectiveness of the proposed model in detecting features of opinion and determining polarity of opinion words along identified features for the Oromia radio and television organizations in news domains. The result of the experiment shows the performance of the method. The experiment we could observe performance of 90 % precision and 87.1 % recall in for positive class, for negative class 87% precision, and 89.7% recall is achieved in the determination of opinion words. The result is performed from 400 collected reviews dataset using eight features.

	System prototype output		
Domain expert judgment		Positive	Negative
	Positive	180	20
	Negative	26	174

Table 5-4Confusion Matrix of System Performance Testing

As presented in the confusion matrix (table5.4), the accuracy of the proposed system is 88.3% is achieved for aspect based opinion summarization. Forty six opinions are miss classified from given four hundred opinion reviews. Comparisons between manpower and automatic opinion mining and aspect based opinion summarization result shows automatic aspect based opinion summarization system is saving time and manpower costs. For example, within 2 minutes 400 opinionated reviews are classified by the proposed system; while it takes 60 minutes, we classify the polarity and aspect based opinion summarization using four person. This means that the proposed system saves more 96 % of the time wasted by individual to clarify manually.

The accuracy of data is very good result because of most of the nature of opinion review we used from ORTO for is directly expressed or direct opinion but if the nature of opinion review is here in ORTO is context based or semantic opinion at most than direct opinion in the review the performance of the system decrease. There are some challenges found throughout the study, which limit the system prototype to score a better performance for opinion mining, and aspect based opinion summarizing. These challenges are discussed as follows: The first one is users give their opinion in context based or/and indirect manner this is another challenging task to identify the polarity of opinions. Means writing reviews in Afaan Oromo, many reviewers use positive opinion terms to express negative opinions. For example: in the review "jalqaba kan ormaa odeessuu irraa osoo waa'ee keenya baree gaarii dha. /before talking about other it's better to know our service. "Even though, the expressed opinion is negative but the system prototype labeled it as positive because of the reviewer used the positive opinion terms 'gaarii' (better) to express negative opinions toward the Oromia radio and television organization in news domain. In additional Afaan Oromo language, have, difficult writing system means if the users cannot write correctly the spelling the system cannot understand the opinions.

CHAPTER SIX

CONCLUSION AND RECCOMENDATION

6.1 Conclusion

The inventions of web 2.0 have concreted the way for the rapid growth of user generated content on the web. Currently we are in the world in which masses of user generated content (i.e. word of mouth) are easily unrestricted online on the web in different domain. This user generated content can be posted online by an opinion holder via different review sites, blogs, discussion forums, social media review, and different organizations and non-organizations site. This massive number of user-generated content can be difficult to collect, understand, summarize, analyze manually. This necessitate the need for good tool for automatically mining and summarizing sentiments at aspect level on a given product, or service reviews in order to understand the attitude of the opinion holder on some particular entity (i.e. object). The ability to automatically extract, classify and summarize opinions from texts would be extremely helpful to individuals, organization, business intelligence, government intelligence and others in making effective decision.

The process of sentiment mining and aspect based opinion summarization in Afaan Oromo language from domain of Oromia Radio and Television news takes place by collecting news review as input via web, and extracts the features that identifies a product aspect which are being commented by customer or user, determines the opinions polarity, and produces a summary extracted from opinion polarity on aspect as an output.

In this study, the problem of identifying and extracting aspect/features and determining opinion polarity and then generating the summary is done using rule based approach.

The result of the experiment shows the performance of the method. From the experiment, we observe performance of 90 % precision and 87.1 % recall in for positive class, for negative class 87% precision, 89.7% recall and is 88.3% system accuracy is achieved in the determination of aspect based opinion summarization.

The Strength of the study is opinion mining and aspect-based summarization of news review that simplifies visualization of opinion. In this study bar chart, based visualization is attempted,

which is also accepted by user into very good user acceptance ratio. Accordingly, the system registers promising result.

However, since the system did not take into account context it is easily affected by indirect comment given by the user about ORTO'S service. The other is Afaan Oromo languages have difficult writing system means if the users cannot write correctly the spelling the system cannot understand the opinions.

6.2 Recommendations

Even though this study attempts to develop and design automatic opinion mining and aspect based summarization, developing and designing a full-fledged, fully functional and a more efficient and effective analysis is still recommended in opinion mining. According, the following future research areas are recommended.

- In this research, we have considered the direct opinion for aspect based opinion summation, so it is difficult to identify the polarity of opinions, which have context, based opinion, for future study context based opinion mining and summarization is another direction that can be considered.
- In this research, we used subjective Afaan Oromo opinionated review, for a future work identifying subjective and objective document considered, which is used to minimize time effort given for classifying subjective and objective texts manually. Subjective text is opinions given on a product or services whereas objective means fact information given on object, event, product and service.
- Opinion spam detection or fake review detection is other research direction that can be considered, these opinions are false positive or false negative opinions, which lead customer in to wrong decision.
- The other direction goes for opinion question answering, which is about mining and answering. Here, a set of questions and a group of documents are obtainable to an automatic NLP system. This method is employed to retrieve the response to the questions in Natural Language
- In this study, we focus only on one language, for future research multi-lingual opinion mining and summarization is another research direction. Multi lingual opinion mining

and summarization helps users who use different language to use the customer of user emotions or feeling on a given product or service.

- The other research direction to be considered is going for text based opinion summarization, that deals on summarizing important opinion in a given reviews on a given domain.

References

[1] Bing Liu, "Sentiment Analysis and Opinion Mining," *Morgan & Claypool Publishers*, April 22, 2012.

[2] Bo Pang and Lillian Lee, "Opinion mining and sentiment analysis," *Foundation and Trends in Information Retrieval*, vol. 2, pp. 1-135, 2008.

[3] Bing Liu, "Sentiment Analysis," *5th Text Analytics Summit ,Boston*, pp. 1-2, June 2009.

[4] Lei Zhang and Bing Liu, "Aspect and Entity Extraction for Opinion Mining," *Data mining and knowledge discovery for big data*, pp. 1-40, 2014.

[5] G.Vinodhini and R M.Chandrasekaran, "Sentiment analysis and Opinion Mining:A survey," *International Journal of Advanced Resarch in Computer Science and Software Engineering*, vol. 2, no. 6, pp. 1-11, June 2012.

[6] Lei Zhang, "Aspect and Entity Extraction from Opinion Documents," Department of Computer Science, Illinois University, Chicago, PhD Thesis 2012.

[7] Esra Akba, "Aspect based opinion mining on turkish tweets," Department Of Computer Engineering , Bilkent University, Msc Thesis 2012.

[8] Alasmar, Ahmed M., "Feature Based Approach in Arabic Opinion Mining Using Ontology," Department of Information Technology, The Islamic University–Gaza, Msc Thesis 2016.

[9] Tulu Tilahun Hailu, "Opinion Mining from Amharic Blog," Department of Computer Science, Addis Ababa University, Ethiopia, Masters Thesis 2013.

[10] Ephrem Alamerew, "Automatic Annotation of Opinionated Amharic Text For Opinion Mining," Department of Information systems, Debre Berhan University, Ethiopia, MSc Thesis 2016.

[11] Selama Gebremeskel, "Sentiment Mining Model for Opinionated Amharic Texts," Department of Computer Science, Addis Ababa University, Ethiopia, Msc Thesis 2010.

[12] Noah Smith and David Smit. (2005) Empirical Research Methods in Computer Science. [Online]. http://www.cs.jhu.edu/~nasmith/erm/ (Accessed date 4/6/2017)

[13] Mcsherry,Frank and Marc Najork, "Computing information retrieval performance measures efficiently in the presence of tied scores," *Advances in information retrieval*, pp. 414-421, 2008.

[14] Samha, Amani Khalaf, "Aspect-Based opinin Mining from customer," Department of

Computer science, Queensland University of Technology, PhD Thesis, July 2016.

[15] Richa Sharma,Shweta Nigam and Rekha Jain, "Mining of Product Reiews at Aspect Level," *International Journal in Foundation of Computer science & Technology(IJFCST)*, vol. 4, pp. 1-15, May 2014.

[16] Dudhat Ankitkumar , R.Badre and Prof.Mayura Kinikar, "A Survey on Sentiment Analysis and Opinion Mining," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 2, no. 11, pp. 1-7, November 2014.

[17] G.vinodhini and R M.Chandrasekaran, "Sentiment analysis and Opinion Mining:A survey," *International journal of Advanced Resarch in Computer Science and Software Engineering*, vol. 2, no. 6, pp. 1-11, June 2012.

[18] Lun wei ku,Yu-Ting Liang and Hsin-Hsi chen, "Question analysis and answer passage retrieval for opinion question answering systems," *Computational Linquistics and chinese Language Processing*, vol. 13, pp. 307-326, September 2008.

[19] Jindal,Nitin and Liu Bing, "Identifying comparative sentences in text documents," *Proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval*, pp. 244-251, 2006.

[20] Asokan, Neethu Kuriana and Shimmi, "Summarizing User Opinions: A Method for Labeled-Data Scarce Product Domains," *International Conference on Information and Communication Technologies (ICICT)*, pp. 93-100, 2014.

[21] Regina Barzilay, "Advance in Automatic Text summarization," *Cambridge*, December 2005.

[22] Neethu Kuriana and Shimmi Asokan, "Summarizing User Opinions: A Method for Labeled-Data Scarce Product Domains," *International Conference on Information and Communication Technologies (ICICT)*, pp. 93-100, 2014.

[23] Arti Buche, Dr. Chandak and Akshay Zadgaonkar, "Ooinion mining and analysis :Asurvey," *International Journal on Natural Language Computing (IJNLC)*, vol. 2, pp. 1307-3336, June 2013.

[24] Richa Sharma,shweta and Rekha Jain, "Mining of Product Reiews at Aspect Level," *International Journal in Foundations of Computer Science & Technology (IJFCST)*, vol. 4, May 2014.

[25] Edison Marrese Taylor Juan D. Velasquez and Felipe bravo-Marquez, "A Novel Deterministic Approach for Aspect-Based Opinion Mining in Tourism Products Reviews," *Expert Systems with Applications*, pp. 7764-7775, 2014.

[26] Mongkol Saensuk, Panida Songram and Phatthanaphong Chomphuwiset, "Feature-Based Opinion Mining on Smart-Phone Reviews," Department of Computer Science,

Mahasarakham University, Kantarawichai District, Mahasarakham, Thailand , 2015.

[27] Lizhen Liu, Zhixin Lv and Hanshi Wang, "Extract Product Features in Chinese Web for Opinion Mining," *Journal of software*, vol. 8, pp. 1-6, March 2013.

[28] Abreham Getachew, "Opinion Mining from Amharic Entertainment Texts," Department of Information Science, Addis Ababa University, Ethiopia, MSc Thesis October, 2014.

[29] Mohammed Tune, "Designing A Graph-Based Opinion Mining Model for Opinionated Text in English, Amharic and Afaan Oromo Language," Department of Information Systems, Debre berhan university, Ethiopia, MSc Thesis February 2015.

[30] Abebe Abeshu Diro, "Automatic Morphological Synthesizer for Afaan Oromoo," Department of computer Science, Addis Ababa university, Ethiopia, Msc Thesis.

[31] Girma D, "Afaan Oromo News Text Summarizer," Faculty of Informatics, Addis Ababa University, Ethiopia, Msc Thesis 2013.

[32] Mandefro Legesse Kejela, "Named Entity Recognition for Afan Oromo," Department of computer Science, Addis Ababa University, Ethiopia, Msc Thesis 2010.

[33] Wakshum Mekonnen, "Development of stemming algorithm for Afaan Oromo texts," Department of Information Science, Addis Ababa University, Ethiopia, Msc Thesis 2000.

[34] Debela Tesfaye, "Designing a Stemmer for Afan Oromo Text: A hybrid approach," Department of Information science, Addis Ababa University, Ethiopia, Msc Thesis.

[35] Gezehagn Gutema Eggi, "Afaan Oromo Text Retrieval System," Department of Information science, Addis Ababa University, Ethiopia, Msc Thesis 2012.

[36] Irresoo Nagi, *Caassefama Afan oromo*. Addis Ababa, Ethiopia: Kuraz International, 2006.

[37] Gumii Qormaata Afan Oromoo, *Caasluga Afan Oromoo Jildi I" Komishinii Aadaaf Turizimii Oromiya*. Finfinnee,Ethiopia, 1995.

[38] Wondimu Tegegne, "The Development of Written Afan Oromo and the Appropriateness of Qubee, Latin Script, for Afan Oromo Writing ," *Historical Research Letter* , vol. 28, 2016.

[39] C Meyer, "On Improving Natural Language Processing through Phrase-based and one to-one Syntactic Algorithm," Kansas State University Manhatan, Kansas, Masters Thesis 2008.

[40] (2017) Substitutions and sentence splitting. [Online]. http://docs.pandorabots.com/tutorials/substitutions-and-sentence-splitting/

[41] Liu Bing and Lei Zhang, "A survey of opinion mining and sentiment analysis," *In Mining text data*, pp. 415-463, 2012.

[42] Mongkol Saensuk, Panida Songram, Phatthanaphong Chomphuwiset, "Feature-Based Opinion Mining on Smart-Phone Reviews," *Proceedings of the 3rd IIAE International Conference on Intelligent Systems and Image Processing*, pp. 1-5, 2015.

[43] Bing liu, *Sentiment Analysis and opinion mining*.: Morgan and Claypool Publishers, May 2012.

[44] Mohammed Sadegn,Roliana Ibrahim and Zulaiha Ali Othman, "Opinion Mining and Sentiment Analysis," *International Journal of Computer and Technology*, vol. 2, pp. 1-8, June 2012.

[45] Milind D. Meshram, "Feature based opinion mining : an overview," *Proceedings of National Conference on Emerging Trends: Innovations and Challenges in IT*, April 2013.

[46] Walaa Medhat,Ahmed Hassan and Hoda Korashy, "Sentiment analysis algorithms and applications: A Survey," *Ain Shams Engineering Journal*, pp. 1093-1113, 2014.

[47] Soujanya Poria,Erik Cambria and Alexander Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowle dge-Base d Systems*, pp. 42-49, 2016.

[48] Soujanya Poria,Erik Cambria,Lun-Wei Ku,Chen Gui,Alexander Gelbukh, "A Rule-Based Approach to Aspect Extraction from Product Reviews," *In Proceedings of International Conference on World*, 2014.

[49] Kaufmann JM. JMaxAlign, "A Maximum Entropy Parallel Sentence Alignment Tool," *In: Proceedings of COLING'12:Demonstration Papers, Mumbai*, pp. 277–88, 2012.

[50] (2009, Apr.) tokenization. [Online]. https://nlp.stanford.edu/IR-book/html/htmledition/tokenization-1.html

[51] Po Pang and Lillian Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, pp. 1–135, 2008.

[52] lu, Yue and chengxiang Zhai, "opinion integration through semi-supervised topic modeling.," *proceeding of the 17th international conference on world wide web.*, pp. 121-130, 2008.

[53] Bing Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, pp. 1-167, 2012.

[54] Eric Brill, "A simple rule-based part of speech tagger," *Proceedings of the workshop on Speech and Natural Language*, pp. 112-116, February 1992.

[55] Raut, Vijay B., and D. Londhe, "Survey on Opinion Mining and Summarization of User Reviews on Web," *International Journal of Computer Science and Information*

Technologies, vol. 5, pp. 1026-1030, 2014.

[56] Marc, Lewis David D and Ringuette, "A comparison of two learning algorithms for text categorization," *Third annual symposium on document analysis and information retrieval*, vol. 33, pp. 81-93, 1994.

[57] Chakrabarti Soumen,Roy Shourya and Soundalgekar Mahesha , "Fast and accurate text classification via multiple linear discriminant projections.," *The International Journal on Very Large Data Bases*, pp. 170-185, 2003.

[58] Ali, A'sim Seedahmed, "Opinion Mining Techniques," *International Journal of Innovative Science, Engineering & Technology*, vol. 2, no. 6, pp. 752-755, June 2015.

[59] Ayesha Rashid, Naveed Anwer,Dr. Muddaser Iqbal and Dr.muhammed sher, "A Survey Paper: Areas, Techniques and Challenges of opinion mining," *International Journal of Computer Science Issues*, vol. 10, no. 6, pp. 18-31, November 2013.

[60] Mily Lal and Kavita Asnani, "Aspect extraction & segmentation in Opinion mining," *International Journal Of Engineering And Computer Science*, vol. 3, no. 5, pp. 5873-5878, May 2014.

[61] S.Mounika and P.Senthil Raja, "Frequent Features on Aspect-Based Opinion Mining using Multiproduct Reviews," *International Journal of Recent Trends in Engineering & Research (IJRTER)*, vol. 2, no. 2, February – 2016.

[62] Hyun Duk Kim,Kavita ganesan,parikshit sondhi, "Comprehensive Review of Opinion Summarization," pp. 1-30, 2011.

[63] Amandeep Kaur, "A Survey on Sentiment Analysis and Opinion Mining Techniques," *Journal of Emerging Technologies in Web Intelligence*, vol. 5, pp. 1-5, November 2013.

Appendix A: Data gathering form for Oromia radio and television organization

ደብረ ብርሃን ዩኒቨርሲቲ

በኮምፒውተንግ ኮሌጅ
የኢንፎርሜሽን ሲስተም

DEBRE BERHAN UNIVERSITY
Computing college
information system

ለ ኦሮሚያ ራድዮና ቴሌቪዥን ድርጅት

ጉዳዩ፡ ትብብር እንድደረግላቸው ስለመጠየቅ

በደብረ ብርሃን ዩኒቨርሲቲ የኢንፎርሜሽን ሲስተም ትምህርት ክፍል ለማስተርስ ድግሪ ምርምር ማግኘ የሚሆን በአሮሞኛ ቋንቋ ላይ ሙሪት ደደረገ ኮምፒተራይዝድ አስተያየት መለያ ሲስተም ለመሪት ማብሉት የሚሆኑ የኦሮሞኛ አስተያት ለመሰብሰብ ነው። ስለዚህ የተሰመደውን ትብብራችሁን እንድታደርጉልን በትህትና እንጠይቃለን።

ስም

መማደረስ ታሪኩ

ID

PGTSR/016/08

ከሰላምታ ጋር

አንዳልካቸው አለ

የኢ.ሰ.ኮ/አስተባባሪ

ግልባጭ

- ለ ኮምፒውተንግ ኮሌጅ

ደብረብርሃንሲቲ

P.O.Box 445
E-mail: dbubethuonet.et
In replying please quote our ref. no.
☎ +251-1168160392
Fax +251-1166120665

Appendix B: List of Afaan Oromo positive and negative keywords

Haqummaa/+

Inaaffaa/-

Gowwaa/-

Abbaa irree/-

Oftuulummaa/-

Qusannaa/+

Rakkachuu/-

Gaabbii/+

Itti fufinsa/+

Dhugaa/+

Boo'icha/-

Fooyya'aa/+

Cimina/+

Bushaa'aa/-
Kallattiin/+

Hattuu/-

Mirga/+

Doqna/-

Gargaarsa/+

Jinnii/-

Quufa/+

Cilee/cilaattii/-

Amala/-

Maamila/+

Duukaa bu'aa/+

Dhabuu/-

Qabeenya/+

Cimaa/+

Gammadaa/+

Iyya/_

Gadhee/_

Wal amanuu dhabuu/_

Obsa/+

Qulqulluu/+

Kan hin amanamne/-

Hanqina/-

Kabaja/+

Badii/-

Seeraan ala/-

Kabaja/+

Kan hin dagatamne/+

Fakkeenya gaarii ta'uu/+

Daba/-

Wal'aansoo/-

Guddina/+

Iyya/-

Jinnii/-

Jibba/-

Sagalee/+

Abbaa buddeenaa/+

Gadhee/-

Bohaarsaa/+

Ta'us/+

Hawwataa/+

Gatii/+

Barataa/+

Gubboo/-

Bayeessa/+

Fakkeenya/+

Gahee/-

Ammayya/+

Nageenya/+

Nuffisiisaa/-

Ammayya/+

Xaxamaa/-

Amanamaa/+

Kan hin mo'amne/+

Sagaagalummaa/-

Waca/-

Jibbisiisaa/-

Amansiisaa/+

Awwaala/-

Dinqisiifannaa/+

Dandeettii/+

Barumsa/+

Yaalii/+

Hirmaannaa/+

Gargaarsa/+

Dhibee/-

Bareedaa/+

Raajii/+

Kan yaadu/

Jabaadhaa/+

Boonuu/

Jal'aa/-

Fokkisaa/-

Walitti bu'insa/-

Komii/-

Mi'aawaa/+

Barbaadamaa/+

Goyyomsuu/-

Jannata/+

Olka'aa/+

Gantuu/-

Aarii/-

Naafa/-

Ajaa'ibsiisaa/+

Wanta mul’atu/+

Fakkeenya/+

Hidhamaa/-

Boonaa/+

Ifa/+

Kan hin dubbanne/+

Mul’ata/+

Kan hin mirkanoofne/-

Walgargaaruu/+

Galagaadhiisee kan jooru/-

Dhuufuu/-

Inaaftuu/-

Kan hin barbaachifne/-

Lola/-

Hacuuccaa/-

Mijataa/+

Gootummaa/+

Kutataa/+

Kan adamsu/-

Iyyuu/-

Tasgabbaa’uu/+

Hongee/-

Eebbifamaa/+

Ergisa/-

Balaa ibiddaa/_

Maseena/-

Jette jettee/-

Jeeqaa/-

Miidhamaa/-

Itti gaafatamummaa/+

Cimina/+

Kan hin dammaqne/-

Ogummaa/+

Faayidaa dhabeessa/-

Fooyya’aa/+

Dhugaa/+

Baacuu/-

Kan hin eegamne/-

Dammaquu/+

Tola ooltummaa/+

Summaa’aa/-

Jarjaraa/-

Haaloo kuusuu/_

Kan faallessu/-

Fiixaan bahuu dhabuu/-

Maqaa/+

Tortoraa/-

Raajii/+

Oduu/-

Mala/+

Kan tasgabbaa’e/+

Ulfina/+

Latuu/+

Diddaa/-

Tokkommaa/+

Beekaa/+

Ba ‘aa/-

Mala dhahuu/+

Maseena/-

Bilisummaa/+

Ganuu/-

Xiqqoo/-

Xiyya/

Kan adda qoqqoodu/-

Halkan/-

Mormituu/-

Middhaa/-

Kan suuta jedhu/-

Amala/-

Ija jabeessa/+

Kan ukkaamsu/-

Soba/-

Kan miidhaa geessisu/-

Mijachuu/+

Adda kan ta’e/+

Dandeettii/+

Garaa kutachuu/+

Kan gaddisiisu/-

Amansiisaa/+

Faayidaa/+

Gabrummaa/+

Sobduu/-

Kan seera jabeessu/-

Qoraa/-

Qulqulluu/+

Goyyomsuu/+

Goomii/-

Kan hin galleef/-

Ifaa/+

Goomataa/-

Shororkeessaa-

Dhiphina/-

Araara/+

Guddina/+

Iyya/-

Jecha/+

Kan hin yaadamne/-

Kan murtessuu hin dandeenye/-

Qomummaa/-

Jallataa/-

Gidduu galluu/-

Gadhee/-

Abaaramaa/-

Kan jeequ/-

Garraamii/+

Kan hin daangeffamne/-

Jeequmsa/-

Jibbisiisaa/-

Balaa/-

Eebba/+

Of eeguu/+

Dhiphina/-

Gammachuu/+

Fudhatama/+

Mormii/-

Kan nama aarsu/-

Himata/-

Sodaa/-

Galma hin ga’iin/-

Kan raawwatamu/+

Balleessuu/+

Humnaa ol/-

Fannisaa/+

Jibbisiisaa/-

Seeraan ala/-

Abbaa humnaa/+

Fakkeenyummaa/ +

Qaruxee/+

Balaa/-

Kan balleessaa qabu/-

Jiraataa/+

Badii/-

Kennaa/+

Hanqina/-

Shakkisiisaa/-

Kennaa/_

Oftuuluu/-

Dhiifama/+

Haftuu/-

Fayyaa/+

Kan jalqabaa/+

Collee/+

Kan ariifatu/-

Badduu/-

Balleessaa/-

Jibbisiisaa/-

Duguuggaa/-

Filatamaa/+

Badii/-

Jooraa/-

Kennaa/+

Baqataa/-

Jeeqaa/-

Saaxilamuu/-

Miidhaa/_

Galma ga’umsa/+

Sodaa/-

Obsaa/+

Dinqisiisaa/+

Mormaa/-

Gogongorsaa/-

Carraa/+

Kan adda ta’e/+

Of eeggannoo/-

Kan hawaasaa/+

Ergaa/+

Onnee/+

Kan maqaa qabu/+

Bareedina/+

Wal’aansoo/-

Buqqee/_

Ajjeechaa/-

Kan saaxilame/-

Wal gargaaruu/+

Badhaasa/+

Kan walgitu/+

Kan shakku/_

Eebbifamaa/+

Kan dhowwu/+

Hunda kan danda’u/+

Kan hin fooyyessine/-

Lola/-

Kan hin leenjine/-

Bareedaa/=

Maqaa qabeessa/+

Hiyyeessa/-

Foo’uu/-

Dogongora/-

Kan ol ka’e/+

Kan barate/+

Leenca/+

Adda/+

Hinqirfuu/-

Yaada qabeessa/+

Milkaa’uu/+

Jibba/-

Fokkisaa/-

Itite/+

Soorummaa/+

Kaayyoof kan dhhabate/+

Samuu/+

Dhabeessa/-

Beekamaa/+

Baaduu/+

Kan fudhatama qabu/+

Ariifachiisaa/-

Gaabbii/-

Amantii/+

Dhiifama/+

Jiraataa/+

Ulfaataa/-

Saree/-

Badii/-

Haaloo kuusuu/-

Gufuu/-

Dabaree/-

Hirkattummaa/+

Qabatama/+

Duwaa/-

Diina/-

Miidhagina/+

Tuullaa/+

Gaarii/+

Mala/+

Gara laafessa/+

Kan seeraa/+

Daangaa/+

Miidhage/-

Gabra/-

Qulqullaa'e/+

Safuu/+

Aadaa/+

Maraatuu/-

Kan namaaf gaddu/+

Duubatti hafaa/-

Kan wal make/-

Walitii naquu/-

Duubatti hafummaa/-

Ajajaa/-

Cimaa/+

Kan hin galleef/-

Ifaa/+

Kan Mirkanaa'e/+

Sodaachisaa/-

Reebumsa/-

Kanjallate/-

Daakaa/+

Kan tuqame/-

Fakkeessituu/-

Irra deebii/-

Ilaalcha gaarii/+

Balaa tasa/-

Kan dhugaa dubbatu/+

Kan afuuraa/+

Obsa/+

Of eeguu/+

Gadi badaa/-

Faayidaa/+

Mo'amuu/-

Cimina/+

Kan haajaa hin qabne/-

Kennaa/+

Kan nama jajjabeessu/+

Gonkumaa/-

Kan ogummaa qabu/+

Obsa dhabuu/-

Galma/+

Kan filame/+

Qabeenya/+

Dambii/+

Bushaa'uu/-

Kan nama kofalchiisu/+

Qaroomina/+

Alagaa/-

Foolii/+

Seera/+

Daguu/-

qalbii harkisuu/+

fakkeessuu/-

saamtuu/-

ergama/+

dansa/+

mil'uu/-

sobduu/-

ajaa'ibsiisaa/+

boo'uu/_

kan badiin hin qabne/+

miidgase/+

addunyaa/+

fakkeessituu/-

dhukkubsataa/-

taphataa/+

obsa/+

kan uffate hin uffanne/-

gatii/-

gammachuu/+

dukkana/-

Dammaqe/+

Dhiibaa/-

Milkii/+

Kan maraate/-

Dhiphummaa/-

Oftuulummaa/-

Kaballaa/-

Firii/+

Miidhaa/-

Kan nama jajjabeessu/+

Bu'uura/+

Gowwomsuu/-

Fayyummaa/+

Kan dhugaa dubbatu/+

Garraamummaa/+

Tola/+

Ajjeechaa/-

Mijataa/+

Hiriyaa/+

Kan hin mijanne/-

Dhara/-

Kan waliin hin makamne/-

Gamna/+

Akka feetee/-

Tasa/-

Cubbuu/-

Rakkoo/-

Gadhee/-

Hadhaawaa/-

Imaanaa/+

Gatii/-

Bareeda/+

Obsa/+

Kan mormu/-

Namummaa/+

Rifate/-

Seera malee/-

Faayidaa/+

Adda/+

Hir'ina/-

Dadhibbina/-

Cittoo/-

Qorumsa/+

Balleessuu/-

Rakkataa/-

Dhuminsa/-

Baayyina/+

Qabeenya/+

Dandeettii/+

Tola ooltummaa/+

Irratti ka'uu/-

Misha/+

Kuttuu/-

Kan hin ganamne/+

Ammayyummaa/ +

Kan hin leenjine/-

Cimaa/+

Haarrii baase/-

Kan sobu/-

Shakkii/-

Ulfina/+

Kan hin mine/+

Oddeessituu/-

Mootummaa/+

Dhiphina/-

Salphina/-

Kennaa/+

Nageeynummaa/+

Tajaajila/+

Xurii/-

Kosii/-

Rakkina/-

Ajajamaa/+

Soda/-

Kan nama quubsu/+

Garaa hanqachuu/-

Namaaf gadduu/+

Danda'aa/+

Gaagura/-

Dheebuu/-

Badii/+

Misooma/+

Hangafa/+

Dursaa/+

Filatamaa/+

Kan nama jibbisiisu/-

Osoo hin shakkiin/+

Morme/-

Dhiittaa/-

Drbinsa/+

Kan waliif hin galle/-

Kan tole hin jenne/-

Sirrii kan hin taane/-

Fayyaa/+

Gaarii/+

Faayidaa/+

Bayeessa/+

Soda/-

Qoricha/+

Ariifachuu/-

Waliif galuu/+

Fayyina/+

Ifa/+

Rifachiisaa/-

Hunaa ol/-

Qabeenya/+

Seera qabeessa/+

Raajii/+

Duudaa/-

Kan hir'isu/-

Mormii/-

Kan qaanessu/-

Gammachuu guddaa/+

Sababa/-

Kanaa olitti/+

Abaarsa/-

Kan jeequ/-

Maqaa qabeessa/+

Qaanii/-

Xuraawaa/-

Kanaa olitti/-

Gatii baasuu/+

Hiyyeessa/-

Qabatamaa/+

Jarjarsa/-

Ho’a/+

Ogummaa/+

Firii/+

Itti fufinsa/+

Addaachuu/+

Mormii/-

Raajii/+

Kan sirrii ta’e/+

kan hin kennine/-

kan miidhage/+

Sabummaa /+

Firii/+

Kan nama hin mufachiifne/+

Haarrii/-

Kaayyoof kan dhaabatu/+

Alagaa/-

Dhukkubbii/-

Diiguu/-

Diina/-

Carraa/+

Kan mirkanaa’e/+

Miidhagduu/+

Bitaa/-

Kadhattuu/-

Dadhabbina/-

Abjuu/-

Xuraawaa/-

Dadhabsiisaa/-

Baqe/-

Qulqullummaa/+

Mala/+

Faallaa/-

Nama amantaa/+

Ariifachiisaa/-

Dhiphaa/-

Dhukkuba/+

Qophii/+

Reebumsa/-

Hamii/-

Si’annaa/+

Mo’uu/+

Qophaa’aa/+

Hubachuu/+

Qorichummaa/-

Salphoo/-

Balleessituu/-

Gaarumaa/+

Kan obbsa qabu/+

Gamna/+

Qonxulummaa/-

Kuusama/+

Jeeqxota/-

Abaaramaa/-

Jibbisiisaa/-

Muuxannoo/+

Kan ganu/-

Oduu dheeressuu/-

Of eeggannoo/+

Arrabsoo/-

Kaka’umsa/+

Gowwaa/-

Balaa/-

Murteessaa/+

Aarii/-

Gara laafummaa/+

Walii galuu dhabuu/-

Gadi bu’aa/-

Garaa wal dhabuu/-

Faayidaa/+

Gufachiisuu/-

Gogaa/-

Suga/+

Kan galma ga’e/+

Dadhabaa/-

Kaayyoo/+

Sooressa/+

Adabbii/-

Sirreessaa/+

Kan ifa hin taane/-

Tokkummaa/+

Walii galuu/+

Jeequmsa/-

Walxaxaa/-

Salphise/-

Kan badii dalagu/-

Walii galuu dhabuu/-

Godaannisa/-

Amala gadhee/-

Naasuu/-

Jarjaraa/-

Olaanaa/+

Himannaa/-

Jarjarsuu/-

Kan ariitii qabu/+

Fedhii/+

Marii/+

Diddaa/-

Yaalii/+

Dadhabaa/-

Doqna/-

Amansiisaa/+

Seera malee/-

Ba’aa/-

Kashalabbee/-

Ogummaa/+

Madaallii/+

Ga’aa/+

Balleessuu/-

Dammaqinsa/+

Kutataa/+

Kabaja/+

Kan hin beekamne/-

Bushaa'uu/-

Gatii /-

Dhiibbaa/-

Iccitii/-

Dhibee/-

Gadda/-

Gargaarsa/+

Boo'icha/-

Maamila/+

Bareeda/+

Tajaajilaa/+

Kan obsa qabu/+

Gara laafessa/+

Miidhage/+

Amala/+

Salphina/-

Amanamummaa/+

Giiftii/+

Dhabuu/-

Goota/+

Kan dhorkame/-

Yaalii/+

kabajamaa/+

wal jaalachuu/+

gammachuu/+

gahee/+

barbaachisaa/+

faalamuu/-

beekkamtii/+

Kan fayyaa qabu/+

Gamna/+

Jabaa/+

Walii galtee/+

Jallataa/-

Rakkina/-

Hojii gaarii/+

Madaa/-

Dhiphisuu/-

Carra qabeessa/+

Kan oolmaa namaaf oolu/+

Nyaata/+

Foolii gaarii/+

Amansiisaa/+

Gooftaa/+

Kan hin taane/-

Kophummaa?-

Kan hin galleef/-

Dhabuu/-

Ajajamuu/+

Qormaata/-

Kan hin argamne/-

Arjaa/+

Hiyyeessa/-

Meetii/+

Affeeraa/+

Wallaalummaa/-

Kan ife/+

Miti/-

Kan hin jenne/-

Ta'us/-

Garuu/-

Garmalee/+

Kan ajaa'ibsiifamu/+

Ajaa'iba/+

Natti hin tolle/-

Barsiisa/+

Sagaagaltuu/-

Sanyummaa/-

Hinqirfuu/-

Tole/+

Mijataa/+

Nageenya/+

Fakkeenya/+

Abdii/+

Guddaa/+

Hidhaa/-

Dhugaa/+

Qulqullina/+

Filatamaa/+

Gammachuu/+

Misha/+

Jaallanneerra/+

Dadhabbina/-

Gaddisiisaa/-

Kan afuuraa/+

Dinqisiisaa/+

Gargaaraa/+

Kan miidhagu/+

Tola/+

Hawwataa/+

Bashannansiisaa/+

Bushaa'uu/-

Bilisummaa/+

Gadi bu'aa/-

Haaraa/+

Miidhagina/+

Ajaa'ibsiisaa/+

Gaarii/+

Gammachiisaa/+

Bashannansiisaa/+

Fagoo/-

Uumaa/+

Summii/-

Dhugaa/+

Jaalala/+

Ergifataa/-

Ergaa/+

Amanamaa/+

Rakkina/-

Balleessawwan/-

Kan milkaa’e/+

Guddina/+

Fokkisaa/-

Galateeffatama/+

Faayidaa/+

Ga’umsa/+

Jaaladheera/+

Shakkii/-

Fooyya’aa/+

Natti toleera/+

Beekamaa/+

Gammachiisaa/+

Hir’ina/-

AppendixCSample sourcecode of aspect based opinion summarization.

```
package opinion.summarize;

import java.io.File;

import java.io.FileReader;

import java.io.IOException;

import java.sql.Connection;

import java.sql.DriverManager;

import java.sql.ResultSet;

import java.sql.SQLException;

import java.sql.Statement;

import javax.swing.JOptionPane;

import org.jfree.ui.RefineryUtilities;


/**

 *

 * @author hp

 */

public class OpinoinSummarize extends javax.swing.JFrame {


public static String title;

public static int posative;

public static int negative;

public static int neutral;


public static String [] category;

public static int [] pos;

public static int [] neg;
```

```
    Connection con = null;
    Statement stm = null;
    ResultSet rs = null;

    /**
     * Creates new form OpinoinSummarize
     */
public OpinoinSummarize() {
initComponents();
        //btnSummarize.setVisible(false);
btnSummarize.enable(false);
btnClear.enable(false);

try {
Class.forName("com.mysql.jdbc.Driver").newInstance();
con = DriverManager.getConnection("jdbc:mysql://localhost:3309/afaan_oromoo_dictionary",
"root", "root");
stm = con.createStatement();
        } catch (Exception e) {
JOptionPane.showMessageDialog(null, e, "Connection",
JOptionPane.INFORMATION_MESSAGE);


        }
    }

if (!(txtDisplay.getText().equals(""))) {

String[] titles = txtDisplay.getText().split("Title: ");
category =new String[titles.length];
```

```
pos =new int [titles.length];
neg =new int [titles.length];
for(int t=1;t<titles.length;t++){
negative=0;
posative=0;
String[] opinion = titles[t].split("\n");

title = opinion[0];
for (int i = 0; i < opinion.length; i++) {
if(i==0){
category[t]=opinion[0];
                }
try {
                    String sufix, prefix;
sufix = "hin";
prefix = "miti";
rs = stm.executeQuery("SELECT * FROM dictionary");
boolean Records = rs.next();
if (Records) {
do {
                        String type = rs.getString(2);
                        String kewWord = rs.getString(1);
if (opinion[i].matches("(?i).*" + kewWord + ".*")) {
int sufixIndex = opinion[i].indexOf(sufix);
int prefixIndex = opinion[i].indexOf(prefix);
int kewWordIndex = opinion[i].indexOf(kewWord);
int sufixLen = sufix.length();
int prefixLen = prefix.length();
```

```
int keyWordLen = kewWord.length();
switch (type) {

case "Posative":
if (opinion[i].matches("(?i).*" + sufix + ".*") || opinion[i].matches("(?i).*" + prefix + ".*")) {
if (((sufixIndex + 2 + sufixLen) == (kewWordIndex)) || ((sufixIndex) == (+kewWordIndex + 2 + keyWordLen)) || ((prefixIndex) == (+kewWordIndex + 2 + keyWordLen))) {
negative = negative + 1;
                    }
                } else {
posative = posative + 1;
                }
break;
case "Negative":
if (opinion[i].matches("(?i).*" + sufix + ".*") || opinion[i].matches("(?i).*" + prefix + ".*")) {
if (((sufixIndex + 2 + sufixLen) == (kewWordIndex)) || ((sufixIndex) == (+kewWordIndex + 2 + keyWordLen)) || ((prefixIndex) == (+kewWordIndex + 2 + keyWordLen))) {
posative = posative + 1;
                    }
                } else {
negative = negative + 1;
                }
case "Negative ":
if (opinion[i].matches("(?i).*" + sufix + ".*") || opinion[i].matches("(?i).*" + prefix + ".*")) {
if (((sufixIndex + 2 + sufixLen) == (kewWordIndex)) || ((sufixIndex) == (+kewWordIndex + 2 + keyWordLen)) || ((prefixIndex) == (+kewWordIndex + 2 + keyWordLen))) {
posative = posative + 1;
                    }
                } else {
```

```
negative = negative + 1;
                                    }
break;
                                 }
                               }
                            } while (rs.next());
                         }
                      } catch (Exception ex) {
JOptionPane.showMessageDialog(null, ex, "Error", JOptionPane.WARNING_MESSAGE);
                      }
                   }
neg[t]=negative;
pos[t]=posative;

          }
             DrawBarChart chart = new DrawBarChart("Opinion Summarization");
chart.pack();
RefineryUtilities.centerFrameOnScreen(chart);
chart.setVisible(true);
chart.setAlwaysOnTop(rootPaneCheckingEnabled);
chart.setLocationRelativeTo(jPanel2);



        } else {
JOptionPane.showMessageDialog(null, "No File is imported\n Please, browse the opnion!",
"Empty File ", 1);
          }
```

```
    }

private void jFileChooser1ActionPerformed(java.awt.event.ActionEvent evt) {
        // TODO add your handling code here:
    }

private void btnBrowseActionPerformed(java.awt.event.ActionEvent evt) {

        //int returnVal = jFileChooser1.showOpenDialog(this);
        //if (returnVal == jFileChooser1.APPROVE_OPTION) {

        File file = jFileChooser1.getSelectedFile();
        String fl = "" + file;
if (fl.equals("null")) {
JOptionPane.showMessageDialog(null, "No File is Selected. Please Browsse the File!\n", "No file is Selected", 2);
        } else {
try {
            // What to do with the file, e.g. display it in a TextArea
txtDisplay.read(new FileReader(file.getAbsolutePath()), null);
        } catch (IOException ex) {
txtDisplay.setText("problem accessing file" + file.getAbsolutePath());
JOptionPane.showMessageDialog(null, file.length(), "No file is Selected", 2);

        }
        }
        //  }
        //else {
```

```
        //   jTextArea1.setText("Opening file is CANCELLED");
// }
    }

private void btnKeyWordActionPerformed(java.awt.event.ActionEvent evt) {
        // TODO add your handling code here:
new Afaan_Oromoo_Dictionary().setVisible(true);
    }

    /**
     * @param args the command line arguments
     */
public static void main(String args[]) {
        /* Set the Nimbus look and feel */
        //<editor-fold defaultstate="collapsed" desc=" Look and feel setting code (optional) ">
        /* If Nimbus (introduced in Java SE 6) is not available, stay with the default look and feel.
         * For details see http://download.oracle.com/javase/tutorial/uiswing/lookandfeel/plaf.html
         */
try {
for (javax.swing.UIManager.LookAndFeelInfo info :
javax.swing.UIManager.getInstalledLookAndFeels()) {
if ("Nimbus".equals(info.getName())) {
javax.swing.UIManager.setLookAndFeel(info.getClassName());
break;
                }
            }
        } catch (ClassNotFoundException ex) {
java.util.logging.Logger.getLogger(OpinoinSummarize.class.getName()).log(java.util.logging.Le
vel.SEVERE, null, ex);
```

```
        } catch (InstantiationException ex) {
java.util.logging.Logger.getLogger(OpinoinSummarize.class.getName()).log(java.util.logging.Level.SEVERE, null, ex);
        } catch (IllegalAccessException ex) {
java.util.logging.Logger.getLogger(OpinoinSummarize.class.getName()).log(java.util.logging.Level.SEVERE, null, ex);
        } catch (javax.swing.UnsupportedLookAndFeelException ex) {
java.util.logging.Logger.getLogger(OpinoinSummarize.class.getName()).log(java.util.logging.Level.SEVERE, null, ex);
        }
        //</editor-fold>

        /* Create and display the form */
java.awt.EventQueue.invokeLater(new Runnable() {
public void run() {
new OpinoinSummarize().setVisible(true);
            }
        });
    }
    // Variables declaration - do not modify
private javax.swing.JButton btnBrowse;
private javax.swing.JButton btnClear;
private javax.swing.JButton btnKeyWord;
private javax.swing.JButton btnSummarize;
private javax.swing.JFileChooser jFileChooser1;
private javax.swing.JPanel jPanel1;
private javax.swing.JPanel jPanel2;
private javax.swing.JPanel jPanel3;
private javax.swing.JPanel jPanel4;
```

```
private javax.swing.JScrollPane jScrollPane1;
private javax.swing.JTextArea txtDisplay;
    // End of variables declaration
}
```