International Journal of Research Studies in Science, Engineering and Technology
Volume 7, Issue 7, 2020, PP 10-18
ISSN 2349-476X

Multi-Class Sentiment Analysis from Afaan Oromo Text Based On Supervised Machine Learning Approaches

Negessa Wayessa*, **Sadik Abas**

Department of Information Technology, College of Informatics, Bule Hora University, Ethiopia

***Corresponding Author:** Negessa Wayessa, Department of Information Technology, College of Informatics, Bule Hora University, Ethiopia*

ABSTRACT

Sentiment analysis is the field of data mining and natural language processing that study about analyzes the opinion data from social media like Facebook, Twitter, organization sites, online news report, user reviews and etc. It is process of classifies opinions under defined polarities classes. Sentiment analysis is a hot research area in the field of data mining and natural language processing. In this work, sentence level sentiment analysis is done into five multiple classes- very negative, negative, neutral, positive and very positive. Here, we proposed two methods of supervised machine learning approaches-Support Vector Machine and Random Forest algorithms to classify sentiment polarity from Oromia broad casting network (OBN) Twitter by Ethiopian language Afaan Oromo. We used tokenization, stop word removal, normalization and stemming as preprocessing and tf-idf used as feature extraction. The Performance of proposed approaches Support Vector Machine and Random Forest achieved accuracy 90% and 89% on OBN Twitter dataset with 1810 corpus size respectively.

Keywords: *Sentiment Analysis, Ethiopian language Afaan Oromo, Supervised machine learning, Support vector machine, Random forest, OBN Twitter.*

INTRODUCTION

One of the most important things to transferring and interchanging idea between peoples is through sharing their idea which says opinion. Peoples are able to express their ideas or opinions in the form of posting on social medial like Facebook, Twitter, Forum discussions and etc. within regard to their day to day life. These opinions can be explores about private and public like individuals, organization, government, product, service, politics and etc. [1]. Since opinions analysis can influences the interest of different private and public organizations or companies are highly interested in analyzing and exploring online opinions [4]. Several private and public are interested to know the opinion of the public with respects to their products and services. Also, many companies are interested to know the public feedback with respect to the new policy, rules and regulations set out as well as public services delivered [6]. With the explosive growth of social media (e.g., reviews, forum discussions, blogs, micro-blogs, Twitter, comments, and postings on social network sites), the opinions posted on this social medias are increased huge volumes of opinions time to time in this real world. Since analyzing huge volumes of opinionated text remains need a technique which was called Sentiment analysis.

Sentiment analysis is the field of data mining and natural language processing that study about analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes [1]. Almost all peoples are share opinions text using their own language and also an organizations or companies have social media page using their official language. So that a huge volumes opinions text on private and public social media are increasing time to time by using Ethiopian language Afaan Oromo.

Afaan Oromo (AO) is one of the major African languages that are widely spoken in Ethiopia. Oromo people are the native speaker of this language; they are largest ethnic group in Ethiopia. According to the Ethiopia Population Census Commission 2007 the total population of Oromo People which use AO language is 34.5% of the total population of the country [2]. According to paper [3] it is also the third most widely spoken language in Africa next to Arabic and Hausa languages. Specially, it is first widely spoken and used in most parts of Ethiopia (Oromia state) and also some parts of other neighbor countries like Kenya, Tanzania, Djibouti, Sudan and Somalia. Oromia broad Casting Network (OBN) is prominent official broad castings which use Afaan Oromo language on different social media like Twitter page, Facebook page, forum discussion and etc. Currently, OBN have several branches like OBN of Adama, OBN of Finfinne, Horn of Africa OBN and international OBN. So, the amounts of Afaan Oromo text on OBN Twitter pages are increasing in day to day. Therefore, we have been proposed sentiment analysis system to analyzed OBN Tweets of Afaan Oromo text. However, there are a few works studied on sentiment analysis system for Afaan Oromo text. Sentiment analysis is not a new field and it has been researched for different natural language like English and other language like Arabic, Chinese, Bangla and Amharic [9] [10] [13] while for Afaan Oromo and other under-resourced language, this field is new and it is at the initial state.

Sentiment analysis can be conducted at different levels such as document level, aspect level and sentence level [1] [10]. The document level deals with classify whether a whole opinion document expresses a positive or negative sentiment [1]. This level is analysis opinions in document on single entity (e.g., single product, single reviews). Aspect based sentiment analysis level is an aspect/feature level in which identifying and extracting the feature of the objects that have been commented on by the opinion holder and determining whether the opinion is positive or negative. Shortly, the idea of aspect based sentiment analysis is not only classify opinions into positive and negative, it also specifies the entity comment was commented on. The other is sentence level sentiment analysis which is determining whether each sentence expressed under defined polarity. Sentence level sentiment analysis is that sentence usually contains single opinion (although not true in any case) in a document typically contains multiple opinions [1]. Document level analysis is analyzed on general that means it is not identify specify area of document part. For this problem sentence level is analyzed each sentence of document into expressed emotion classes. As an example:

“I have registered for leaning data mining program. This program have amazing course. It has very good handout. But, it is very expensive to handout”.

First we analyzed this paragraph as sentence level according to the following: The first sentence expresses no opinion, it doesn’t offer any sentiment and it is neutral. Sentence two and three are expresses a positive opinion about the program and hangout of program as whole. The last sentence is expressing a negative opinion about a pricing. Generally, we extract four polarities of opinion classes. But, according to document level, this paragraph is analyzed into majority of polarity in a paragraph which is positive.

This research work extracts emotion by using supervised machine learning techniques. These two techniques are Support Vector Machine and Random Forest. We used hand created dataset with 1810 Afaan Oromo text for training and testing purpose from OBN Twitter page. The result was measured using performance metrics such as precision, recall, f-score and accuracy for evaluating the effectiveness of the proposed techniques.

The rest of this thesis is organized as follows. In section 2, some previous works related to our work has been described. Section 3 gives a simple view of the overall proposed model. Section 4 presented the experimental results of the proposed system. Finally, Section 5 concluded the thesis and provides some directions for future works.

Related Works

Recently, some sentiment analysis researches were studied regarding Afaan Oromo text. As far as researchers know four researches have been proposed within various feature and techniques. The first paper [4] Proposed Aspect based summarization of Afaan Oromo news text on the news domain. The researchers used lexicon-based approach and rule to specify the opinions from ORTO news service dataset. They have trained the system with corpus of size 400 reviews and achieved good results. Due to un-availability of lexical database and linguistic resources such as POS, feeling on social media indirectly cannot handle in their system. [5] Developed unsupervised Opinion Mining Approach for Afaan Oromo Sentiments. The researchers used lexicons, POS and N-gram as feature to identify the opinions from social medial OPDO official Facebook page. They have trained the system with corpus of size 600 reviews. [6] Designed Sentiment Analysis for Afaan Oromo Using Deep Learning Approach. They focused on investigating Convolutional

Neural Network and Long Short Term Memory deep learning approaches for the development of sentiment analysis of Afaan Oromo social media content such as Facebook posts comments.

They system has been trained and tested on dataset size 1452 comments that were collected manually from the official site of the Facebook page of Oromo Democratic Party/ODP. They system has been analyzed document level of Afaan Oromo review. They used preprocessing such as normalization, tokenization and stop word removal on manually collected dataset. They used word embedding as feature to specify opinions and the result showed that Convolutional Neural Network achieved the accuracy of 89% as well as The Long Short Memory achieved accuracy of 87.6%. [11] Studied on Constructing Sentiment Mining Model for Opinionated Afaan Oromo Texts on Ethiopian Politics. They used two machine learning techniques such as Decision Tree (DT) and Naïve Bayes (NB). Their corpus size 1800 reviews of Afaan Oromo text Ethiopian politics collected from different blogs and websites. They used tokenization, normalization, stop words removal and stemming in text preprocessing and bag of word in unigram as feature. Lastly, their experimental result showed that Naïve Bayes algorithm (88%) performed better than Decision Tree algorithm (83%) in polarity classification.

In addition to previous work of AO sentiment analysis, there were developed for others language. From different worked on other language, we reviewed three papers from other language. We select these three papers based on their working multiple classes. [9] Investigated a machine learning approach to multi-scale sentiment analysis on Amharic online posts. They usedNaïve Bayes machine learning algorithm and used unigram, bigram and hybrid variants as features. They multi-scale were consisted -2,-1, 0, +1 and +2. They have corpus size 608 posts that collected from Facebook, Twitter, Dire Tube and Ethiopian Reporter websites online sources. Lastly, their model achieved a performance accuracy of 43.6%, 44.3% and 39.5% for unigram, bigram and hybrid language models respectively. [10] Developed an Automated System of Sentiment Analysis from Bangla Text using Supervised Learning Techniques. These Supervised Learning Techniques- Naïve Bayes Classification Algorithm and Topical approach to extract the emotion from any Bangla text. They studied sentiment analysis at sentence level and document level. They classified sentiment with six individual emotion classes- happy, sad, tender, excited, angry and scared. They created manually data corpus with 75000 sentences and 7400 sentences used for training dataset as well as 100 sentences used for test data. Their model achieved highest in topical approach 90% accuracy. [7] Designed KNN classifier based approach for multi-class sentiment analysis of twitter data. They were done sentiment classification into multiple classes. They proposed methodology KNN were compared with other approach SVM. They extracted Twitter data by using python Tweepy. They used n-gram as feature and achieved 86% accuracy.

In the regarding to Afaan Oromo text sentiment analysis, researches [4], [6] and [11] have been proposed on Afaan Oromo review sentiment analysis on binary classification (positive and negative) and [5] add neutral on this class. They also studied on document and aspect level sentiment analysis that means they didn't studied on sentence level. Today's world focused on multi-label classes of sentiment analysis rather than binary classes such as: negative and positive. However, to the best knowledge of the researcher, there is no investigated on multi-class sentiment analysis and also, there is no researched sentiment analysis at sentence level for Afaan Oromo text. Therefore, in this study, we proposed on investigating sentence level sentiment analysis for Afaan Oromo tweets based on supervised machine approaches. The aim of this study was classifying the polarity a given text whether the expressed emotion of Afaan Oromo tweets into five major classes such as: very negative, negative, neutral, positive and very positive based on trained dataset.

PROPOSED MODEL

In this section, to investigate an automate system for classifying opinions from Afaan Oromo sentence; we applied two supervised machine leaning approaches: Support vector machine and Random forest. These approaches have been proposed for classifying opinions from Afaan Oromo text at sentence level into five multiple sentiment classes.

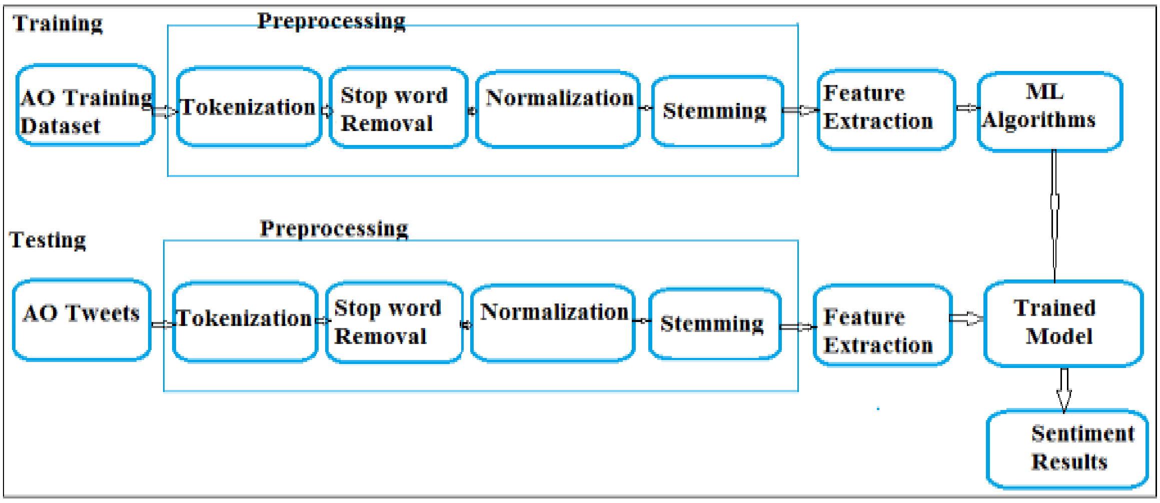

Figure1. *Proposed Model Architecture*

To conduct the procedure with a better performance, the process working of this study has been consisted the following sub sections.

Data Collection And Preparation

For this study, primary data sources have been taken from Ethiopian governmental broad casting cooperation, Oromia broad casting network (OBN) twitter page. This page have a huge user that give different opinion for improving their life through politics, culture, business, sport and some others topics which released every day by OBN the people of sound. In the activity of dataset collection step, we got above 10,000 un-labeled tweets text from OBN. Therefore, we acknowledged OBN for getting these huge amounts of tweet dataset without payment. For this study, we used supervised machine learning approaches. Since supervised machine learning approaches needed labeled dataset for training purposes. Therefore, we labeled dataset within discussion of experts. But, these 10,000 un-labeled tweets were contained different un-opinion text as well as other mixed language text. We take more time with experts of language to clarify Afaan Oromo text and labeled 1810 dataset into multiple classes-positive, negative, neutral, very negative and very positive for training phase. These classes categorized depend on the power of emotion word in sentences. For example the tweets which have high power to sense the reader, we categorized under very positive or very negative rather than positive or negative. There is also the repetition of adjective, adverb and some others predicate of sentence give high sense emotion to say very positive and very negative. Otherwise, when the tweets are no opinion sentiment and the sensitive of negative or very negative and positive or very positive are equal, it is neutral. For this assumption, there is an example according to the following table1. The bold sentence is opinion in Afaan Oromo language and normal sentence is meaning in English.

Table1. *An Example of dataset style*

No	Tweets	Polarity
1	**QophiisirnaGadaakeessanbaayyeebayyeegaariidha.** (Your Gadaa System program is very very good.)	Very Positive
2	**OBN gidduugidduusagantaakeessaanittiosoosirbootaHaacaaluuHundeessaa nu affeertaniigaariidha.** (If OBN can invite Hacalu Hundessa voice in between programs, it is good for us.)	Positive
3	**Qophiinqe'eeqorannoodansaamiti.** (The program studying about environment is not good.)	Negative
4	**OBN Seenaa fi DudhaaOromooirrattisirrittihojjechaahinjiru.** (OBN is not working perfectly on history and culture of Oromo.)	Very Negative
5	**Magaalaanmuummeeoromiyaafinfinneedha.** (The Capital city of Oromia is Finfinne.)	Neutral

From 1810 dataset size, 286 very negative, 476 negative, 351 neutral, 401 positive and 296 is very positive opinion text.

Text Preprocessing

Another step is text preprocessing in order to remove un-necessary part for opinion mining from dataset. Text pre-processing is the process of applying any type of computation on shapeless raw data and transforms it into an arrangement that more easily and effectively processed in another procedure. The architecture of proposed model in this work has been consisted two phases, training and testing. In the training phase, the machine learning needed to learn from a set of annotated AO tweets and it was used to classify un-labeled tweets in testing phase. These two phases have the following preprocessing steps: tokenization, stop word removal, normalization and stemming.

Tokenization is the first step in preprocessing of both phases used to tokenize a sequence of text or document into sentences and changes sentences into word. In this study, we proposed sentence level sentiment analysis for Afaan Oromo tweets text. Since sentence level needed tokenized document text into sentence. Document tokenization into sentence in Afaan Oromo is like English language based on period (.), question mark (?) and exclamatory mark (!). After this, sentence is tokenized into word on space between sequences of each word in sentences for easy to apply the next step stop word removal.

Stop word removal is also another step of preprocessing used to remove most frequently word in text that are not relevant or have no impact to classify sentiments. In this research, some Afaan Oromo stop words are significant for the sentiment classification and need to remain in the text [6]. As an example "*hin*" stop word of Afaan Oromo that influence the sentiment of classification. For example,

1. OBN Seenaa fi DudhaaOromooirratti**sirritti** hojjechaa**hin**jiru. "OBN is **not** working **perfectly** on history and culture of Oromo".
2. OBN Seenaa fi Dudhaa Oromooirratti**sirritti** hojjechaajira. "OBN is working **perfectly** on history and culture of Oromo".

In the first sentences, "***hin***" stop word is influencing opinion into negative emotion and "***sirritti***" (**perfectly)** is adverb that comes for more explaining the negativity of OBN. This perfectly word is also having another significant which classify sentiment form normal into very negative or very positive. So for this study, we filtered removed stop words through a manual process that is not relevant for the classification process.

Normalization: Many people use different character form to write the similar word and they can use uppercase, lower case or mix of the two in similar word. They also used elongated texts and numbers into explain their emotion. Therefore, normalization need for improving performance of the system. In this work normalized according to the following rules:

The word "baay'ee" and "baayyee" to mean many has the same meaning with different writing. The only difference is that the apostrophe "'" is replaced by "y".

The word "Bilisummaa" to mean Freedom is written as "Bilisummaa" or "BILISUMMAA" at the beginning of the sentence and it is written as "bilisummaa" at the middle of a sentence while it has the same meaning at both locations in the sentence. We normalized lower letter which most common in a text.

The other is emotion in numbers. For example: "100% sin jaallanna" to mean we love 100% normalized into "persentiidhibbatokko sin jaallanna" to mean: we love one hundred percent.

There are also elongated texts, for example, OBN sin jaallannaaaaaaaaa to mean: we loveeeeeeeee OBN is normalized to OBN sin jaallanna to mean we love OBN.

Stemming is a component that reduces morphological variants of words into root form. In morphologically complex languages like Afaan Oromo, a stemmer will lead to important improvements in opinion mining systems [4]. For example in Afaanoromo, "afeeramuu", "afeerraa", "afeeramtani", "afeeraman", "afeeramaniiru", all are rooted to "afeer" . For our case, we used Afaan Oromo stemmer algorithms developed by [12].

Feature Extraction

Feature extraction is the process of plotting from textual data to real-valued vectors. In our case, after tweets cleaned from stop words, normalized and stemmed, we used Tf-idf as feature extraction. Tf-idf (term frequency- inverse document frequency) is a common feature extraction method in text mining. We used Tf-idfvectorizer from scikit-learn feature extraction tools.

The Methodology of Proposed Approaches

Support Vector Machine

Support vector machine (SVM) is a type of supervised machine learning classification algorithm which can be used for both binary label classification and multi-label classification. SVM is originally formulated for binary classification [14]. For example, in the paper [13], they used SVM for sentiment analysis into two classes such as positive and negative class. However, in our case, we have been proposed multi-class sentiment analysis. For multi class problem, support vector machine can solve multi classification problem using one-against-one or one-against-rest method [14]. In the one-against-one approach, instead of trying to distinguish one class from all the others, they seek to distinguish one class from another one. In the one-against-rest binary of problem, SVM is trained for each class in order to distinguish that class and the rest.

Support vector machine (SVM) has many classes library that is capable of performing binary classification and multi classification such as: SVC, NuSVC and linearSVC library. SVC and NuSVC are similar methods and use kernel parameters. On the other hand, linear SVC is another implementation of support vector classification for the case of a linear kernel, but linear SVC does not accept keyword kernel. In the multi-class classification process, SVC and NuSVC implement the "one-against-one" approach. If K is the number of classes, then K(K-1)/2 classifiers are constructed and each one trains data from two classes. To provide a consistent interface with other classifiers, the decision function shape option allows transforming the results of the "one-against-one" classifiers to a decision function of shape (n_sample, K). Linear SVC automatically uses the one-against-all strategy by default. It can also specify it explicitly by setting the multiclass parameter to ovr (one-vs-the-rest). Theoretically, the term 'One-against-one 'is an approach used to classify one class from others classes. In this approach, instead of trying to distinguish one class from all the others, it seeks to distinguish one class from another one. As a result, the train one classifier per pair of classes, which leads to K(K-1)/2 classifiers for K classes. Each classifier is trained on a subset of the data and produces its own decision boundary. In this work, we used SVC library. We select this library form other support vector machine based on achieved highest performance.

Random Forest Classifier

Random forest is a popular supervised machine learning used for classification and regression. A random forest is an ensemble of tree structured classifiers and every tree of the forest gives a unit vote, assigning each input to the most probable class label [8]. Random forest is a tree based machine learning algorithm that leverages the power of multiple decision trees for making decisions. Random forest is a forest of trees. The main advantage of random forest is that it does not suffer from over fitting, even if more trees are appended to the forest [8]. Random forest classifier uses many of the parameters as Decision tree classifier. However, being a forest rather than an individual decision tree, Random forest classifier has certain parameters that are either unique to random forests or particularly important [15].

Experimental Results

In this paper, our experiment was conducted by used supervised machine learning methods on OBN twitter annotated dataset. These methods are Support vector machine and Random forest classifier. For Support vector machine, we used ***SVC*** library from scikit-learn and for Random forest, we used ***Random Forest Classifier*** library from scikit-learn of python libraries by using Anaconda programming. This SVC library is with kernel parameter. The term kernel is a support vector machine parameter which is to take data as input and transform it into the required form. For user interface build, we used tkinter library from python libraries.For example, when an input Tweets "***Qophiisirna Gadaakeessanbaayyeebayyeegaariidha.***"/'Your Gadaa System program is very very good.' entered and outputted the system was captured the following contexts as in figure 2 below. After an inserted tweet, system was applied preprocessing steps and the proposed models were predicted the tweet categories based on trained dataset.

Figure1. *Sample Input Tweet opinion and the Polarity Value*

The performance of the proposed system is analyzed on the basis of analysis metrics scikit-learn namely recall, precision and accuracy. These methods were evaluated using the prepared testing dataset that 20% of dataset and the left is training dataset. We reported the evaluated proposed system results according to the following captured of python code.

Support Vector Machine Model

```
x_train1, x_test1, y_train1, y_test1=train_test_split(x1, y1, test_size=0.2,random_state=0)
polarity_model = SVC(kernel="rbf", random_state=0, gamma=1, C=1)
polarity_model.fit(x_train1, y_train1)
prediction = polarity_model.predict(x_test1)
evaluation=classification_report(prediction,y_test1)
print(evaluation)
print("Accuracy=",accuracy_score(prediction,y_test1))
```

```
              precision    recall  f1-score   support

    Negative       0.91      0.77      0.84       109
     Neutral       0.93      0.99      0.96        68
    Positive       0.98      0.98      0.98        92
VeryNegative       0.62      0.81      0.70        42
VeryPositive       1.00      1.00      1.00        51

    accuracy                           0.90       362
   macro avg       0.89      0.91      0.89       362
weighted avg       0.91      0.90      0.90       362

Accuracy= 0.9005524861878453
```

Random Forest Model

```
polarity_model2 = RandomForestClassifier(random_state=0, n_jobs=-1)
polarity_model2.fit(x_train1, y_train1)
prediction2 = polarity_model2.predict(x_test1)
evaluation2=classification_report(prediction2,y_test1)
print(evaluation2)
print("Accuracy=",accuracy_score(prediction2,y_test1))
```

```
              precision    recall  f1-score   support

    Negative       0.88      0.78      0.83       104
     Neutral       0.93      0.97      0.95        69
    Positive       0.97      0.97      0.97        92
VeryNegative       0.64      0.76      0.69        46
VeryPositive       1.00      1.00      1.00        51

    accuracy                           0.89       362
   macro avg       0.88      0.90      0.89       362
weighted avg       0.90      0.89      0.89       362

Accuracy= 0.8922651933701657
```

In our final experiment, we have compared our study with other papers regarding sentiment analysis from Ethiopian language, Afaan Oromo text. All of the previous researchers were analyzed sentiment at document level. The numbers of sentiment classes analyzed in their research were not more than three classes and they used small dataset. Beside this, our dataset is some more than the other research papers and we analyzed at sentence as well as we used multiple classes.

Table2. *Comparative Analysis with other papers for Afaan Oromo text sentiment analysis*

Paper Title	Features	Methods	Classes	Dataset	Accuracy
Aspect based summarization of Afaan Oromo news text[4]	lexicons and rule based	Rule based	2 (Positive and Negative)	ORTO news service dataset (400 corpus size)	88.5% for Pos 88.3% for Neg
Opinion Mining for Afaan Oromo [5]	lexicons, POS and N-gram	Unsupervised	3 (Positive, Negative and Neutral)	OPDO official Facebook page (600 corpus size)	Good in bigram and low in tri-gram
Sentiment Analysis for Afaan Oromo [6]	Word embedding	Convolutional Neural Network and Long Short Term Memory	2 (Positive and Negative)	Facebook page of ODP (1452 corpus size)	LSTM: 89% CNN: 87.6%
Constructing Sentiment Mining Model for Opinionated Afaan Oromo Texts[11]	Bag of word (unigram)	Decision Tree and Naïve Bayes.	2 (Positive and Negative)	Afaan Oromo text Ethiopian politics (1800 corpus size)	NB: 88% DT: 83%
Our paper	TF-IDF	Support vector machine and Random forest	5(Very Negative, Negative, Neutral, Positive and Very Positive)	OBN Twitter (1810 corpus size)	SVM: 90% RF: 89%

Conclusion and Future work

In this paper, we analyzed the performance of supervised machine learning approaches to classify multiple class sentence level sentiment analysis from Ethiopian language, Afaan Oromo text. These approaches are Support vector machine (SVM) and Random forest (RF). The dataset were used to train and test a sentiment classifier based on SVM and RF uses tokenization, stop word removal, normalization and stemming as preprocessing. Then, we used Tf-idf (term frequency- inverse document frequency) as feature extraction as input to machine learning algorithms. The experiment results that proposed approaches SVM performed accuracy 90% and RF achieved an accuracy of 89% on OBN Twitter dataset with 1810 corpus size. But, due to un-availability of standard sentiment dataset of Afaan Oromo text on social media, still there exist some limitations. Therefore, in the future work, we forward the following points on sentiment analysis regarding to Afaan Oromo text. One of the great challenges in sentiment and opinion mining is having Afaan Oromo standard dataset. So, future research should consider creating standard dataset within large size that can be used for future research in this area. It is also improving performance by developing a hybrid approach mechanism using these two methods and will also implement others approaches of sentiment analysis. The other research direction to be considered is going for Aspect based sentiment analysis regarding Afaan Oromo text. The previous researchers were considered on document level, feature level and our paper focused on sentence level. Therefore, we forward research direction of combination document and sentence level sentence analysis.

Reference

[1] Bing Liu, “Sentiment Analysis and Opinion Mining”, Morgan & Claypool Publishers, May 2012.

[2] Summary and Statistical Report of the 2007 Housing Census: Population Size by Age and Sex, Addis Ababa, 2008.

[3] Tabor Wami's, new book titled, Yewugena Dirsetochena Yetarik Ewunetawoch, 2004.

[4] Wegderes Tariku. “Sentiment Mining and Aspect Based Summarization of Opinionated Afaan Oromo News Text”, Debre Berhan University, unpublished, 2017.

[5] Jemal Abate. “Unsupervised Opinion Mining Approach for Afaan Oromoo Sentiments”, Haramaya University, 2019.

[6] Megersa, Oljira. “Sentiment Analysis of Afaan Oromoo Facebook Media Using Deep Learning Approach”, Institute of Technology, Ambo University, 2020.

[7] Soudamini Hota and Sudhir Pathak. “KNN classifier based approach for multi-class sentiment analysis of twitter data”, International journal of Engineering and Technology, Chandigarh University, Mohali, India.

[8] Breiman Leo, Random forests- Random features,Statistics Department University of California Berkeley, CA 94720

[9] Wondwossen Philemon and Wondwossen Mulugeta, "A Machine Learning Approach to Multi-Scale Sentiment Analysis online posts".

[10] Rashedul. Amin. et. “An Automated System of Sentiment Analysis from Bangla Text using Supervised Learning Techniques”IEEE 4th International Conference on Computer and Communication Systems, 2019.

[11] Dursa Abduselam. “Constructing sentiment mining model for opinionated Afaan Oromo texts on Ethiopian politics” Gondar University, unpublished 2017.

[12] Debela Tesfaye and Ermias Abebe Designing a rule based Stemmer for Afaan Oromo Text.

[13] Nurulhuda Zainuddin and Ali Selamat. “Sentiment Analysis Using Support Vector Machine”, IEEE 2014 International Conference on Computer, Communication, and Control Technology, 2014

[14] Chih-Wei Hsu and Chih-Jen Lin,” A Comparison of Methods for Multi-class support vector Machines”.

[15] Chris Albon, Machine learning with python Cookbook, Published by O’Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, 2018.

Citation: *Negessa Wayessa*, Sadik Abas, “Multi-Class Sentiment Analysis from Afaan Oromo Text Based On Supervised Machine Learning Approaches” International Journal of Research Studies in Science, Engineering and Technology, 7(7), 2020, pp. 10-18.*