DEEP LEARNING APPROACH FOR AMHARIC SENTIMENT ANALYSIS

UNIVERSITY OF GONDAR

Yeshiwas Getachew(Msc) **Abebe Alemu (Assistance professor)**

Faculty of informatics **Faculty of informatics**

Department of IT **Department of computer science**

yeshiwasg@gmail.com **Abecom21@gmail.com**

Abstract

Deep learning has emerged as a powerful machine learning technique that learns multiple layers of representations or features of the data and produces state-of-the-art prediction results. Along with the success of deep learning in many other application domains, deep learning is also popularly used in sentiment analysis in recent years.

Social media is now playing a vital role in influencing people's sentiment in favors or against a government or an organization. Therefore, to understand the sentiment of any posting in social media an efficient method is an ultimate necessity. We have analyzed some Facebook postings to understand socio-political sentiments. Among this broad scope we have done on this research applying the state of the art in sentiment analysis on Amharic language using deep learning approach in socio-political domain. The further preparation of the dataset on the other domain enhance the language so on this research try to show the data extracted from Fana broadcasting corporation official Facebook page using Graph Application interface of Facebook social media on immigration, war and public relation issues and prepare the data for further preprocessing.

After collecting thus data from the FBC using post_id all preprocessing steps tokenization, stop word removal, stemming of the sentence are undertaken. The manual annotation of the sentence extracted data contain both the text file and Emoji are annotated using linguistic experts in seven class, positive, very positive, extremely positive, neutral, negative, very negative and extremely negative class by considering the effect of most common Emoji. Using the Scikit-learn feature extraction classes, Count Vectorizer and TF-IDF Vectorizer the researcher build the feature extraction method to use our custom vocabulary. We test both feature extractors and find that our model performs better with Count Victories as it offers a simple representation of our data.

To evaluate the performances of the systems we have collected 1600 reviews from immigration, war and public relation domains. Then we evaluate our system training and validation accuracy using three experiments by changing training and testing split 90%,10%:80%,20%:70%,30%, the size of the dataset, the number of epochs and network layers. Accordingly, on the first experiment register 90.1 % average training accuracy and 90.1% average validation accuracy performances were obtained by the first method. The second method achieves an average training of 82.4% and an average validation accuracy of 40% performances obtained. The third experiment conducted by increasing the number of data set 1600 and five network layers we get 70.1 training accuracy and 40.1 validation accuracy. The results show that this study is promising.
Keywords: *deep learning, sentence level sentiment analysis, Amharic social media sentiment*

I. Introduction

Sentiment analysis refers to the management of sentiments, opinions, and subjective text [1]. Sentiment analysis provide the comprehension information related to public views, as it analyzes different tweets and reviews. It is a verified tool for the prediction of many significant events such as box office performance of movies and general elections [2]. Public reviews are used to evaluate a certain entity, i.e., person, product or location and might be found on different websites like Amazon. The opinions can be categorized into negative, positive or neutral. The purpose of sentiment analysis is to automatically determine the expressive direction of user reviews [3]. The demand of sentiment analysis is raised due to increase requirement of analyzing and structuring hidden information which comes from the social media in the form unstructured data [4].

Recently social media platforms have become a main medium for people to express their daily activities, reactions and emotions. Blogs and micro blogs are the most common form of social media [5]. Blogs are informal sites on the worldwide web where users are used to post ideas, discussions, thoughts on a particular issue [6]. While microblogs are smaller blogs with short posts up to a limited number of signs [7]. One thing they have in common is that they consist of entries that are listed in a chronologically descending order (i.e. the latest news is on top). They are tools that enable discussions and comments on information shared with other users. They are characterized by its dynamic and up to datedness.

About 16,037,811 Internet users on June/2017, 15.4% of the Ethiopian population per ITU 4,500,000 Facebook users on June/2017, 4.3% penetration rate[8].Among this cyber users citizen most of them participating on different issues and give feedback towards the issue raised by the his/her friends, blogs, medias ,newspapers by mother tongue language. However still now as my concern there is no work done on sentiment mining for socio-poltical domain based on opinioned Amharic text using deep learning approach.

Previous opinion mining techniques and sentiment mining models are often developed for Amharic language in movie review and product service areas, the previous researchers Abebe [9], Mengistu [10] and Abraham[11] followed a supervised machine learning approach which employs Naïve Bayes Decision Tree, Support Vector machine, Multinomial Naïve Bayes and Maximum Entropy algorithms using Bag of words, Information gain, n-grams presence, n-grams frequency & n-grams-TF-IDF features for document level Amharic sentiment classification. In spite of thus researchers our research focus on sociopolitical domain and we collect the data set from the official Facebook page of Fan broadcasting by using Facebook data extraction tool known as Facepager using graph API. In addition to this we follow a deep learning approach to train and test the model which is novel method for the area in our language.

One of the current trend in social media people express their felling towards publicly posted issues using emotion icon(Emojis). Thus, emotion icon associated with Amharic text express different sentiment of the people about issues raised by the online community. The researcher come up with novel idea which is incorporation of emotion icons with Amharic comments in sentiment analysis.

In terms of data required to train and test the model the deep learning efforts to learn high level abstraction by exploiting the hierarchical architectures. It is a promising approach and has been extensively applied learning, semantic parsing, natural processing and many more. We select deep learning approach because of improved ability of chip processing (GPU unit), extensive lower expenditure of hardware and significant enhancement in machine learning algorithm.

Research Questions

Based on the above research problem, the researcher raised the following research questions: -

- What are the best features extracted from opinionated Amharic texts that have associated with socio ploltical data?
- What preprocessing tools and procedures applied to come up with quality data sets for training and testing?
- Which deep learning approch produce a better classification model for mining the domain area?
- Could the inclusion of emotion icon effect on Amharic sentiment analysis?

In this paper, Section I covers introduction and research gap, Section II covers the detailed literature review, Section III Methodology we follow presents, Section IV the experimental result, lastly in section V shows the conclusion part are presented.

II. Literature review

In Literature Review section, various studies of sentiment analysis both on local language and other language using Deep Learning techniques are discussed.

The current state of the art on sentiment analysis and clearly define what are the gap of Amharic opinion mining and sentiment analysis we done extensive review on different thesis. Sentiment analysis and opinion mining model done for Amharic language using machine learning and lexicon-based approach by six researchers thus research papers [9] [10] [11] [12] [13] [14] summarized in table 1 are discussed in detail. Moreover related works which is done to opinion mining work done on English and other languages thus are [15] [16] [17] [18] 19].

Table 1 summarized Amharic sentiment analysis and opinion mining literature

Author	Objective	Features	Model	Classes	Selected Model	Domain, Amharic (data source)	Num dataset
Selama, (2010), MSc Thesis (AAU)	Assign docs sentiment	lexicon + Context valence shifter	Rule & Lexicon Based	+ve, -ve & Neutral	General lexicon with valence shifter	Movie review + news doc	303
Abebe, (2013), MSc Thesis (UoG)	Assign docs sentiment	BoW, IG	NB and DT	+ve & -ve	Naïve Bayas with Information Gain	Movie review+news	456
Mengistu, (2013) MSc Thesis (AAU)	Assign docs sentiment	n-grams presence, n-grams frequency & n-grams-TF-IDF	NB, MNB & SVM	+ve, -ve & Neutral	Support Vector Machine with unigram	ERTA.com + FanaBC.com + diretube.com	576
Tulu, (2014) Journal Article	Assign feature sentiment	Feature + Opinion + Context valance shifter	Rule & Lexicon Based	+ve, -ve & N/A	Feature with adjacent left & right adjective	Hotel, University & Hospital	484
Abreham, (2014) MSc Thesis (AAU)	Assign docs sentiment	BoW & IG	NB, DT & ME	+ve & -ve	Naïve Bayas with Information Gain	EBC, diretube.com & habesha.com	616
Wondwossen (2014) Journal Article	Assign docs sentiment	Unigram, Bigram & Hybrid	NB	-2, -1, 0, +1, +2 (multi-scale)	Naïve Bayas with Bigram	Social media, Product marketing & News	608

III. Methodology

The researcher follows Empirical research design because of the nature of both people opinion have been subjective towards the topic. the annotation of extracted data need experts and experimental setup needs to validate the model. Therefore, the Sentiment analysis on this research has the potential of employing Hybrid techniques.

Deep learning

As our data is numerical an ANN have created a vector product of each input with a randomly generated number often in the range (-1, 1) along with a bias value which are feed into a summation or activation function to determine an output value. A bias value is used in the event of all data points are equal to zero which means there would be no value to multiply with weights. Therefore, the neuron would not generate an output. The function f represents the activation function used to determine a non-linear output of the neuron and can use mathematical functions such as in following which are also visualized in Figure 4.1 below:

Sigmoid: takes a real-valued input and squashes it to range between 0 and 1

$\sigma(x) = 1 / (1 + \exp(-x))$

tanh: takes a real-valued input and squashes it to the range [-1, 1]

$\tanh(x) = 2\sigma(2x) - 1$

ReLU: ReLU stands for Rectified Linear Unit. It takes a real-valued input and thresholds it at zero (replaces negative values with zero) $f(x) = \max(0, x)$

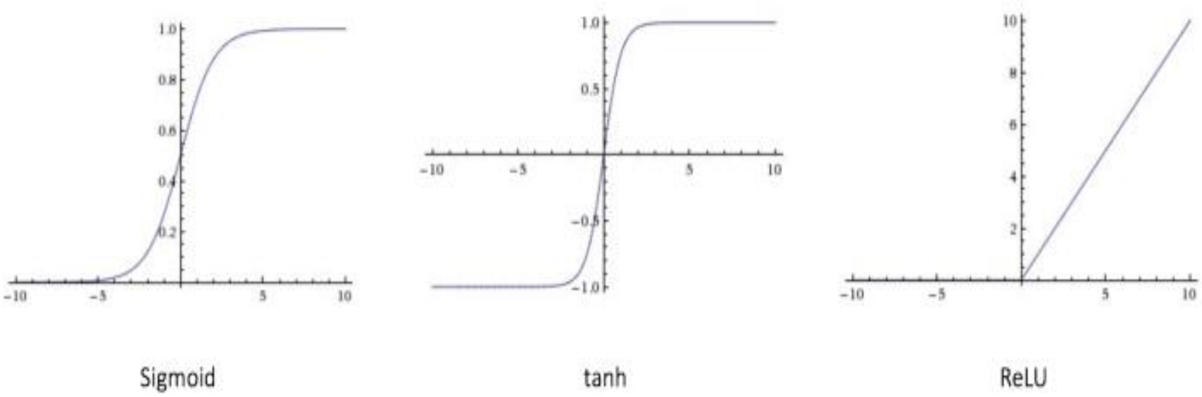

Figure 1 Activation functions [20]

An ANN can comprise of many different neurons each with different activation functions depending on the context of the problem. In our case we expect our final layer the output layer to use a Sigmoid activation function as our class labels would be in the range (0, 1). Functions such as tanh and reLU would be useful in the input and hidden layers of a deep network in order to improve the network's performance as it grows more complex and depending on the size of our input data.

A multi-layer network contains many neurons, and depending on the network structure, alters the weights through forward and back-prorogation, using the labeled training data to determine the best model. Multi-layer networks are useful for non-linear boundaries and the outputs of neurons in earlier or hidden layers, feed into later layers to determine a classification [21]. The neural network training process and error handling through methods such as back-propagation are computationally more expensive than the NB or SVM but are also more powerful and can achieve better results in more than one context particularly for multi-classification problems [20].
A simple multi-layer preceptor or deep network can be visualized in Figure 2 below:

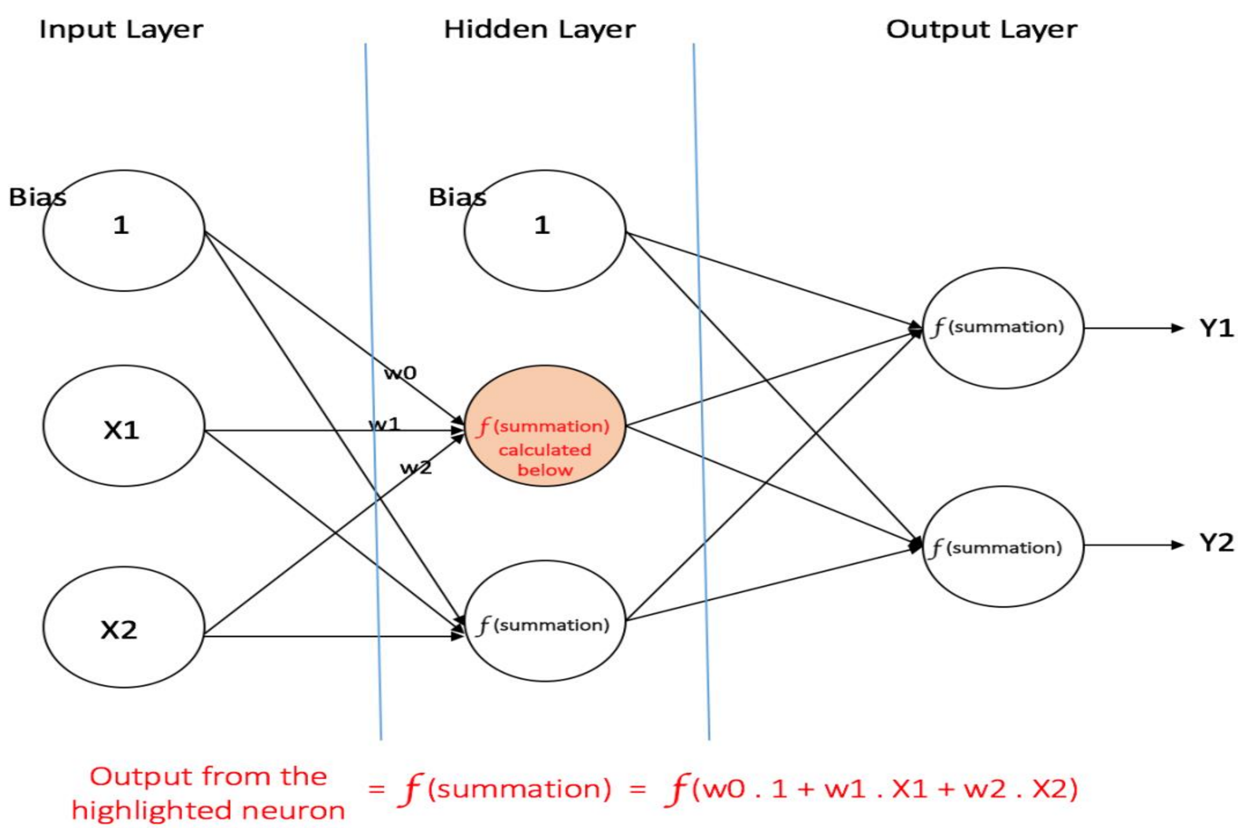

Figure 2 a simple Deep network, adopted [54]

Input features, labeled X1 and X2, represent data in numerical terms and as we mentioned earlier they are associated with a weight value and passed into an activation function. A single or multiple hidden layer can contain any number of neurons or nodes which take the previous layer's activation function result and pass it forward for further calculation which is subsequently used in the output layer to determine a final value. The number of features in our system would be the size of the dictionary file or vocabulary Xn, where n is the number of words in the file and each value of Xn-1 of our inputs would represent a one-hot encoded value of our raw text data. Finally, the output layer would Consist of the number of class labels in our annotated data ranging in (0,1).

An ANN is trained through the processes of forward and back-propagation. Forward propagation is the process of applying the dot product of our input features with randomly assigned weights and passed into an activation function, where the output can be represented as the following:

$Y = f\,(x1*w1 + x2*w2 + \ldots + xn\text{-}1 *wn\text{-}1)$

Where denotes the number of features of our data f as the activation function and Y as the output result. These output values are passed into deeper layers and eventually to the output layer which compares the predicted result against the training data to determine whether the model has correctly predicted the outcome or not. Often forward propagation returns incorrect results and therefore requires further processing methods to help the network learn perform better, which can be achieved by updating the weights along with back-propagation.

We calculate the total error at the output nodes and propagate these errors back through the network using Back propagation to calculate the gradients. Then we use an optimization method such as Gradient Descent to adjust all weights in the network with an aim of reducing the error at the output layer [20]. The name for one commonly used optimization function that adjusts weights according to the error they caused is called "gradient descent. Gradient is another word for slope in its typical form on an x-y graph represents how two variables relate to each other. In this particular case the slope we care about describes the relationship between the network's error and a single weightlike. how does the error vary as the weight is adjusted [22].

The network repeats these processes via iterations or epochs until the model has correctly learnt to classify the data. All of these factors affect the networks performance and as we discuss the design, implementation and evaluation of our model. we learn through time how each of these functions and algorithms can greatly improve or reduce the accuracy of the network and so inform the improved designs of our ANN.

There are many works using deep learning for natural language processing and most use variations of the simple neural network in order to achieve their respective goal within their relevant context. However, they share a common foundation in taking raw textual data and representing the individual words or word pairs as vectors, to be used for further processing. Word embedding is a distributed representation of a word often a one-hot representation which is suitable for the input of neural networks [23] where a word corresponds to a one-hot vector indicating a binary value denoting its presence in a document.

A convolution neural network is similar to a normal neural network in its operation but its architecture is better suited for image processing problems as the input vectors are in 3d form and can perform better than regular neural networks on a larger scale. The efficiency of a convolution neural network or CNN has inspired use cases for natural language tasks by transforming one dimensional textual data into a matrix as the input image for the network. A big argument for CNNs is that they are fast. Convolutions are a central part of computer graphics and implemented on a hardware level on GPUs.

Deep learning approach for Amharic sentiment analysis Architecture

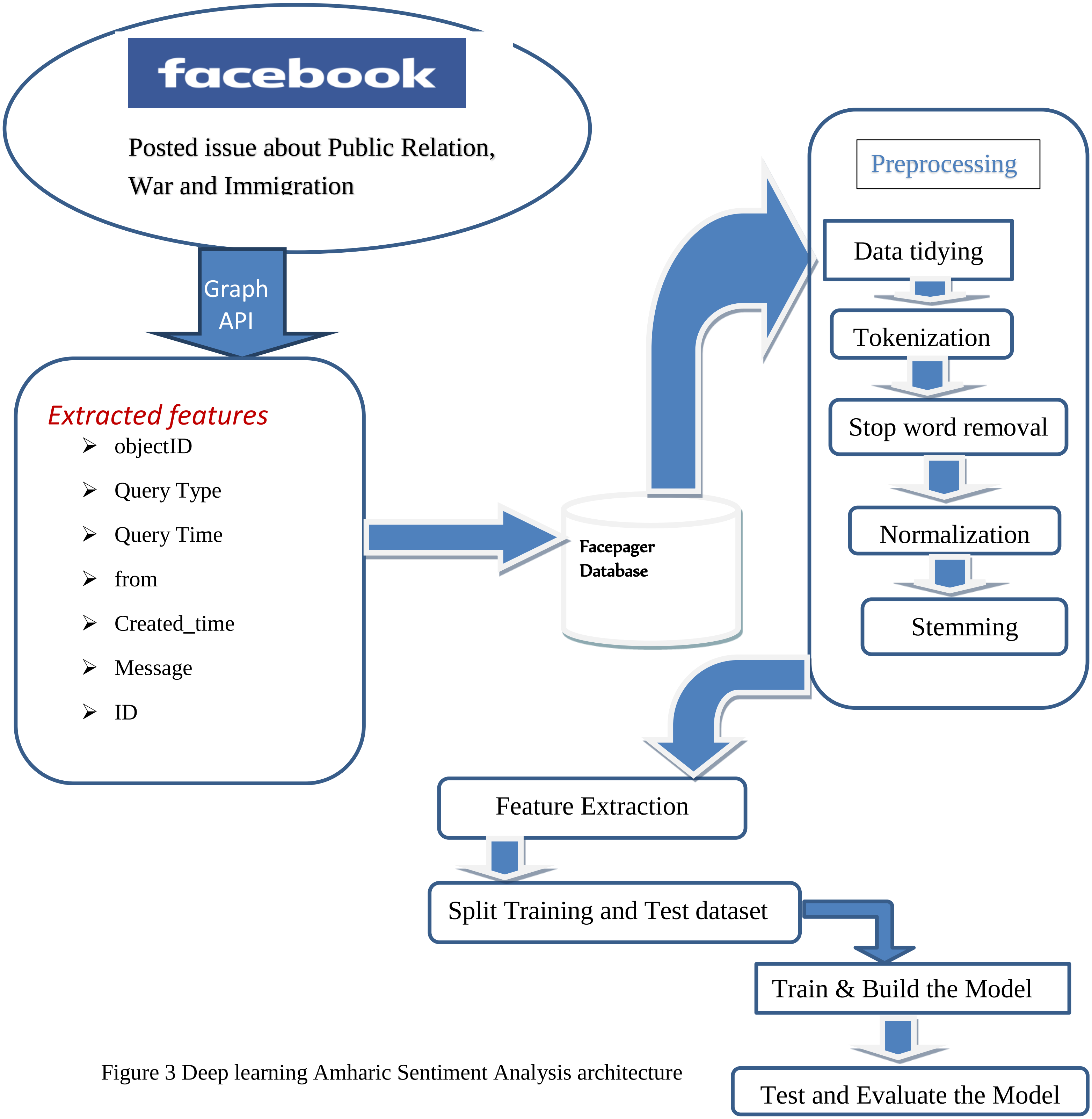

Figure 3 Deep learning Amharic Sentiment Analysis architecture

Data collection

A. Data source selection

We focus on this resarch primary data source FBC offical Facebook page because this page was legal under the Facebook company terms and condition and people express his/her idea freely on social media. Despite the researcher concentrate on the following aspect only from the broad concept of sociopoltical domain area thus are immegration data, public relation and war .

We collect the data by using Facebook API by creating developer account and scrape using the Facepager software text processing tool. The main attribute extracted during the data collection are listed on appendix A.

We use 1800 reviews collected from thus issue among that we use 1602 reviews including emojis and common 4000 vocabulary file for training and testing the model.

Table 5.1 reviews collected from Fana broadcasting corporation

Issue	Number of reviews
Immigration	652
Public relation	482
War	468

IV. Experimentation

In this subsection we discuss the experimental result and evaluation of our system architecture in an agile and test-driven development style. During each iteration or sprint as is the agile term we test our model and report the results of its performance and conclude with a reflection where we discuss the findings. The reflection serves as a means of improving the system in the next sprint by setting goals or targets to improve upon the last iteration.

Opening each epoch, we present a table of results for the average training and validation accuracy along with the size of the data set and input dimensions (vocabulary size) as well as the number of layers in the network. Following this we discuss the findings in more detail and points of critique to suggest improvements for the next iteration to achieve. During the course of each sprint, we develop and refine the network structure and discuss features added or removed in order to achieve greater performance scores.

Experiment one

In the first iteration of testing, we use a three layer, fully connected network as defined in the previous chapter. Use a data set of around 600 labeled comments, with a training split of 90%, 80%,70% and test of 30%,20,10% batch size of 10 and 50 epochs. As a first run, we use a small data set of 600 comments and a vocabulary file of size 4000 which returns an initial training accuracy of 96.11% however, which suggest a great degree of over fitting.

Experiment one result

Network layers	Train and test split	Training Score	Validation Score	Input Dimension	Data set size
3	70,30%	95.5	90.50	4000	600
3	80,20%	95.83	91.60	4000	600
3	90,10%	96.11	96.61	4000	600

Experiment Two

Upon inspection of our vocabulary file many of the words in the list were punctuation, hashtags, usernames, variations and elongations of base words. After thoroughly cleaning and reducing the list of words and emojis to the smallest number possible we remain with 1052 words and emoji that we hope can achieve maximum coverage in capturing sentiment of the whole dataset. Reflection from the previous iteration revealed that many of the words treated the hash symbol and other hash tags as words which would invariably fragment the rules we used to classify our data and so we chose to only include the exclamation and question mark as they are valuable in emphasizing sentiment.

Repeating the training shows similar results to the previous iteration, suggests that the network structure may need more layers or neurons. Researching the Keras user group reveals that changing the activation function in the second to last layer to 'tanh' helps to balance the results and reduce false positives. Another recommendation mentioned was to implement a random generator in the network, in order to be able to reliably reproduce results.

Through testing many different network configurations such as increasing layers and neurons, implementing various learning rates and weight initializations we found that a four-layer network with neurons greater than 100 return more accurate results for our data. The training accuracy drops to a more accurate range between 80-85% and validation of around 83% accuracy.

Experiment Two result

Network layers	Train and test split	Training Score	Validation Score	Input Dimension	Data set size
4	70,30%	84.05	79.89	1052	600
4	80,20%	84.03	83.54	1052	600
4	90,10%	84.63	83.05	1052	600

Experiment Three

Parallel to further annotation of the training data we spent a considerable amount of time in furthering our knowledge of deep learning and implementing methods to improve the network. Learning from resources in [24] [25] [26] [27] [28] [29][29] [30] several improvements have been made to the network. Firstly, through testing the researcher found that the network improves its accuracy with increased neurons per layer except for the final layer that remains fixed to the size of our encoded labels. Having experimented with many different configurations the best results were returned from a network structure with configurations of between 300 and 1000 neurons per layer. Two improvements have been made in this iteration a new layer has been added which also matches the number of neurons of the input later.

Our network is now a five-layer system with three hidden layers, and greater number of neurons per layer compared to that last iteration. The notable improvements arise from the second layer which is the first hidden layer to have more neurons than the input layer which then further pass their learned weights to the second hidden layer. The second hidden layer features less neurons compared to the first as we found that further deeper layers with equal or more neurons degrades performance of the model. We limit the second hidden layer to feature a number of neurons less than the first hidden layer and more than the input later.

The second improvement that has been implemented is the use of a weight constraint where "a constraint is imposed on the weights for each hidden layer ensuring that the maximum norm of the weights does not exceed a value of 3" [29]. Testing different values for the weight constrained showed that values of either 2 or 3 perform the best. Combined with a Dropout layer for each Dense layer we can expect better performance in training and reduce the risk of over fitting on our data. Further improvements to the forward and back propagation processes we modify the weight initializations of each hidden layer to use "glorot_normal" and "lecun_uniform" for the input layer, which [27] Gaussian initialization scaled by fan_in + fan_out (Glorot 2010) and Uniform initialization scaled by the square root of the number of inputs (LeCun 98) respectively. The justification for this decision is that through testing accuracy results between epochs displayed fewer variances and therefore represent minor but more stable scores.

Finally, we choose to use the "adamax" optimization algorithm instead of the standard 'adam' optimizer as it is better suited to sparse data, such as ours, and so we observe minor improvements in our network's performance.

Experiment three result

Network Layers	Training Score	Validation Score	Input Dimension	Data set size
5	70.00	41.41	1311	1000

V. Conclusion

In this research we would try to attempt the first deep learning model that includes both the sentence level and emotion icons sentiment analysis model for our language. as we know Amharic is one of the semantic language low resource-oriented language despite this scarcity the researcher tries to explore thus sentence from online review in immigration, war and public relation issue by its post ID and extract using Facepager tool which is a software used for different social medias like Facebook to extract different data. With the Facebook module you can get data via the so called Graph API. You can access public available data or any data you are allowed to see with your account that is at least data about you and your friends.

We discussed issues with collecting and processing Amharic socio-political data, where foreign languages, punctuation, hash tags usernames and information like hyperlinks is difficult to filter out on a large scale. To overcome this number of natural language processing books and online resources were studied in detail to learn the various techniques for solving these problems. After collecting the data, we have preprocessing, tokenization, stop word removal, stemming and normalization is undertaken to get clear data set.

Furthermore, a notable challenge was in the conceptualization of representing text as numerical data, which was a difficult problem to understand as well as solve, particularly in the context of emoji. Libraries such as Word2Vec represent words in a vector form, where related words in sentiment are based on their distance between each other. Had our problem been that of traditional text, a utility such as this would have been very useful but as we have learnt in this thesis, a single numerical representation of an emoji is not a true rendition of its meaning and or its context. It is equally difficult for a human to interpret social media dialect, particularly with the use of emoji, and so we chose to implement our own rules rather than automate the process based on method such as vector distances for words. This decision meant that a large portion of the research time was dedicated to the collection and manual annotation of comments.

Perhaps the most demanding of these problems was in the manual annotation of the data, which required that I read through a large sample of the data and assess its sentimental value by hand. To establish a sense of the magnitude of the problem from a sample of around 1500 post-processed comments, only around 1000 were viable for training our model, and each of these were required to be classified against the rules we presented in chapter four which ultimately consumed a large portion of the research time. However, the time and work required for generating a quality labeled data set was necessary as current works or existing labeled Amharic data are only text based but also use fewer or different class labels.

Reference

[1] A. f. c. linguistics., "http://www.aclweb.org/archive/misc/what.html," 18 september 2015. [Online].

[2] "Introduction to Sentiment Analysis," [Online]. Available: https://lctmaster.org/files/MullenSentimentCourseSlides.pdf. . [Accessed 17 june 2017].

[3] L. B., "Sentiment analysis and opinion mining," in Synthesis Lectures on Human Language, 2012.

[4] T. M. a. P. T, "Survey on mining subjective data on the web.," in Data Mining and Knowledge Discovery, 2012.

[5] Ilona Buchem, Ralf Applet, Sascha Kaiser, Sandra Schon, and Martin Ebner, "microblogging-," 2011.

[6] Danah Boyd. , "Exploring the definition of a medium. Reconstruction," in A blogger’s blog.

[7] Akshay Java, Xiaodan Song, Tim Finin, and Belle Tseng. , "Why we twitter understanding microblogging usage and communities.," ACM, vol. In Proceedings of the 9th , no. Web mining and social network analysis, pp. 56-65, 2007.

[8] [Online]. Available: http://www.internetworldstats.com/africa.htm#et.

[9] A. Alemu, "MSc thesis,Sentiment Mining for Discovering Hidden Knowledge in Opinionated Amharic Text Review," University of Gondar, Gondar, 2013.

[10] B. Mengistu, "MSc thesis, Sentiment Analysis for Amharic opinionated text," Ababa university, Addis Ababa, Ethiopia, 2013.

[11] G. ABREHAM, "OPINION MINING FROM AMHARIC ENTERTAINMENT TEXTS," Addis Abeba Universty, Addis Abeba, 2014.

[12] T. Tilahun, "Linguistic Localization of Opinion Mining from Amharic Blogs," Addis Abeba Universty, Addis Abeba, 2014.

[13] Wondwossen P., and Wondwossen M, "A Machine Learning Approach to Multi-Scale Sentiment Analysis of Amharic Online Posts.," HiLCoE Journal of Computer Science and Technology, Vols. Vol. 2, No. 2 , 2014 .

[14] G. Selama, "Master Thesis,,Sentiment Mining Model for Opinionated Amharic Texts," Addis Ababa University, Addis Ababa,Ethiopia, 2010..

[15] Baccianella, S., Esuli, A., and Sebastiani, F, "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining," in ‘Book SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining’ (2010, edn.), 2010, pp. 2200-2204.

[16] Shahab Saquib Sohail and Jamshed Siddiqui, R.A, "Feature extraction and analysis of online reviews for the recommendation of books using opinion mining technique," Aligarh Muslim University, 18 July 2016.

[17] Jandail, R.R.S, "A proposed Novel Approach for Sentiment Analysis and Opinion Mining," International Journal of UbiComp, , vol. 5, p. 1, 2014.

[18] Liu, B., "Opinion mining and sentiment analysis," Springer, no. Web Data Mining, pp. 459-526, 2011.

[19] Cambria, E., Havasi, C., and Hussain, "SenticNet 2:Semantic and Affective Resource for Opinion Mining and Sentiment Analysis," pp. 202-207, 2012.

[20] "A Quick Introduction to Neural Networks," [Online]. Available: Internet: https://ujjwalkarn.me/2016/08/09/quick-intro-neural-networks/. [Accessed 04 july 2017].

[21] WalaaMedhata,AhmedHassanb,HodaKorashy:, "Sentiment analysis algorithms and applications," [Online]. Available: Internet http://www.sciencedirect.com/science/article/pii/S2090447914000550. [Accessed 4 november 2017].

[22] "Introduction to Deep Neural Networks," [Online]. Available: http://deeplearning4j.org/neuralnet-overview.html#concept. [Accessed 23 september 2018].

[23] "What is the role of the activation function in a neural network? How does this function in a human neural network system?," [Online]. Available: internet: https://www.quora.com/What-is-the-role-of-the-activation-function-in-a-neural-network.

[24] Amanuel, "Probablistic Information Retrieval," in master thesis, Addis Ababa , 2012.

[25] Tessema Mindaye, Meron Sahlemariam, and Teshome Kassie,, "The need for AmharicWordNet," in Global WordNet Conference, , Mumbai, India, 2010.

[26] "Cross-validation for time-series," [Online]. Available: internet: https://stackoverflow.com/questions/tagged/python. [Accessed 24 june 2017].

[27] "Keras Documentation," [Online]. Available: https://keras.io. [Accessed 12 augest 2016].

[28] "Develop Your First Neural Network in Python With Keras Step-By-Step," [Online]. Available: http://machinelearningmastery.com/tutorial-first-neural-network-python-keras/. [Accessed 26 04 2017].

[29] "5 Step Life-Cycle for Neural Network Models in Keras," [Online]. Available: http://machinelearningmastery.com/5-step-life-cycle-neural-network-models-keras/. [Accessed 07 05 2017].

[30] "Display Deep Learning Model Training History in Keras internet," [Online]. Available: http://machinelearningmastery.com/display-deep-learning-model-training-history-in-keras/. [Accessed 25 july 2017].