SENTIMENT ANALYSIS USING DEEP LEARNING IN ALBANIAN: AN OVERVIEW

Muhamet KASTRATI and Marenglen BIBA
Department of Computer Science, University of New York Tirana, Albania

---

ABSTRACT

Sentiment analysis aims to automatically classify people's opinions, attitudes and emotions expressed towards a particular entity, topic, or event. Here, deep learning has been widely used, and proved useful to sentiment analysis problems. Efforts are mostly devoted to the rich-resource languages; there is only little research on sentiment analysis for the low-resource languages. Therefore, the present paper aims to provide an overview of prominent deep learning models successfully applied on sentiment analysis for low-resource languages such as Albanian language. It begins with a short introduction, motivation and some basic concepts on sentiment analysis, and most prominent deep learning models applied for sentiment analysis. Papers providing information about deep learning approaches and some of the current trends in the field are subsequently here highlighted. In the end, discussions are here made, and limitations of deep learning approach in sentiment analysis in Albanian language reported.
***Keywords*:** Albanian language, sentiment analysis, machine learning, deep learning

1. INTRODUCTION

Sentiment analysis or opinion mining aims to automatically classify people's opinions towards a particular topic or event (Liu 2012; Singh *et al*., 2020). Over the past 20 years, many researchers from different areas or disciplines are paying more and more attention to sentiment analysis, which is made possible thanks to the availability of the huge amounts of data on social media platforms such as Facebook, Twitter, WhatsApp, LinkedIn, Instagram, to name a few.

Sentiment analysis applications, among others, include domains such as business, politics, healthcare, communications, education, financial markets, etc.

During the last two decades, sentiment analysis has been widely studied in NLP, data mining, web mining, text mining, and information retrieval. It also has been studied in other areas from computer science to management sciences and social sciences. Over the last few years, many efforts have been made by machine learning research community to study sentiment analysis focusing on social networks.

There are several learning approaches for sentiment analysis to be grouped into five major categories including lexicon-based, machine learning, deep learning, a combination of lexicon and machine learning, concept-based learning approaches.

Deep learning is a kind of representation learning (hierarchical feature learning), which makes it easier to extract useful information (automatic feature engineering) when building classifiers or other predictors (Bengio 2013). Deep learning approaches have recently become the dominant methodology in sentiment classification and of great interest for the community and industry. Sentiment classifiers relying on deep learning (neural based) approaches have shown state-of-the-art results on various tasks (Zhang *et al*., 2018).

Research on sentiment analysis for rich-resource languages has already achieved great success especially for English due to the availability of large enough datasets, plenty of tools and opinion lexical resources. On the other hand, greater efforts ought to be made to develop and adopt existing tools and resources for similar performances in other low-resource languages, such as Albanian language.

Despite significant result achieved in the field of sentiment analysis, some challenges (limitations) regarding this area such as identification of sarcasm and irony, noisy texts, sentiment to objective statements, etc. still remain. In addition, one major limitation for sentiment analysis today is most tools and resources are available only for rich-resource languages; in contrast, many low-resource languages that are spoken and written by millions of people have no such resources or tools available.

The present paper reviews and assesses the performance of deep learning models that have achieved state-of-the-art results on sentiment analysis for low-resource languages.

The reminder of the paper is organized as follows: Section 2 overviews briefly the background theory through a general introduction of sentiment analysis, learning approaches used for sentiment analysis, especially from the context of sentiment analysis for the low-resource languages. Section 3 presents state-of-the-art review on sentiment analysis using machine learning and deep learning-based approach, focusing on low-resource languages. The paper concludes with some future directions presented in Section 4.

2. BACKGROUND

This section introduces shortly the basic concepts that are more than necessary to facilitate understanding of the sentiment analysis, learning approaches used for sentiment analysis and a brief overview of deep learning methods that are mainly used in this area.

Sentiment Analysis

Sentiment analysis, otherwise named opinion mining, is the field of study that analyses people's opinions, attitudes, and emotions expressed towards a particular topic or event (Liu 2012; Singh *et al*., 2020). The term sentiment analysis is mostly used in industry. Both sentiment analysis and opinion mining terms are used in the academic enterprises in this field. Here, the terms ''sentiment analysis'' and ''opinion mining'' are used interchangeably (Liu 2012).

Machine learning

Machine learning (ML) is a subfield of artificial intelligence that deals with building a computer system that learns from data without being explicitly programmed (Samuel 1959; Domingos 2012). It has made significant progress and received enormous attention in the research community and industry.

Over the last two decades, ML techniques have been used successfully to a wide range of applications ranging from speech recognition, document categorization, document segmentation, part-of-speech tagging, word-sense disambiguation, named entity recognition, parsing, machine translation, sentiment analysis and opinion mining and so forth.

In general, there are three major learning approaches in ML field: supervised, unsupervised and reinforcement learning. However, supervised learning dominates compared to other ML approaches, as almost all the significant results are achieved by using this approach.

Sentiment analysis using ML techniques deals with the classification of unstructured data and text into positive, neutral, and negative categories (Ahmad *et al.*, 2017).

Deep learning

Artificial neural networks (ANNs) or simply neural networks (NN) are well-known class of machine learning models that are loosely inspired by concepts from biological brains. On the other hand, deep learning is a subfield of machine learning that deals with deep learning in artificial neural networks (Schmidhuber 2015).

Despite the opportunities that offer the machine learning methods, these methods are struggling with their ability to deal with natural data in their raw form. In other words, these methods rely on manual feature engineering.

On the other hand, deep learning is a kind of representation learning that makes possible automatic feature engineering when building classifiers or other predictors (Zhang *et al*., 2018). Furthermore, deep learning benefits from the availability of huge amount of data and fast enough computers which make it possible to train large neural networks, which further benefit in their performance (Ng 2017).

Deep learning has made significant progress and received enormous attention in the research community and industry, due to their state-of-the-art result achieved in various applications including computer vision, speech recognition, natural language processing, machine translation, online advertising, web search, recommendation systems, etc. Further details about these models could be found in (LeCun *et al.*, 2015; Schmidhuber 2015; Kastrati and Biba 2021).

3. STATE-OF-THE-ART REVIEW

More recently, researchers have made considerable progress with regard to sentiment classification by employing several deep learning and machine learning concepts. In this section, we will briefly describe several studies concerning sentiment analysis of web contents regarding end user opinions, attitudes and emotions expressed towards a particular product, topic or event using deep learning approach.

Machine Learning for sentiment classification

Conventional machine learning models have been widely used for sentiment analysis tasks. Ahmad *et al*., (2017) through their comprehensive review on ML algorithms and techniques used for sentiment analysis, they provided a comparative study on some of the most popular machine learning models used in the area including SVM, Naïve Bayes, Logistic Regression, Maximum Entropy, Random Forest, and Decision Trees. They concluded that SVM outperforms the other machine learning algorithms regarding sentiment analysis tasks in terms of accuracy and efficiency.

Biba and Mane (2014) presented the first approach for sentiment analysis in Albanian. They developed a machine-learning model to classify text documents belonging to a negative or positive opinion regarding the given topic. To train their machine learning models they built a corpus of 400 documents containing political news consisting of five different topics. In this case, each topic was represented by 80 documents classified as positive or negative. The authors made an empirical comparison between 6 different

machine learning algorithms, including Bayesian Logistic Regression, Logistic Regression, SVM, Voted Perceptron, Naïve Bayes, and Hyper Pipes for classification of text documents, achieving an accuracy between 86% and 92% depending on the topic. The authors conclude that, in general, to achieve higher accuracy in sentiment analysis, a larger corpus in the Albanian language is needed.

Kote *et al*., (2018) presented a comprehensive experimental evaluation of machine learning algorithms applied for opinion mining in Albanian language. Among the algorithms tested that performed better include Logistic and Multi-Class Classifier, Hyper Pipes, RBF Classifier, and RBF Network with the classification accuracy ranging from 79% to 94%. Classification algorithms are trained on a corpus of 500 news articles in Albanian consisting of five different topics. Each topic was represented with a balanced set of articles classified as positive or negative. The experimental results were interpreted concerning several evaluation criteria for each algorithm showing interesting features on the performance of each algorithm.

Later, Kote *et al.,* (2018b) further extended their research from an in-domain to a multi-domain corpus (for the Albanian language). They created 11 multi-domains corpuses combining opinions from 5 different topics, which was later split for training and testing of 50 classification algorithms in Weka. The result obtained showed that from all evaluated algorithms seven of them performed better and three of them are based on Naïve Bayes. In addition, the authors concluded that the availability of a larger corpus in the Albanian language could further improve the performance of classification algorithms for opinion mining.

Kote and Biba (2021) presented their results of experimental evaluation about the performance of machine learning techniques on opinion classification tasks in the Albanian language. Their approach to opinion mining addresses the problem of classifying text document opinions as positive or negative opinions for Albanian language. Authors here conducted an experimental evaluation of fourteen techniques used for opinion mining. They tested several machine learning algorithm's performances by the help of Weka. Authors reported that Naïve Bayes Multinomial outperformed other machine learning algorithms tested, and the best result obtained in terms of correctly classified instances was 84.88%.

Deep Learning for sentiment classification

This section briefly overviews some of the most prominent deep neural networks employed in sentiment analysis tasks including Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and their extended version such as Bidirectional Recurrent Neural Network (BiRNN), Long-Short Term Memory (LSTM), Bidirectional Long-Short Term Memory

(BiLSTM) and Gated Recurrent Unit (GRU) and hybrid neural networks that takes the advantages of the two complementary deep neural networks.

BiRNNs are extended variant of RNNs architecture. While unidirectional RNNs are influenced from the previous inputs to make predictions about the current state, BiRNNs pull in future data to improve its accuracy. In other words, the output is influenced from previous inputs (backwards) and future states (forwards) simultaneously. LSTM is an RNN architecture specifically designed to handle the problem of long-term dependencies and performs better at storing and accessing information than standard RNNs. On the other hand, GRU was proposed with idea to make each recurrent unit to adaptively capture dependencies of different time scales, respectively, to reduce computational burden caused by additional parameters used in LSTM (Kastrati and Biba 2021).

In general, each of these neural networks is specialized to certain tasks, due to the different nature in their text modelling capabilities. For example, CNNs are specialized for using in computer vision, and RNNs are mainly used for sequential inputs such as time series or natural language. However, their extended versions have also achieved significant result, such as the case with 1D-CNN that can extract local features from the text. On the other hand, BiLSTM has proved to perform good at capturing contextual information from both direction as well as the long-range dependencies. Furthermore, the hybrid models benefit by taking the advantages of both complementary architectures (Kastrati *et al*., 2021).

There is a considerable number of research studies about sentiment analysis problems, mostly in English in the last decade, but only a few research studies on sentiment analysis for low-resource languages, such as Albanian language. Skënduli *et al*., (2018) presented an approach for user-emotion detection on microblogging texts and postings in the Albanian language. The authors studied user-emotion at the sentence level by using deep learning methods. Their approach was based on the idea of classifying a text fragment into a set of pre-defined emotion categories (based on Ekman's model), and therefore aims at detecting the emotional state of the writer conveyed through the text. To perform their experiments, they built a new dataset that contained manually annotated posts on Facebook belonging to some active Albanian politicians, which were classified using Ekman's model, and finally divided into six smaller datasets. A set of experiments on these datasets to evaluate deep learning and conventional machine learning models subsequently followed. In addition, during their analysis, they also adopted a domestic stemming tool for the Albanian language to pre-process the datasets, which showed a slight improvement in the classification accuracy. In general, the obtained results showed that deep learning outperformed the other conventional machine learning models, i.e., Naïve Bayes (NB), Instance-

based learner (IBK), and Support Vector Machines (SMO) with a given accuracy ranging from 70.2% to 91.2% for (correctly classified unstemmed instances), and 67.0% to 92.4% for (correctly classified stemmed instances).

Later, Skenduli and Biba (2020) proposed an analysis of micro-blogging content to characterize the users individually when writing posts with emotional content. First, the present investigation aimed at collecting textual units that allowed them to summarize the lexicon used by the user. They focused on sentence-based emotion detection problems. The main aim of this approach was to classify the textual units into a set of pre-defined emotion categories based on the Ekman model. Second, the analysis was carried out through a keyword extraction approach, which aimed at finding representative generic word sets in the form of prototypes of textual unit clusters. In addition, several experiments were carried out from different points of view, yet always focusing on the user. They provided useful information about classification accuracy, clustering, and other valuable information concerning user emotion profiling.

Kastrati *et al*., (2021) presented the sentiment analysis of people's opinions expressed on Facebook with regards to the pandemic situation in Albanian language. They developed a deep learning model to classify people's opinions belonging to a neutral, negative, or positive opinion regarding the given topic. To train the deep learning models they created a dataset containing 10.742 manually annotated Facebook comments in the Albanian language. The authors reported their efforts on the design and development of a sentiment analyser that relies in deep learning. The researchers reported their experimental results obtained using deep learning models with static and contextualized word embeddings, that is, fastText and BERT. The authors concluded that combining the BiLSTM with an attention mechanism outperformed the other approaches in their sentiment analysis task. The best results achieved by their proposed model given in terms of Precision Recall, and F1 score were 72.31%, 72.25%, and 72.09%, respectively.

Vasili *et al*., (2021) tested and reviewed several approaches in sentiment analysis on Twitter messages for the Albanian language. In addition, they compared the results among several methods and noted the challenges that arise when dealing with sentiment analysis for Albanian language, and finally made their suggestions for the future work. Authors here investigated the performance of sentiment classification techniques using three main approaches: traditional machine learning, lexicon-based and deep learning-based approach. Their experiments revealed that LSTM based RNN with Glove as a feature extraction technique provides the best results with F-score = 87.8%, followed by Logistic Regression.

Table. 9 A summary of the studies on sentiment analysis for low-resource languages (the case of Albanian language) using machine/deep learning classifiers

Ref no.	Motivation and aim of the work	Methods used in the work	Dataset used	Obtained Accuracy	Limitations
Biba and Mane (2014	To analyze polarity of documents containing political news in Albanian language consisting of five different topics	Six different machine-learning algorithms: Bayesian Logistic Regression, Logistic Regression, SVM, Voted Perceptron, Naïve Bayes, and Hyper Pipes.	400 documents containing political news consisting of five different topics.	Achieved accuracy was between 86% and 92% depending on the topic.	Small and in-domain corpus. Dataset is not publicly available.
Kote *et al*., (2018)	To analyse the sentiments of the news articles in Albanian language consisting of 5 different topics	Machine learning: Logistic and Multi-Class Classifier, Hyper Pipes, RBF Classifier, and RBF Network	500 news articles in Albanian consisting of five different topics.	Accuracy obtained was ranging from 79% to 94%.	Small and in-domain corpus.
Kote *et al*., (2018b)	To further extend their research from an in-domain to a multi-domain corpus	Machine Learning: classifiers. From 7 of them that performed better, 3 were based on Naïve Bayes.	11 multi-domains corpuses combining opinions from five different topics	Weka was used for analysis, and the result obtained were promising.	Small corpus. Datasets are not publicly available.
Kote and Bi (2021)	To evaluate the performance of machine learning techniques on opinion classification tasks in the Albanian language	Machine Learning: classifiers. Several algorithms evaluated: Naïve Bayes algorithm showed the best performance in terms of the average value.	500 news articles in Albanian language consisting of five different topics such as higher education law, waste import, VAT in small businesses, tourism, and politics	Naïve Bayes Multinomial outperformed others with 84.88% of correctly classified instances.	Small corpus. Datasets are not publicly available.
Kastrati *et a* (2021)	To analyze people's opinions expressed on Facebook regarding Covid-19 pandemic situation in Albanian language	Deep Learning: 1D-CNN, BiLSTM, and a hybrid 1D-CNN + BiLSTM model	10742 comments collected from the NIPHK's Facebook page between 13th March and 15th August, 2020	The best results achieved given in terms of P: 72.31%, R: 72.25%, and F1: 72.09%.	Small and in-domain corpus.
Skënduli *et al*., (2018)	To analyse user-emotion detection on microblogging texts and postings in the Albanian language	Deep Learning: Convolutional neural networks (CNNs)	6,358 Facebook posts belonging to Albanian politicians	Accuracy for: Unstemmed: 70.2% - 91.2% Stemmed: 67.0% - 92.4%	Small and in-domain corpus. Dataset is not publicly available.
Skënduli an Biba (2020)	To analyse micro-blogging content to	Deep Learning: Convolutional neural	6,358 Facebook posts belonging	Accuracy for: Unstemmed:	Small and in-domain

	characterize the users individually when writing posts with emotional content	networks	to Albanian politicians	70.2% - 91.2%	corpus. Dataset is not publicly available.
Vasili *et al.*, (2021)	To test and review different approaches in Sentiment Analysis on Twitter messages in the Albanian language	Deep Learning: Convolutional neural networks (CNNs)	Was used the dataset consisting of 15 languages from (Mozetič *et al.*, 2021)	By sing LSTM-RNN with Glove, was achieved the best results given in terms of F1: 87.8%	"low quality annotators which should be eliminated from further considerations "

4. CONCLUSIONS

Here papers about sentiment analysis and emotion detection in Albanian as summarized in Table 1 are reviewed. The performance of conventional machine learning and deep learning methods clearly depends on the pre-processing and size of the dataset. Furthermore, deep learning benefits from using representation learning, large amount of training data and higher computational power that make possible training large and complex deep neural networks for sentiment and emotion classifiers. As it can be noticed, in situation where the dataset is vast, deep learning methods such as 1D CNN, LSTM, BiLSTM and combined CNN with BiLSMT architectures generally outperform conventional machine learning methods (Naïve Bayes, SVM, Logistic Regression and Decision Trees). Overall, we can conclude that deep learning methods are the method of choice for sentiment analysis and emotion detection in Albanian.

REFERENCES

Ahmad M, Aftab S, Muhammad SS, Ahmad S. 2017. Machine learning techniques for sentiment analysis: A review. *International Journal of Multidisciplinary Sciences and Engineering*, **8(3):** 27.

Bengio Y, Courville A, Vincent P. 2013. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, **35(8):** 1798-1828.

Biba M, Mane M. 2014. Sentiment analysis through machine learning: an experimental evaluation for Albanian. *Recent Advances in Intelligent Informatics*. Springer, 195–203.

Domingos P. 2012. A few useful things to know about machine learning. *Communications of the ACM*, **55 (10):** 78–87.

Kastrati M, Biba M. 2021. A State-of-the-art survey on deep learning methods and applications. *International Journal of Computer Science and Information Security* (IJCSIS), **19(7):** 53-63.

Kastrati Z, Ahmedi L, Kurti A, Kadriu F, Murtezaj D, Gashi F. 2021. A Deep learning sentiment analyser for social media comments in low-resource languages. *Electronics*, **10 (10):** 1133.

Kote N, Biba M, Trandafili E. 2018. A thorough experimental evaluation of algorithms for opinion mining in Albanian. International Conference on Emerging Internetworking, Data & Web Technologies. Springer, 525–536.

Kote N, Biba M, Trandafili E. 2018 (b). An experimental evaluation of algorithms for opinion mining in multi-domain corpus in Albanian. International Symposium on Methodologies for Intelligent Systems. Springer, 439–447.

Kote N, Biba M. 2021. Opinion Mining in Albanian: Evaluation of the Performance of Machine Learning Algorithms for Opinions Classification. *International Journal of Innovative Science and Research Technology*, **6 (6)**: 964-973. ISSN No-2456-2165.

LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. *Nature*, **521(7553):** 436-444.

Liu B. 2012. Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, **5(1)**: 1-167. Morgan & Claypool Publishers, Vermont, Australia.

Mozetič I, Grčar M, Smailović J. 2016. Multilingual Twitter sentiment classification: The role of human annotators. PloS one, 11(5), e0155036.

Ng A. 2017. Machine learning yearning. URL: http://www.mlyearning.org/ (96), 139.

Samuel AL. 1959. Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, **3 (3):** 210–229. doi: 10.1147/rd.33.0210.

Schmidhuber J. 2015. Deep learning in neural networks: An overview. *Neural networks*, vol. 61 85–117. https://doi.org/10.1016/j.neunet.2014.09.003.

Singh NK, Tomar, DS, Sangaiah AK. 2020. Sentiment analysis: A review and comparative analysis over social media. *Journal of Ambient Intelligence and Humanized Computing*, **11**, 97–117.

Skënduli MP, Biba M, Loglisci C, Ceci M, Malerba D. 2018. User-emotion detection through sentence-based classification using deep learning: A case-study with microblogs in Albanian. *International Symposium on Methodologies for Intelligent Systems*. Springer, 258–267.

Skënduli MP, Biba M. 2020. Classification and clustering of emotive microblogs in Albanian: Two user-oriented tasks. *Complex Pattern Mining*. Springer, 153–171.

Vasili R, Xhina E, Ninka I, Terpo D. 2021. Sentiment analysis on social media for Albanian language. *Open Access Library Journal*, **8(6):** 1-31.

Zhang L, Wang S, Liu B. 2018. Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, **8(4):** e1253.