User-Emotion Detection Through Sentence-Based Classification Using Deep Learning: A Case-Study with Microblogs in Albanian

Marjana Prifti Skenduli[1(✉)], Marenglen Biba[1], Corrado Loglisci[2], Michelangelo Ceci[2], and Donato Malerba[2]

[1] University of New York Tirana, Tirana, Albania
{marjanaskenduli,marenglenbiba}@unyt.edu.al
[2] Universita' degli Studi di Bari, Bari, Italy
{corrado.loglisci,michelangelo.ceci,donato.malerba}@uniba.it

Abstract. Human emotion analysis has always stimulated studies in different disciplines, such as Cognitive Sciences, Psychology, and thanks to the diffusion of the social media, it is attracting the interests of computer scientists too. Particularly, the growing popularity of Microblogging platforms, has generated large amounts of information which in turn represent an attractive source of data to be further subjected to opinion mining and sentiment analysis. In our research, we leverage the analysis performed on micro-blogging texts and postings in Albanian language, which enables the use of technologies to monitor and follow the feelings and perception of the people with respect to products, issues, events, etc. Our approach to emotion analysis tackles the problem of classifying a text fragment into a set of pre-defined emotion categories and therefore aims at detecting the emotional state of the writer conveyed through the text. In order to achieve this goal, we perform a comparative analysis between different classifiers, using deep learning and other classical machine learning classification algorithms. We also adopt a domestic stemming tool for Albanian language in order to preprocess the datasets used in a second round of experiments. Experimental evaluation shows that deep learning produces overall better results compared with the other methods in terms of classification accuracy. We present also other findings related to the length of the texts being processed and the impact on the classifiers' accuracy.

1 Introduction

The use of social media has essentially brought two main novelties in the society, the facilitation of intertwining social relationships and the possibility to express and share feelings and emotions. Human emotion analysis has always stimulated studies in different disciplines, such as Cognitive sciences and Psychology, and, thanks to the diffusion of the social media, is attracting the interests of computer

scientists. The reason may be seek in the impact that the analysis of emotions may have on real-world applications and emerging challenges. A representative case is the analysis of micro-bloggings and postings, which enables the use of technologies to monitor and follow the feelings and perception of the people with respect to products, issues and events. Emotion analysis is chiefly concerned with the problem of classifying a text fragment into a set of pre-defined emotion categories and therefore aims at detecting the emotional state of the writer conveyed through the text. To define the categories, many works resort to cognitive-based emotion theory, for instance, the Ekman' model [7]. The recognition of emotions within a text fragment is more complex than other tasks of text classification for several reasons related. First, there are no grammatical or syntactic structures able to characterize a category and discriminate others. Second, the same emotion can be expressed with different lexical forms. Third, emotions can be attributed to different causes and can be induced by different behaviors of human beings or events.

The length and nature of the text distinguish the problem of emotion detection into sentence-level classification and document-level classification. The sentences are prone to the sparseness due to the limited content and quality of the writing. The typical case is represented by the short texts in social media, which are often characterized by abbreviations, emoticons and typos [10]. The documents are often formal, written in correct language but may contain impersonal content which does not express an emotional state [5]. Contrary to the documents, the sentential forms are often subjective, personal and represent manifestations of the personality of individuals. Moreover, the sentences are information units of the documents, therefore the analysis of emotions at the level of the sentences can support the analysis at the level of the documents.

For both problems, research works are essentially based on unsupervised learning and supervised learning [21]. The approaches of the first category largely rely on the availability of lexical resources, which often depend on the specific language of the resources and thus offer coverage of a specific language in emotive texts. The supervised learning approaches rely on the annotated datasets, which are generated especially for frequently spoken languages and are subject to manual labeling.

However, many linguistic resources and annotated texts have been generated for wide-coverage languages [2], such as English, Chinese and Arabic, while no attempt has been made for other rare Indo-European languages, such as the Albanian. Nowadays, Albanian is spoken by 7.6 million Albanians living in Albania, Kosovo, Montenegro, northwest of Macedonia, northwest of Greece, in some Western Europe countries and in North America where thousands of Albanians have migrated and currently live. Despite its late documentation, the Albanian language is of great interest to linguists and not only, due to its unique and archaic traits. Albanian has distinctive features which range from morphological to lexical viewpoints. It has a large alphabet (thirty-six letters), and constitutes a lot of words with particles in two units, which are difficult to label [16]. The words used to express the same idea can have different types of grammatical

relations, they can be nouns and verbs. In addition, they can be associated to general semantic categories or specific categories [3]. Albanian is a very rich language in words that hold more than one meaning. They are called polysemantic words, which may mislead an automatic classifier since the emotion being expressed might be related to only one of the meanings of the word [4]. The richness of the lexical characteristics and syntactic rules however are not enough to stimulate the generation of linguistic resources, such as thesaurus and dictionary, which drives us towards application of supervised learning to analyse emotive texts.

In this paper, we present a case study for the classification of emotion on micro-blogging sentences. Motivated by the growing relevance of the social media to mirror political information influenced by emotion states, we generated a dataset of Facebook statuses posted by Albanian politicians and manually selected the sentences and annotated them by using the Ekman's emotion categories [7]. Then, we performed extensive experiments with several state-of-art classifiers under different perspectives.

2 Related Work and Contribution

In recent research, there has been great attention in the study of various aspects of emotion analysis, such as emotion resource creation, emotion cause event analysis, reader emotion detection, and emotion detection for writers authoring long-texts (documents) and short-texts (messages), which is the focus of the current work.

In [14], the authors generated a lexicon of pairs word-emotion based on hashtags from an annotated Twitter dataset. Experiments show an improvement in accuracy by using SVM classifiers for the six basic emotion categories. Different lexical resources, such as, WordNetAffect, SentiWordnet and SenticNet have been used in [6] with a Conditional Random Field classifier. The algorithm relies on three scoring methods and is able to outperform many systems which use only one lexicon.

An approach which does not need lexicon has been described in [1]. The algorithm works on semantic and syntactic relatedness. The approach described in [10] does not resort to lexicons but exploit other features internal to the texts written in Chinese. It incorporates both the label dependence among the emotion labels and the context dependence among the contextual instances into a factor graph, where the label and context dependence is modeled as various factor functions.

In [21], the authors use intra-sentence based features to determine the emotion label set of a target sentence coarsely through the statistical information gained from the label sets of the most similar sentences in the training data. Then, they use the emotion probabilities between neighboring sentences to refine the emotion labels of the target sentences. The proposed algorithm is evaluated on Ren-CECps, a Chinese blog emotion corpus. The same corpus has been analyzed in [17] which use a polynomial kernel to compute similarities between sentences and basis sets of emotions. A different language is studied in [5], which

consider a Bengali blog dataset annotated at the word-level and proposed a scoring technique for the constituents of the sentences.

The method reported in [19] identifies emotional patterns from part-of-speech tags of emotion triggered terms and its co-occurrence terms. Patterns are classified hierarchically into categories referred to positive and negative emotions. Sentences are categorized by capturing the degree of emotive content with respect to the semantics of patterns.

Recently, some studies overcome the constraint to require emotive words, common to many works, by considering emotion signals, such as polarity shifters, negations, emoticons and slangs [2]. The authors propose a rule-based classification framework which combines a pipeline of classifiers which learned from the emotion signals.

A research that is recently attracting attention is that of focusing the study on the emotions manifested by specific users [15,20]. This has a two-fold motivation. First, in social media platforms, the content of the messages is often expression of the reaction to or influence of the messages posted by specific users. A typical example is represented by Twitter users with many followers. Second, the distribution of the emotion categories could be not fair because some users could have never been expressed some emotions. In these cases, a user-centric study could be more reliable because we could better learn the lexical characteristics. The current work leverages upon these considerations and reports a case study focused on the emotive tweets of (influential) users. We propose a supervised learning approach based on a Deep learning architecture, which compared against some competitors, is able to accurately account for the specific distribution of the emotion categories over the posts of the selected politicians.

3   Construction of the Sentence-Level Datasets

The extant studies report sentiment classification at varying levels of granularity (document and sentence based), mainly for popular languages like English, German, Spanish, Chinese etc. While fully recognizing the scarcity of Albanian linguistic resources i.e. corpora, gazetteers and dictionaries, we decided to build our own corpuses using as a primary source public Facebook pages belonging to high rank Albanian politicians. It goes without saying that the finished corpuses represent an important contribution of this work. It is fitting therefore that we describe the steps followed for the construction of corpuses, which were time consuming, yet imperative to the aim of this work.

3.1   Data Collection and Assembly

For the purpose of collecting abundant quantities of micro blogging data, we used RestFB[1] - a simple and flexible Facebook Graph API client written in Java language. It is an open source software released under the terms of the

---

[1] https://restfb.com//.

MIT License. Expanding upon this step, specifically, we had to reuse the source code and setup a whole framework that allowed us to fetch posts out of public/community pages belonging to Albanian public figures. The framework we propose entails the following depicted modules as shown in Fig. 1. We fetched around 60000 posts belonging to 119 Albanian politicians, shortlisted among the ones who are pretty active in Facebook and popular to the general perception of the social media audience. The fetched posts where captured and stored in a local SLQ database, from where we could extract and build sample datasets for experimental purposes.
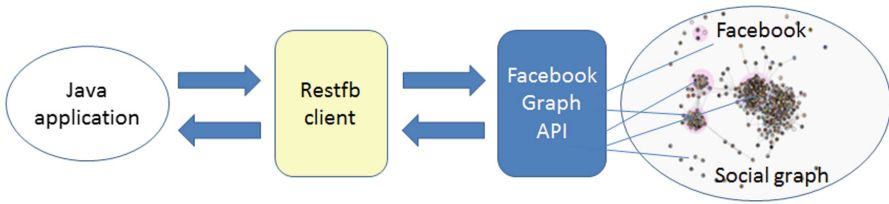

Fig. 1. Data collection framework.

3.2 Data Preprocessing

Prior to execution of sentence-based classification, we had to perform the preparation tasks on the raw text datasets, such as data cleaning and stemming. Initially, we exported six sample datasets from our SQL raw collection. The six datasets were manually cleaned and annotated using the Ekman' model [7] - a vocabulary consisting on seven emotions: JOY, ANGER, DISGUST, FEAR, SHAME, GUILT and SADNESS. As a result, we constructed six datasets, further referred to as per their respective ID numbers: D1, D2, D3, D4, D5 and D6. Dataset D1 is a multiuser dataset that assembles 2325 posts from 119 users, not longer than 200 characters distributed across the period of time January-March 2018. The remaining datasets D2, D3, D4, D5 and D6 are single-user datasets consisting of 159, 322, 1002, 1481 and 1069 posts respectively, not longer than 200 characters each and distributed across the period of time 2008–2018. Moreover, we manually annotated the sentiment polarity of each sentence in both training and testing corpuses. We chose a 70:30 split of our datasets into training and testing sets. The annotation for training corpuses is used to train the classifier, while the annotation for testing corpuses is used to test the accuracy of sentiment classification at sentence level.

4   Sentence-Based Classification

In this work, we address the categorization of the emotions expressed in the collected sentences with a Deep learning (DL) architecture, which revises a state of the art model originally designed for binary classification [8].

DL architectures have attracted much interest for their peculiarities to learn with small intervention on the data representation and feature engineering. In Natural Language Processing, we can enumerate many DL methods which have been proposed to investigate as many tasks, such as speech recognition and sequence labeling (e.g., POS tagging).
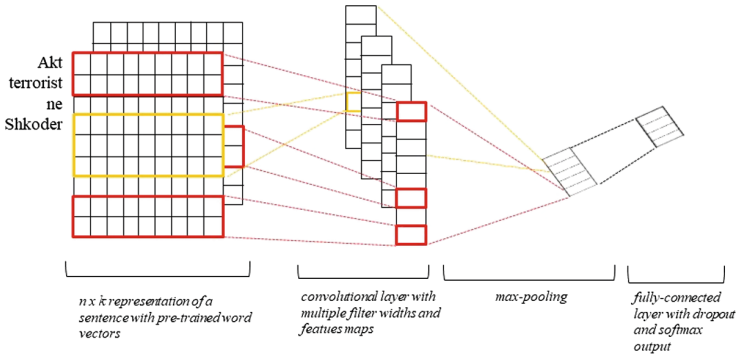

Fig. 2. Deep learning architecture for sentence-based classification of emotions.

In this paper, we extend the model proposed in [8] to the multi-class problem (the categories refer to the Ekman's basic emotions) and apply it to Albanian language. The model is sketched in Fig. 2 and summarized in the following. The Input layer embeds words into low-dimensional vectors. More precisely, a raw sentence is represented as a concatenation of the vectors, each referring a word of the sentence. All the vectors have length k and the sentence representation has length n, so all the sentences have n vectors. A sentence is padded if necessary.

The Convolutional layer performs filters over the embedded word vectors using multiple widths. More precisely, we apply convolution operations to windows of width h (that is, h words), in order to produce a long feature vector of length hk. Then, we apply a max-pool operation on the feature map, in order to select the most important feature. A dropout regularization is applied to the Convolutional layer, in order to prevent the co-adaptation of hidden units and force them to learn individually useful features. Finally, there is fully connected *softmax* layer whose output is the probability distribution over labels. Further details can be found in [8]. The experimental configuration in terms of the architecture parameters is reported in Sect. 5.

As to the word embeddings in Albanian, which is a specific characteristic of the current paper and it is beyond the original model, we consider the word vectors pre-trained on a Twitter corpus in English [9] and build the dictionary of the words by performing a translation of the words, which have no social tags (e.g., hashtag), into Albanian. Further details on this vocabulary are reported in Sect. 5.

Table 1. Experimental results in terms of accuracy (correctly classified unstemmed instances).

Dataset ID	Deep Learning	NB	IBK	SMO
D1	**91.2%**	75.5%	78.3%	81.8%
D2	**86.3%**	82.9%	75.3%	83.5%
D3	**81.9%**	79.1%	80.7%	80.7%
D4	70.2%	77.2%	79.9%	84.8%
D5	**90.5%**	81.2%	89.9%	90.1%
D6	**88.8%**	65.6%	75.4%	75.4%

5 Case Study and Experimental Evaluation

In this section we report experimental details on the architecture setup, vocabulary preparation in Albanian and performance of the learner in terms of accuracy.

The Deep learning architecture has been designed by using the Keras framework[2] and Tensorflow[3] as back-end. The hyper-parameters are configured as follows: sentence length $(n) = 33$, word embedding size $(k) = 300$, windows width $(h) = 3$, dropout rate = 0.2, number of neurons of the Convolutional layer =100, number of epochs = 10. The size of the vocabulary is almost 1M of words and the translation step has been performed with public API libraries[4].

We have performed experiments along two main perspectives: learning emotive posts regardless of authors (politicians) and learning emotive posts of specific individual authors. The results are discussed by following this distinction. As to the first perspective, we considered 2325 annotated posts (training set), equally distributed over the authors under examination. In the second perspective, we considered the posts posted by five specific authors[5] and built the annotated training sets with 159, 322, 1002, 1486 and 1069 posts. The performance has been evaluated in terms of classification accuracy on the testing sets.

The datasets are structured as follows: Dataset D1 is composed of 119 users whose posts have been labelled using the seven emotional tags while datasets D2-D6 are single-user datasets which represent the data of every single user over the period of reference. We decided to measure the classification accuracy for both cases: multi-user context and single-user context so that we could understand how well algorithms perform in both different scenarios.

In addition, we decided to investigate the effect of linguistic processing such stemming for Albanian, on the classification accuracy. For this purpose we have used the algorithm developed in [18], which uses a rule-based approach for stemming of texts in Albanian language. The experiments are divided into two groups, the first without stemming and the second with stemming.

---

[2] https://keras.io/.
[3] https://www.tensorflow.org/.
[4] https://github.com/matheuss/google-translate-api.
[5] We do not report their names due to privacy reasons.

Table 2. Experimental results in terms of accuracy (correctly classified stemmed instances).

Dataset ID	Deep Learning	NB	IBK	SMO
D1	**92.4%**	66.7%	73.6%	76.1%
D2	**84.5%**	81.6%	78.5%	82.3%
D3	**83.1%**	77.9%	81.0%	81.0%
D4	67.0%	78.1%	80.0%	84.8%
D5	**91.0%**	83.5%	90.3%	90.3%
D6	**85.3%**	65.4%	75.4%	75.3%

Experimental results are shown in Tables 1 and 2. We have used algorithms implemented in Weka for the other three classical machine learning classification algorithms, such as Naive Bayes (NB), Instance-based learner (IBK), and Support Vector Machines (SMO).

Table 1 shows the results on texts without being stemmed, with 10-fold cross-validation. As we can see, the DL approach has produced better results in terms of classification accuracy. Only in one case the DL approach has scored worse than the other algorithms. We suspect this worse result may be due to shorter sentences in this dataset in general. However, this case needs further investigation which we leave for future work as it requires extensive experiments by varying the length of the sentences to understand how this relates to the performance of the various algorithms.

Table 2 shows results on stemmed texts. As it can be seen, again the DL approach performs in general better than the other classical algorithms. Again only in one case the DL approach has performed worse which leaves again room for future exploration of the reasons of this specific result.

Another interesting finding of the experiments, is that using the stemming step, did not lead to relevant improvement in the overall performance of all the algorithms, with some worse results in some cases. We suspect this is due to short length of the sentences which after being stemmed carry even less information than before being stemmed. We believe this matter requires further investigation on the relationship between length of the texts being stemmed and the performance of the classifier.

6   Conclusions

In this paper we have presented an approach for analyzing micro-blogging texts and postings in Albanian language. Our approach to emotion analysis is by classifying a text fragment into a set of pre-defined emotion categories and therefore detecting the emotional state of the writer conveyed through the text. We have performed a comparative analysis between different classifiers, using deep learning and other classical machine learning classification algorithms. We have

also used a stemming tool for Albanian language to perform a second round of experiments, preprocessing our datasets. Experimental evaluation shows that deep learning produces overall better results compared with the other methods in terms of classification accuracy. Interesting findings related to the length of the texts being processed and the impact on the classifiers' accuracy, show that this matter requires future investigation. As future work, we plan to investigate three research directions: *(i)* sentence-based classification by considering semantic relations among named entities [13], *(ii)* time-variability of the user emotion [12] and *(iii)* correlation between the emotive status of different users [11].

References

1. Agrawal, A., An, A.: Unsupervised emotion detection from text using semantic and syntactic relations. In: 2012 IEEE/WIC/ACM International Conferences on Web Intelligence, WI 2012, Macau, China, 4–7 December 2012, pp. 346–353 (2012). https://doi.org/10.1109/WI-IAT.2012.170
2. Asghar, M.Z., Khan, A., Bibi, A., Kundi, F.M., Ahmad, H.: Sentence-level emotion detection framework using rule-based classification. Cogn. Comput. **9**(6), 868–894 (2017). https://doi.org/10.1007/s12559-017-9503-3
3. Beci, B.: Gramatika e gjuhes shqipe. Publisher, Logos-A (2005)
4. Biba, M., Mane, M.: Sentiment analysis through machine learning: an experimental evaluation for Albanian. In: Proceedings of the Second International Symposium on Intelligent Informatics, ISI 2013, India, pp. 195–203 (2013). https://doi.org/10.1007/978-3-319-01778-5_20
5. Das, D., Bandyopadhyay, S.: Sentence to document level emotion tagging a coarse-grained study on Bengali blogs. In: Advances in Pattern Recognition - Second Mexican Conference on Pattern Recognition, Mexico, pp. 332–341 (2010)
6. Das, D., Bandyopadhyay, S.: Sentence-level emotion and valence tagging. Cogn. Comput. **4**(4), 420–435 (2012). https://doi.org/10.1007/s12559-012-9173-0
7. Ekman, P.: Facial expression and emotion. Am. Psychol. **48**(4), 384–392 (1993). https://doi.org/10.1037/0003-066X.48.4.384
8. Kim, Y.: Convolutional neural networks for sentence classification. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, 25–29 October 2014, Doha, Qatar, pp. 1746–1751 (2014)
9. Li, Q., Shah, S., Liu, X., Nourbakhsh, A.: Data sets: word embeddings learned from tweets and general data. In: Proceedings of the Eleventh International Conference on Web and Social Media, ICWSM 2017, Canada, 2017, pp. 428–436 (2017)
10. Li, S., Huang, L., Wang, R., Zhou, G.: Sentence-level emotion classification with label and context dependence. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics, ACL 2015, China, pp. 1045–1053 (2015)
11. Loglisci, C.: Time-based discovery in biomedical literature: mining temporal links. IJDATS **5**(2), 148–174 (2013). https://doi.org/10.1504/IJDATS.2013.053679

12. Loglisci, C., Ceci, M., Malerba, D.: Relational mining for discovering changes in evolving networks. Neurocomputing **150**, 265–288 (2015). https://doi.org/10.1016/j.neucom.2014.08.079

13. Loglisci, C., Ienco, D., Roche, M., Teisseire, M., Malerba, D.: Toward geographic information harvesting: Extraction of spatial relational facts from web documents. In: 12th IEEE International Conference on Data Mining Workshops, ICDM Workshops, Brussels, Belgium, 10 December 2012, pp. 789–796 (2012). https://doi.org/10.1109/ICDMW.2012.20

14. Mohammad, S.M., Kiritchenko, S.: Using hashtags to capture fine emotion categories from tweets. Comput. Intell. **31**(2), 301–326 (2015). https://doi.org/10.1111/coin.12024

15. Oleri, O., Karagoz, P.: Detecting user emotions in twitter through collective classification. In: Proceedings of the 8th International Joint Conference on Knowledge Discovery, Portugal, 2016, pp. 205–212 (2016). https://doi.org/10.5220/0006037502050212

16. Piton, O., Lagji, K., Përnaska, R.: Electronic dictionaries and transducers for automatic processing of the Albanian language. In: 12th International Conference on Applications of Natural Language to Information Systems, NLDB 2007, France, 2007, pp. 407–413 (2007). https://doi.org/10.1007/978-3-540-73351-5_38

17. Quan, C., Ren, F.: Recognizing sentence emotions based on polynomial kernel method using Ren-CECps. In: Proceedings of the 5th International Conference on Natural Language Processing and Knowledge Engineering, NLPKE 2009, China, 2009, pp. 1–7 (2009). https://doi.org/10.1109/NLPKE.2009.5313834

18. Sadiku, J., Biba, M.: Automatic stemming of albanian through a rule-based approach. J. Int. Res. Publ. Lang. Individuals Soc. **6** (2012)

19. Shaila, S.G., Vadivel, A.: Cognitive based sentence level emotion estimation through emotional expressions. In: Selvaraj, H., Zydek, D., Chmaj, G. (eds.) Progress in Systems Engineering. AISC, vol. 366, pp. 707–713. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-08422-0_100

20. Williams, G., Mahmoud, A.: Analyzing, classifying, and interpreting emotions in software users' tweets. In: 2nd IEEE/ACM International Workshop on Emotion Awareness in Software Engineering, Argentina, 2017, pp. 2–7 (2017). https://doi.org/10.1109/SEmotion.2017.1

21. Xu, J., Xu, R., Lu, Q., Wang, X.: Coarse-to-fine sentence-level emotion classification based on the intra-sentence features and sentential context. In: 21st ACM International Conference on Information and Knowledge Management, USA, pp. 2455–2458 (2012). https://doi.org/10.1145/2396761.2398665