An Overview of Sentiment Analysis: Levels, Approaches and Challenges

Loukmane Maada [1], Khalid Al Fararni[2], Badreddine Aghoutane[1], Yousef Farhaoui[3], Mohammed Fattah[4]

[1]*IA Laboratory, Science Faculty, Moulay Ismail University, Meknes, Morocco*

[2]*LISAC Laboratory, Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco*

[3]*L-STI, T-IDMS, University of Moulay Ismail, Faculty of Science and Technics, Errachidia, Morocco*

[4]*Image laboratory, Moulay Ismail University Meknes, Morocco.*

Keywords: Sentiment analysis, machine learning, deep learning, lexicon-based

Abstract: Text is a huge data source; it contains opinions, facts, feelings ... basically embedded knowledge. Sentiment Analysis (*S.A*) primary goal is to analyse the text and determine its polarity (positive, negative, neutral). This field of research has been on the rise since the beginning of this century. A lot of approaches, from word preprocessing and embedding techniques to complex big data architecture., have been modelled, tested and proposed. This paper provides an overview of the different sentiment analysis approaches, namely the traditional machine learning approach, the deep learning approach, the lexicon-based approach, and the hybrid approach. In addition, a brief insight into the challenges *S.A* faces and some proposed solutions are displayed.

1 INTRODUCTION

The emergence of Web 2.0 has changed how people use the internet, from a read-only paradigm to an interactive paradigm. This digital revolution allowed the users to express their opinions and feelings about different topics: politics, finance, health and more; Especially with the emergence of social media, micro-blogging websites with millions of users generate an enormous amount of data per day. For instance, about 500 million tweets are sent out per day. This embedded knowledge within social media drew the attention of researchers; however, the classical NLP methods were suboptimal, leading to the development of an NLP sub-field known as sentiment analysis.

Sentiment analysis, also known as opinion mining or emotion A.I, is a set of analytic methods that extract and identify information (sentiment, opinion, and attitude) from text using natural language processing, text analysis, and computational linguistics. This information is used to analyse customers ' satisfaction with a product. Or to predict particular behaviours, e.g. forecasting the election winner.

Opinion mining has been used in different domains and multiple real-world applications. From predicting the stock market (Chiong Raymond et al.,2018) to detect event popularity (Mariana Daniel et al.,2016) and predicting the election (F. Nausheen et al.,2018); to fighting Covid-19 (A.H. Alamoodi et al., 2020), increasing the benefits of tourism‘s company (Gianpierre Zapata et al., 2017) and to improve tourists' experience (L.Maada et al., 2021).

Our aim is to build a big data solution based on hybrid recommendation and sentiment analysis using machine and deep learning techniques to recommend the most suitable tourist offer, improve the customer experience and forecast the tourist demand in Morocco (K. AL Fararni et al., 2021).

[1] https://orcid.org/0000-0003-4165-1486
[2] https://orcid.org/0000-0001-5907-6948
[3] https://orcid.org/0000-0002-9555-6786
[4] https://orcid.org/0000-0001-6128-9715
[5] https://orcid.org/0000-0003-0870-6262

Maada, L., Al Fararni, K., Aghoutane, B., Farhaoui, Y. and Fattah, M.
An Overview of Sentiment Analysis: Levels, Approaches and Challenges.
DOI: 10.5220/0010735700003101
In *Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning (BML 2021)*, pages 413-418
ISBN: 978-989-758-559-3

This paper gives an overview of sentiment analysis, which is beneficial for newcomer researchers. It provides a comprehensive summary of the different sentiment analysis levels(sentence level, document level, and aspect level). Followed by the basic approaches of sentiment analysis(machine learning approach, lexicon-based approach, hybrid approach, and deep learning approach), then the paper will expose the challenges sentiment analysis faces and some proposed solutions.

2 LEVELS OF SENTIMENT ANALYSIS

Sentiment analysis has been investigated on three different granularity levels: sentence level, document level, and aspect level.

2.1 Sentence-level

This level focuses on the sentences in a text and tries to extract opinions from them. It first classifies the sentence as either objective or subjective. Then it extracts the polarity (positive or negative) of each sentence, assuming that objective sentences are always neutral. However, objective sentences are not always neutral and may include some opinions. For example, "My phone broke yesterday," which contains an implicit negative sentiment (Singh Nikhil Kumar et al., 2020); we refer to this sort of sentiment as a mild sentiment (Zimba David et al., 2018).

2.2 Document-level

At this level, we suppose that the whole document contains one opinion; in other words, it assumes that each document has a single entity opinion. Hence, Sentiment analysis at this level does not apply to documents that contain multiple entities (Zimba David et al., 2018).

2.3 Aspect-level

This level is considered to be the hardest out of the three levels. It assumes that every opinion is dependent on an aspect and has a target which without it is meaningless (Zimba David et al., 2018). Therefore the aspect level will aim to extract the entities and their respective aspects. To illustrate the sentence "The PS5 design is great, but it's costly", evaluate two parts: the great design and the high price.

3 SENTIMENT ANALYSIS APPROACHES

In this section, we will display the different approaches to sentiment analysis. This section will be divided into four subsections: the lexicon-based approach, the machine learning approach, the hybrid approach, and the deep learning approach. Each of those subsections provides an overview of different techniques that those approaches use.

3.1 Lexicon-based Approach

The lexicon-based approach is based on using a corpus or dictionary to determine the polarity. This approach requires a high-quality sentiment lexicon to yield good results. There are mainly two types of lexicons: The general-purpose lexicon which, as its name suggests, is used for general-purpose classifiers, and the domain-specific lexicon used for the domain-related classifiers. In this section, we will display the lexicon-based approaches and some of the most used worldwide lexicons.

3.1.1 Manual Approach

This approach utilises an existing dictionary or corpus, then associates the sentiment strength to each of the sentiments. For example, if we consider a 3-class sentiment strength {-1,0,1} the words "good" and "bad" will be associated respectively with the sentiment strength 1 and -1. The lexicon's advantages are that it works well and is quite accurate. In contrast, this method has some drawbacks. It is very time-consuming, cannot match the language's evolution, and is not qualified for specific domains. Generally speaking, rather than its direct use. The manually developed lexicon is used to blend with other methods.

3.1.2 Dictionary Approach

The dictionary approach was first presented by (Hu Minqing et al., 2004). It consists of building a small dictionary with a few sentiment words whose polarity has been manually set. The number of words is increased iteratively by gathering antonyms and synonyms from well-known lexicons. e.g. WordNet or its enhanced version SentiWordNet.

3.1.3 Corpus Approach

The corpus approach, like the dictionary approach, starts with a small set of words;

then increases among the large corpus according to some specific rules or formulas. The major difference between the two methods is that the corpus approach can find opinion words within a particular context. There are two types of corpus-based techniques: statistics-based (PMI (Kenneth Ward Church et al., 1989), LSA (Thomas K Laudauer et al., 1998)...) and semantic-based approaches when it comes to generating techniques.

3.1.4 Lexicon Examples

WorldNet: is a cognitive linguistics-based English dictionary suggested by (Miller et al.,1990). It associates words with semantic relations such as synonyms, hyponyms, and meronyms. The synonyms are organised into synsets, each with a brief definition and usage examples.

SentiWorldNet: is a lexical resource for opinion mining based on WordNet. It classifies the word into positive, negative, and neutral. SentiWordNet takes into consideration that a world may have different polarities. For instance, "cool" may mean having a low temperature as in a cool afternoon or calmness as in he recovered his cool. SentiWord uses glosses for each word entry to distinguish one from another (Esuli Andrea et al., 2006).

WordNet-Affect: is a hand-curated collection of emotion-related words (nouns, verbs, adjectives, and adverbs), divided into 28 subcategories ("Joy", "Love", "Fear", etc.) and classified as "Positive", "Negative", "Neutral", or "Ambiguous".

3.2 Machine Learning Approach

The machine learning approaches consist of two categories that are supervised learning and unsupervised learning. In this section, we will be exposing the widely used algorithms in both categories.

3.2.1 Supervised Learning

Supervised learning is a set of learning algorithms trained based on labelled data. In the following, we will briefly expound on the widely used techniques in sentiment analysis.

Naive Bayes is a probabilistic model based on the Bayes formula Eqs.1 with an independence assumption between the entities(features).

$$P(A/B) = \frac{P(A)*P(B/A)}{P(B)} \qquad (1)$$

The Naive Bayes (N.B) or simple Bayes is a fast, robust, and simple method that is patriotically successful in NLP.(Fitri et al., 2018) studied the customer's satisfaction with cellular data service using N.B algorithm on a Tweeter-based data set. They reached on average 94,5% precision, 93,31% recall, F1-score of 93,15% and accuracy of 99,09%. (Risky Novendri et al., 2020) worked on analyzing YouTube comments using the N.B method with which they achieved 81%, 74.83%, and 75.22% for accuracy, precision, and recall, respectively.

Support Vector Machine or SVM is a binary classification method that determines the optimal hyperplane separating the data to positive and negative in S.A. The SVM proceeds by maximising the margin between the training data. If the training set is not linearly separable, it is mapped to a higher dimension using the kernel function. (Lavanya and Deisy, 2017) present a multi-class SVM model to analyse customer's satisfaction with the company's(Google, Apple, Microsoft...) products using Tweets. The model reached an accuracy of 96%, a recall of 83%, and an F1-score of 88%. In the tourism field, (Laoh et al., 2019) used SVM with n-gram to examine the Bali hotel review, they achieved the best result using the 2-gram method combined with SVM, 94%, 75%, 72%, 74% for accuracy, precision, recall, and F1-measure respectively.

Maximum entropy is a probability distribution technique widely used in NLP. It consists of finding the distribution that satisfies the constraints and should have the maximum entropy. Consider the following data set $T = \{(x_1, y_1), (x_2, y_2\}),..., (x_n, y_n)\}$ and feature function $f_i(x, y)$, $i = 1, 2,..., n$. The equations that make up the model are as follows:

$$p_{maxE(y/x)} = \frac{exp(\sum_i \lambda_i f_i(x,y))}{Z_x} \qquad (2)$$

where $\lambda_i$is the weight parameter of the feature function and

$$Z_x = \sum_y exp(\sum_i \lambda_i f_i(x,y)) \qquad (3)$$

(Dewanti Putri et al., 2019) led a study on the Grab user reviews in which they utilised the SVM algorithm and the MaxE algorithm. The results were close, with the SVM reaching a maximum accuracy of 89%, whereas MaxE reached 90.46%.

3.2.2 Unsupervised Learning

Unlike supervised learning that heavily relies on labelled data, unsupervised learning methods detect patterns in the data set without pre-associated labels.

K-means clustering is a method that aims to partition the data set into a k cluster in which each data point belongs to the cluster with the nearest mean. (Riaz et al., 2019) used the K-means clustering to partition the words sentiment strength into three clusters, positive, negative, and neutral. The review analysis technique they proposed reached an accuracy of 95%.

Topic modelling is an unsupervised machine learning approach that scans the text, detects the words and sentence patterns, clusters word groups and similar expressions that best describes a set of texts.

(K. Lavanya et al., 2018) suggested a topic modelling-based framework to do multi-aspect S.A on Chinese Online social reviews. They used the Latent Dirichlet Allocation (LDA) to extract the global topic, the local topic (the aspect) and their associated sentiment. The empirical results revealed that the proposed approach enhanced the performance of the model.

3.3 Deep Learning Approach

Deep learning approaches are well known for trying to mimic the human's brain work. The deep learning algorithms were mostly used in image recognition, but with the embedding tools (word2vec and GloVe) emerging, deep learning methods become the research focus in the sentiment field. In this section, we will display a few deep learning algorithms.

3.3.1 Recurrent Neural Network

A recurrent neural network (R.N.N) is a neural network algorithm derived from feedforward neural networks. R.N.N passes information from its previous time step to the current time step, making it capable of processing sequential data like text.

3.3.2 Long Short-Term Memory Model

L.S.T.M model is an extension of the R.N.N model that utilises gates to ease the issues of blow up and vanishing gradients in the R.N.N model. The L.S.T.M, unlike the R.N.N model, discards the useless information (forget gate) and passes the useful one to the next unit.

(Fu et al., 2018) proposed a lexicon-enhanced L.S.T.M model. They trained a word sentiment classifier with a sentiment lexicon to obtain the embedded sentiment of each word, the word embedding, and its relative sentiment is passed as an input to L.S.T.M. The results show that this method enhances the performance. For instance, it improved the accuracy by about 4% on the IMDB dataset.

3.3.3 Convolutional Neural Network

A deep convolutional neural network (D.C.N.N) is mostly applied in the field of visual imagery analysis. However, it has high efficiency when combined with good word representation models.

(Minaee et al., 2019) proposed a framework that combines L.S.T.M and C.N.N models using GloVe embedding representation. They average the prediction from L.S.T.M and C.N.N to get the final prediction. The experience showed a slight gain in accuracy compared to the use of CNN and L.S.T.M independently. For instance, the proposed model achieved an accuracy of 90% on the IMDB data set, whereas the CNN and L.S.T.M alone reached 89.3% and 89%, respectively.

Both CNN and L.S.T.M are one of the most used neural network approaches in the field of sentiment analysis. However, each one of them does a better job at a specific task. C.N.N excels in features extraction, whereas L.S.T.M achieves superior sentiment understanding, hence the combination of both in recent research.

3.4 Hybrid Approach

The hybrid approach is a relatively new approach to sentiment analysis that combines different approaches. It uses each technique's strengths. This method is typically used for complex sentiment analysis tasks.

(Asghar et al., 2018) developed a hybrid technique that incorporates four classifiers: a slang classifier, an emoticon classifier, the SentiWordNet classifier, and an enhanced domain-specific classifier; to improve the performance of Twitter-based sentiment analysis systems. The sentiment classification was performed at the sentence level; the multistage hybrid framework overcame the limitations of previously proposed models.

(Yadav et al., 2019) compared the performance of a hybrid based approach that uses SVM, N.B, and Genetic algorithm (G.A) and each of those algorithms performance on a hotel reviews data set. The hybrid approach achieved a 93% accuracy, whereas SVM alone achieved 85.2%, N.B 85%, and G.A 85.3%.

(Sosa, 2017) combined two neural network models namely, C.N.N and L.S.T.M, in a hybrid model. Experiment results showed that the L.S.T.M-C.N.N model performed better than the average model.

The hybrid approach has shown to be the most effective approach among all existing ones since it combines different approaches to overcome each one's weaknesses and makes the most of their strengths.

4 CHALLENGES

Text is a complicated sequential structure that includes information, making sentiment analysis difficult. In this section, we will explore some of those challenges and some proposed solutions.

4.1 Sarcasm

Sarcasm can be defined as the use of irony to mock or annoy for a humorous purpose. For instance, "The movie plot was so good that I knew who the killer was in the first 5 min :)" this sentence contains first the word "good" and second the happy face emoji; this review has a high probability of being classified as positive whereas it is negative. Much research has been done to detect sarcastic reviews. The state of art performance is about 80% accuracy (N.Majumder, 2019).

4.2 Suggestions, Questions and Advice

Suggestions, questions, and advice are mistaken for either positive or negative sentiment, whereas they should be classified as neutral. For instance, "It would be great if the store installs a better security system", this review is an objective suggestion that is mistaken for a positive sentiment since it contains "better" and "great", which are strong positive words.

4.3 Negation

Negation is a common language technique for expressing displeasure. This technique makes deducing the meaning of narrative content harder. For instance, "I'm not happy with the quality of the product" this review is negative; however, it has a high probability of getting mistaken for positive. Some research has been done in this field; (Pröllochs et al., 2020) proposed a framework based on a reinforcement learning model that detects the parts of the sentence that use negation, this method yields some promising results.

5 CONCLUSION

In this paper, we investigated the levels of sentiment analysis (aspect-level, document-level, sentence-level). We presented the different approaches used in opinion mining: the machine learning approach, the lexicon-based approach, the deep learning approach and finally, the hybrid approach. We listed some of the challenges sentiment analysis faces. We believe that sentiment analysis techniques are on a constant rise; and will offer a better and deeper understanding of complex sentiments in the incoming years.

REFERENCES

Chiong, Raymond and Fan, Zongwen and Hu, Zhongyi and Adam, Marc T. P. and Lutz, Bernhard and Neumann, Dirk, 2018 A Sentiment Analysis-Based Machine Learning Approach for Financial Market Prediction via News Disclosures. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion, page 278–279.*

Daniel, Mariana and Neves, Rui Ferreira and Horta, Nuno, 2017, Company event popularity for financial markets using Twitter and sentiment analysis. In *Expert Systems with Applications, volume 71, page 111–124.*

F. Nausheen and S. H. Begum, 2018, Sentiment analysis to predict election results using Python. In *2018 2nd International Conference on Inventive Systems and Control (ICISC), page 1259–1262.*

Alamoodi, A.H.and Zaidan, B.B. and Zaidan, A.A. and Albahri, O.S. and Mohammed, K.I. and Malik, R.Q. and Almahdi, E.M. and Chyad, M.A. and Tareq, Z. and Albahri, A.S. and Hameed, Hamsa and Alaa, Musaab, 2021, Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review. In *Expert Systems with Applications, volume 167, page 114155.*

Gianpierre Zapata. and Javier Murga. and Carlos Raymundo. and Jose Alvarez. and Francisco Dominguez, 2017, Predictive Model based on Sentiment Analysis for Peruvian SMEs in the Sustainable Tourist Sector. In *Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - KMIS, page 232-240.*

David Zimbra and Ahmed Abbasi and Daniel Zeng and Hsinchun Chen, 2018 The State-of-the-Art in Twitter Sentiment Analysis: A Review and Benchmark Evaluation. In *ACM Trans. Manage. Inf. Syst, Article 5.*

Singh, Nikhil Kumar and Tomar, Deepak Singh and Sangaiah, Arun Kumar, 2020 Sentiment analysis: a review and comparative analysis over social media. In *Journal of Ambient Intelligence and Humanized Computing, volume 11, page 97-117.*

Minqing Hu and Bing Liu, 2004 Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference, page 168-177.*

Kenneth Ward Church and Patrick Hanks, 1989 Word association norms, mutual information, and lexicography. In *Proceedings of the 27th annual meeting on Association for Computational Linguistics, page 76–83.*

Thomas K Landauer and Peter W. Foltz and Darrell Laham, 1998 An introduction to latent semantic analysis. In *Discourse Processes, volume 25, page 259–284.*

Miller, George A. and Beckwith, Richard and Fellbaum, Christiane and Gross, Derek and Miller, Katherine J., 1990 Introduction to WordNet: An On-line Lexical Database*. In *International Journal of Lexicography, volume 3, page 235–244.*

Esuli, Andrea and Sebastiani, Fabrizio, 2006 SENTIWORDNET : A Publicly Available Lexical Resource for Opinion Mining.*Proceedings of the Fifth International Conference on Language Resources and Evaluation.*

F. S. Fitri and M. N. S. Si and C. Setianingsih, 2018 Sentiment Analysis on the Level of Customer Satisfaction to Data Cellular Services Using the Naive Bayes Classifier Algorithm. In *2018 IEEE International Conference on Internet of Things and Intelligence System (IOTAIS), page 201–206.*

Novendri Risky and Callista, Annisa Syafarani and Pratama, Danny Naufal and Puspita, Chika Enggar, 2020 Sentiment Analysis of YouTube Movie Trailer Comments Using Naïve Bayes. In *Bulletin of Computer Science and Electrical Engineering, volume 1, page 26–32.*

K. Lavanya and C. Deisy, 2017 Twitter sentiment analysis using multi-class SVM. In *International Conference on Intelligent Computing and Control (I2C2), page 1–6.*

E. Laoh and I. Surjandari and N. I. Prabaningtyas, 2019 Enhancing Hospitality Sentiment Reviews Analysis Performance using SVM N-Grams Method. In *2019 16th International Conference on Service Systems and Service Management (ICSSSM), page 1–5.*

K. B. A. Dewanti Putri and A. Uswatun Khasanah and A. Azzam, 2019 Sentiment Analysis on Grab User Reviews Using Support Vector Machine and Maximum Entropy Methods. In *2019 International Conference on Information and Communications Technology (ICOIACT), page 468-473.*

Riaz, Sumbal and Fatima, Mehvish and Kamran, M. and Nisar, M. Wasif, 2019 Opinion mining on large scale data using sentiment analysis and k-means clustering. In *Cluster Computing, volume 22, page 7149-7164.*

K. Lavanya and C. Deisy, 2012 Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and HowNet lexicon. In *Knowledge-Based Systems, volume 37, page 186-195.*

X. Fu and J. Yang and J. Li and M. Fang and H. Wang Bo Pang and Lillian Lee, 2018 Lexicon-Enhanced LSTM With Attention for General Sentiment Analysis. In *IEEE Access, volume 6, page 71884-71891.*

Shervin Minaee and Elham Azimi and AmirAli Abdolrashidi, 2019 Deep-Sentiment: Sentiment Analysis Using Ensemble of CNN and Bi-LSTM Models.

Asghar, Muhammad Zubair and Kundi, Fazal Masud and Ahmad, Shakeel and Khan, Aurangzeb and Khan, Furqan, 2018 MT-SAF: Twitter sentiment analysis framework using a hybrid classification scheme. In *MT-SAF: Twitter sentiment analysis framework using a hybrid classification scheme, volume 35, page e12233.*

N. Yadav and R. Kumar and B. Gour and A. U. Khan, 2019 Extraction-Based Text Summarization and Sentiment Analysis of Online Reviews Using Hybrid Classification Method. In *2019 Sixteenth International Conference on Wireless and Optical Communication Networks (WOCN), page 1-6*.

P. M. Sosa, 2017 MT-SAF: Twitter sentiment analysis using combined LSTM-CNN models.

N. Majumder and S. Poria and H. Peng and N. Chhaya and E. Cambria and A. Gelbukh, 2019 Sentiment and Sarcasm Classification With Multitask Learning. In *IEEE Intelligent Systems, volume 34, page 38-43*.

L. MAADA, K. AL FARARNI, B. AGHOUTANE, M. FATTAH, Y. FARHAOUI. A proof of concept web application for sentiment analysis in tourism in the region of Draa-Tafilalet, presented at *the 2nd Edition of International Conference on Big Data, Modelling and Machine learning (BML'21), Kenitra, Morocco, July. 15 – 16, 2021.*

K. AL Fararni, F. Nafis, B. Aghoutane, A. Yahyaouy, J. Riffi, A. Sabri. Hybrid Recommender System for Tourism Based on Big Data and AI: A Conceptual Framework. Big Data Mining and Analytics *2021, 4(1): 47-55*

Nicolas Pröllochs and Stefan Feuerriegel and Bernhard Lutz and Dirk Neumann, 2020 MT-SAF:Negation scope detection for sentiment analysis: A reinforcement learning framework for replicating human interpretations. In *Information Sciences, volume 536, page 205-221*.