Deep Learning Technique of Sentiment Analysis for Twitter Database

https://doi.org/10.3991/ijim.v16i01.27575

Atif M. Gattan(✉)
Faculty of Arts and Humanities, King Abdulaziz University, Jeddah, Saudi Arabia
amgattan@kau.edu.sa

Abstract—Sentiment investigation is the progression of calculating, recognizing as well as classifying the people mentality stated in the outline of the manuscript and the outlook of the individual concerning the particular subject matter can be investigated with the assist of that statistics. The information is accumulated by way of the assistance of API in favor of classifying of the end outcome based on the information collected as in Negative way, Neutral way and Positive way with the support of the scoring polarity allocated in favor of every statement that are collected. These data are fruitful for discovering and enhancing the consumer requirements and to acquire better tune-up. Major improvement of using the concept of sentiment analysis is to develop the user needs by straightforwardly collecting the information from the outsized set of consumers.

Keywords—deep learning, sentiment analysis, twitter, twitter database

1 Introduction

Big data is an extensive term for the non-conventional approaches and procedures desirable to collect, categorize, progression, and assemble insights from outsized datasets. Despite the fact that the trouble of functioning with bulky dataset that exceeds the processing authority or storage of a particular computer is not recent, the occurrence, scale, and worth of this sort of processing has very much expanded in modern years. The fundamental constraints for functioning with big data are the equivalent as the requirements for processing with datasets of any dimension. The objective of the majority of big data systems is to exterior insights and associations from huge volumes of assorted and mixed data that would not be potential using conservative techniques.

Data mining is the development of mining through bulky data sets to discover patterns and found relationships to resolve troubles with the support of data analysis. Tools of Data mining allow activities to estimate future trends in various real time domains. Data mining techniques are applied in numerous research areas, associating mathematics, cybernetics, genetics and marketing. Despite the fact that data mining techniques are a means to make efficiencies and discover the behavior of the customer, if employed appropriately, a production can set itself apart from its struggle through the purpose of predictive investigation.

Sentiment investigation is one of the contextual mining of text content which categorize and extracts subjective details in source material, and serving a business to recognize the social sentiment of their trademark, product or service while observing online discussions. Sentiment Analysis is the major universal text classification tool that investigating a received message and notify whether the essential sentiment is positive, negative or neutral.

2 Investigate twitter sentiments during big data

The rise of social media has generated remarkable concentration along by way of Internet users today. Information from commencing the authentic social community networking websites can be employed for a number of intentions, similar to prediction, promotion or sentiment analysis. Twitter is an extensively used social media website for commenting through short statuses. The millions of twitter comments established each time possibly will be bringing up to sentiment analysis. However managing such a massive quantity of shapeless data is a monotonous mission to engage in. The present and existing tools of analytics and models used that are accessible in the market are not adequate to manage big data. Hence, there is a demand to apply Cloud storage for such category of applications.

With the support of the tool Hadoop, the twitter data can be applied for intelligent investigation and storage of outsized data. The planned method of improved sentiment investigation on tweet comments in the surrounding of cloud is too helpful in determining the opinions of people when it comes to diverse subject matters interrelated to some other field. Further compare the services of various contributors and judge which one is the best. Using Hash tags can endow with a straight forward programmed method to evaluate what people think.

2.1 Deep learning

The vast blooming of the world wide web is incorporated with social networks similar to facebook, twitter, Linkedin, instagram etc. have demonstrate the way to considerable consumers communication and interactions has authorized clients to convey their attitudes in relation to goods, services, proceedings, their favorites along with other users. Furthermore it afforded openings to the users to distribute their understanding knowledge and experiences with everyone. The more rapid enlargement of social media is origin tremendous development of digital substance. Social media have warped online judgments, blogs, tweets, and comments into an extremely expensive asset for the corporate concerns to acquire approaching from the information and prepare their tactic. Production industries associations require to progress and learn the emotional reactions to explore information and to grow in sights of the production industry [1,2]. Established approach to physically mine difficult features, recognize which characteristic is appropriate, and obtain the outlines from its massive data is extremely time overwhelming and need considerable individual endeavors. Technique of Deep Learning demonstrates tremendous concert through NLP approaches to carry

out sentimental investigation on this substantial data. The foundation inspiration of Deep Learning approaches is to be familiar with intricate features removed from the out sized quantity of information exclusive of greatly outside involvement by means of deep learning neural networks. Similar this approach mechanically gain knowledge of novel complex features. Mutually mechanical feature pulling out as well as accessibility of properties is especially essential while contrasting the established ML approach and DL techniques. At this point the ambition is to categorize the outlook and sentiments conveyed by consumers [2–4].

Deep Learning employs prominent and dominant neural network approaches to imitate the manner individual common sense procedure information for converting languages, identifying speech, discovering objects and constructing decisions. Approaches of Deep Learning are accomplished to discover and find out the prototypes from mutually shapeless and unlabeled information exclusive of human involvement. Deep Learning techniques are assists to train through various layers of demonstration and produce ultra modern anticipating consequences. Previously, DL techniques contain extra ordinarily booming in achieving the sentiment investigation. It affords mechanical feature pulling out, wealthy illustration abilities as well as enhanced routine than conventional feature based approaches. These deep-rooted techniques can stipulate physically powerful baselines, and their analytical abilities can be employed in combination with the happening the techniques of deep learning. Commonly there are two kinds of neural networks techniques are referred namely the image processing technique of Convolutional Neural Networks (CNN) and the NLP task of RNN. The concert of DL is employed in the direction of to boost the suggestions based on the sentimental investigation achieved on top of the dissimilar assessments, which are attained since various communal networking websites. Trial assessment achieved point towards the RNN based DL Sentiment Analysis (RDSA) originates the activities through ever growing precision of the sentimental investigation, which it revolve to furnish healthier suggestions by the consumer and thus facilitate in the direction of recognizing specific arrangement as per the constraint of the consumers demand. Figure 1 depicts about the Work outline of ML and DL approaches [5].

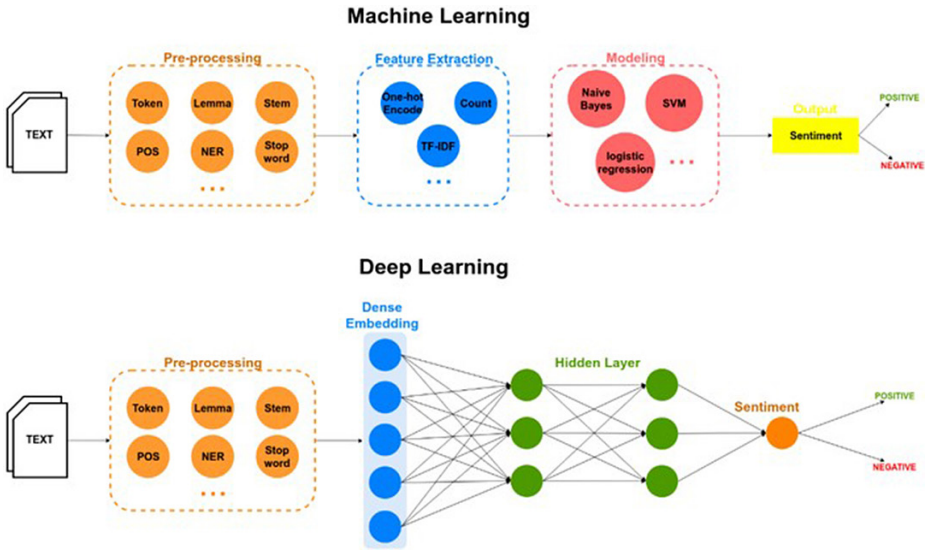

Fig. 1. Work structure of machine learning and deep learning [2]

DL influences the technique of multilayer through the concealed layer of NN. By tradition, the features of machine learning models are acknowledged and discovered through manual process or using the approaches of feature selection. On the other hand, the features of Deep Learning are trained, removed mechanically ensuing in elevated correctness and functionality [6–7].

Deep Learning (DL) and Machine Learning (ML) are occasionally employed interchangeably. Deep learning (DL) is in reality, the techniques of machine learning but it is further superior. Once essential machine learning formulates a blunder mistake, human being input is involved to correct it and to revolutionize the production and "force" the representation to be trained. Within deep learning the perception of neural network can gain knowledge to correct itself all the way through its highly developed algorithm chain. The hierarchical machine learning of Deep learning is to use multiple algorithms in a progressive chain of measures to solve complex problems and permits to deal with substantial amounts of information, accurately and with incredibly little individual effort [8].

The machine learning process will initiates with the text data. These data must be pre-processed with the assistance of unique techniques and further the extraction techniques are applied for predicting the fruitful features. Then the modeling approaches like SVM, Navies Bayes and etc will help to design the algorithm according to the requirements. Finally the sentiment analysis technique is applied for predicting the positive and negative reviews of the contents. Likewise the Deep Learning technique is pertained in the same way, with the difference of including the concepts of neural networks which will assist to train the sentiment analysis. By the part of deep learning accomplishment in most of the other application domains, it vitally used the concept of

sentiment analysis in the modern days. Deep learning is premeditated an advancement of machine learning. It chains collective approaches that aspire to reproduce like the human brains activities using artificial neural networks and have allowed numerous realistic domains of machine learning associating with client support automation like self automatic driving car [9–11].

3 Proposed methodology

Sentiment analysis is a efficient manuscript investigation tool that without human intervention extracts the formless data like social medium customer services and email services are according to emotion and opinion are performed with the assistance of the approaches of DL and ML approaches. In general sentiment investigation be one kind of NLP algorithm working in the direction of establishing the data is either in neutral or whether in positive direction or in negative way. It is regularly carry out on textual data to assist businesses examine product and brand sentiment in feedback of the customer and recognize the customer requirements.

It seems to be a set of approaches or various approaches utilized to discover the sentiment investigation like neutral, negative or positive of a particular specific manuscript. It is one of the great and dominant automated systems of NLP and discovers into the great count according to the industry. It mentioned in the direction of the intention of NL Processing, manuscript examination, calculation linguistics as well as biometric clarification to analytically recognize, remove, count and revise various states and individual information. Most of the times the sentiment investigation goes away from the classification of manuscripts to discover estimations and classifies them as undesirable or desirable, negative or positive [12–14].

3.1 Sentiment analysis

Figure 2 illustrates Researchers have essentially learned the concept of sentiment investigation at three various kinds of granularity levels are explicitly mentioned as document level of sentimental investigation, analysis level, aspect level of sentimental investigation along with sentence level of sentimental investigation [15].

The first category of document level of sentimental investigation catalogsa narrow-minded document like reviews of the product as expressing an entire opinion of negative or positive. It believes the whole document as the essential data unit and imagines that the document is known to be opinionated and enclose opinions about a single attribute like a particular phone. Most efficient sentiment classification is Sentence level, which catalog personage sentences in a document. On the other hand, every text could not be unspecified to be narrow-minded [16].

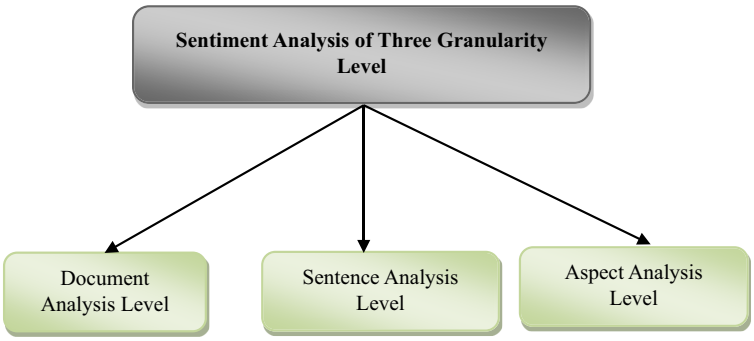

Fig. 2. Three granularity levels of sentimental investigation

By tradition, a particular sentence is classified as narrow-minded or not open-minded, which is referred preconception taxonomy. After that resultant narrow-minded concepts are cataloged as communicating negative judgment or positive judgment. Sentence granularity level of sentiment cataloging can also be originated by the concept of three class categorization trouble, it means, to sort a sentence as neutral view point, negative view point or positive view point. Evaluating with the support of manuscript sentiment analysis level as well as sentiment investigation through sentence level, sentiment investigation of aspect level or aspect-based sentiment examination be further excellent-grained. Its chore is in the direction of take out in addition to recapitulate the attitude of the people articulated on things and characteristics or features of entities which mentioned as objectives. In support of illustration, in an assessment of the products, it aspires to recapitulate negative estimations and positive estimations taking place due to dissimilar directions of the productions correspondingly, though the common emotion on the product could be negative direction or positive direction. The complete mission of aspect supporting sentiment investigation includes of plentiful sub processes similar to feature mining, entity mining, and aspect sentiment cataloging. For instance, based on the complete sentence, "the voice quality of iPhone is great, but its battery sucks", attribute removal must recognize "iPhone" as the attribute, and aspect mining ought to categorize that "voice quality" as well as "battery" are two phases. The direction of sentiment cataloging should categorize the reaction conveyed on the quality of the iPhone's voice as positive way as well as on the battery of the iPhone as in negative direction. Notice that for straightforwardness, in most of the algorithms aspect is feature removal and object extractions are come together and are described as extracting the aspects of sentiment [10–12].

ALGORITHM OF SENTIMENT ANALYSIS

STEP 1 – Analyze the Tweet in Twitter Database

STEP 2 – Collecting Tweet data from Twitter

STEP 3 – Remove the Twitter Handles using Natural Language Processing (NLP). It takes two arguments, among the former argument is original string of text and the later one is original string of text.

STEP 4 – Removing short words having length three or less.

STEP 5 – Using Tokenization the entire string is split a string of text of tokens.

STEP 6 – Stemming is a rule based process of stripping the suffix.

STEP 7 – Calculating polarity which to helps the dataset to discover whether positive or negative. And the outcome is divided into three kinds of processes respectively.

a) Matplot Lib ⟶ Plotting results in the form of graphs.
b) Word cloud ⟶ Represent Most used words
c) Text Blob ⟶ Live Natural Language Processing from Tweet

4 Experimental result

The following figures (Figure 3 and Figure 4) depicts about the sentiment analysis of Tweet database. Every tweet comments are assembled in the form of unpleasant with the consideration of low confidence levels and pleasant with the consideration of high confidence levels. The feature of unpleasant category is partitioned into the emotions of upset, stressed, nervous, tense, active, bored, depressed, unhappy and sad. Likewise, the features of pleasant type of emotions are divided into the people alert, excitement, elated, happiness, comfortable, peaceful as well as quiet, stress-free in addition to calm. Table 1 explains about the category of sentiment analysis similar to unpleasant and pleasant with the emotions which explained below.

Table 1. Category of sentiment analysis

Category of Sentiment Analysis	Emotions	Outcome
Unpleasant Investigation	Upset, stressed, nervous, tense, active, bored, depressed, unhappy, sad	Negative
Pleasant Investigation	Attentive, excitement, elated, happiness, comfortable, peaceful as well as quite, stress-free in addition to calm	Positive

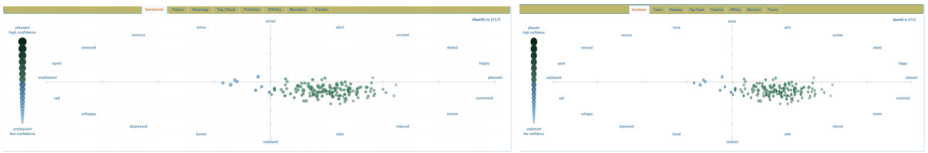

Fig. 3. Sentiment analysis of tweet database & unpleasant and pleasant twitter database in sentiment analysis

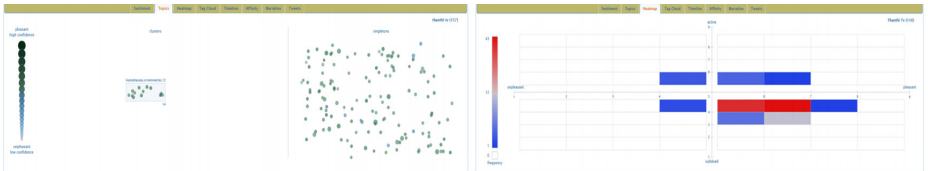

Fig. 4. Unpleasant and pleasant clusters & heatmap of unpleasant and pleasant analysis

Especially, in Figure 5 depicts about the time line of various emotions of the people like relaxed, happy, unhappy and sad for a particular tweet in various schedule period. Different color denotes various emotions of the specific timeline of the particular month period. The purpose of this sentiment analysis is to understand the people mentality, according to the study will lead to improve the specific activity, event or performance of the product. This analysis will help the human like to understand the current scenario and trend of the business for further improvements.

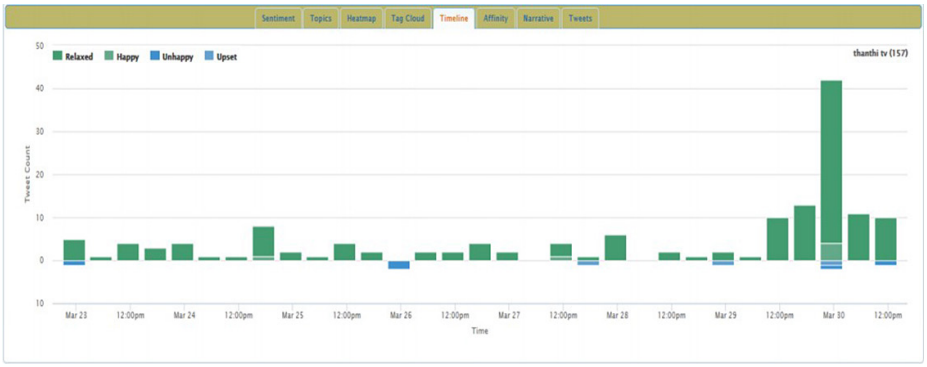

Fig. 5. Timeline of various emotions

5 Conclusions

An innovation method of sentimental investigation is according to the regent news activities which growing through the communal society network of massive, big data. The fore most objective is to progression on news information and discovers the reaction from this information in the form of human emotional. Able to discover this emotional level from current database furthermore will investigate the information commencing the bulky databases and present in the graphic and illustrative representation. Furthermore, the tweet databases are using the API of various websites to investigate the information. In future, can able to include the novel element of storing the hunt history of public community in favor of the benefits of the citizens in comprehending the obstacle further professionally and acquire the existent time sentiment analysis statistics from the community human being.

6 References

[1] Joshi, A., Balamurali, A. R., and Bhattacharyya, P. (2010). A fall-back strategy for sentiment analysis in a new language: A case study for Hindi. In International Conference on Natural Language Processing (pp. 1081–1091).

[2] Go, A., Bhayani, R., and Huang, L. (2009). Twitter sentiment classification using distant supervision. CS224N project report, Stanford, 1: 2009.

[3] Agarwal, A., and Bhattacharyya, P. (2005). Sentiment analysis: A new approach for effective use of linguistic knowledge and exploiting similarities in a set of documents to be classified. In Proceedings of the International Conference on Natural Language Processing (ICON), June 2005 (Vol. 22).

[4] Kennedy, A., and Inkpen, D. (2006). Sentiment classification of movie reviews using contextual valence shifters. Computational Intelligence, 22: 110–125. https://doi.org/10.1111/j.1467-8640.2006.00277.x

[5] Aone, C., Okurowski, M. E., Gorlinsky, J., and Larsen, B. (1990). A trainable summarizer with knowledge acquired from robust NLP techniques. Advances in Automatic Text Summarization, 71–80.

[6] Araque, O., Corcuera-Platas, I., Sánchez-Rada, J. F., and Iglesias, C. A. (2017). Enhancing deep learning sentiment analysis with ensemble techniques in social applications. Expert Systems with Applications, 77: 236–246. https://doi.org/10.1016/j.eswa.2017.02.002

[7] Dang, N. C., Moreno-García, M. N., and De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. Electronics, 9: 483. https://doi.org/10.3390/electronics9030483

[8] Ghorbani, M., Bahaghighat, M., Xin, Q., and Özen, F. (2020). ConvLSTMConv network: A deep learning approach for sentiment analysis in cloud computing. Journal of Cloud Computing, 9: 1–12. https://doi.org/10.1186/s13677-020-00162-1

[9] Goularas, D., and Kamis, S. (2019, August). Evaluation of deep learning techniques in sentiment analysis from twitter data. In 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML) (pp. 12–17). IEEE. https://doi.org/10.1109/Deep-ML.2019.00011

[10] Preethi, G., Krishna, P. V., Obaidat, M. S., Saritha, V., and Yenduri, S. (2017, July). Application of deep learning to sentiment analysis for recommender system on cloud. In 2017 International Conference on Computer, Information and Telecommunication Systems (CITS) (pp. 93–97). IEEE. https://doi.org/10.1109/CITS.2017.8035341

[11] Wang, Z., and Fey, A. M. (2018). Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery. International Journal of Computer Assisted Radiology and Surgery, 13: 1959–1970. https://doi.org/10.1007/s11548-018-1860-1. Da Xu, L., He, W., & Li, S. (2014). Internet of things in industries: A survey. IEEE Transactions on industrial informatics, 10: 2233–2243.

[12] Qaiser, S., Yusoff, N., Ahmad, F. K., and Ali, R. (2020). Sentiment analysis of impact of technology on employment from text on twitter. International Journal of Interactive Mobile Technologies, 14(7). https://doi.org/10.3991/ijim.v14i07.10600

[13] Saeed, S., Shaikh, A., and Memon, M. A. (2018). Impact of social networking sites on personality & attitude of young adults (Research covering the young adults' lives within Korangi, Karachi). International Research Journal of Arts & Humanities (IRJAH), 46(46).

[14] Khalafat, M., Alqatawna, J. F. S., Al-Sayyed, R., Eshtay, M., and Kobbaey, T. (2021). Violence detection over online social networks: An Arabic sentiment analysis approach. International Journal of Interactive Mobile Technologies, 15(14). https://doi.org/10.3991/ijim.v15i14.23029

[15] Shaikh, A., Ali, S., Memon, N., and Karampelas, P. (2010). SOA security aspects in web-based architectural design. In From Sociology to Computing in Social Networks (pp. 415–430). Springer, Vienna. https://doi.org/10.1007/978-3-7091-0294-7_22

[16] AbdelFattah, M., Galal, D., Hassan, N., Elzanfaly, D., and Tallent, G. (2017). A sentiment analysis tool for determining the promotional success of fashion images on Instagram. International Journal of Interactive Mobile Technologies, 11(2), 66–73. https://doi.org/10.3991/ijim.v11i2.6563

7 Author

Atif M. Gattan, Information Science Department, Faculty of Arts and Humanities, King Abdulaziz University, Jeddah, Saudi Arabia. E-mail: amgattan@kau.edu.sa

Article submitted 2021-10-15. Resubmitted 2021-11-28. Final acceptance 2021-11-29. Final version published as submitted by the authors.