

## A Proportional Sentiment Analysis of MOOCs Course Reviews Using Supervised Learning Algorithms



Anuj Kumar Singh<sup>1</sup>, Sandeep Kumar<sup>1</sup>, Shashi Bhushan<sup>2\*</sup>, Pramod Kumar<sup>1</sup>, Arun Vashishtha<sup>1</sup>

<sup>1</sup> Department of Computer Science & Engineering, Krishna Engineering College, Ghaziabad 201007, India

<sup>2</sup> School of Computer Science, University of Petroleum & Energy Studies, Dehradun 248001, India

Corresponding Author Email: [sbhushan@ddn.upes.ac.in](mailto:sbhushan@ddn.upes.ac.in)

<https://doi.org/10.18280/isi.260510>

### ABSTRACT

**Received:** 15 September 2021

**Accepted:** 13 October 2021

#### Keywords:

*sentiment analysis, MOOC, Naive Bayes, SVM, logistic regression, multilayer perceptron*

When anyone is looking to enroll for a freely available online course so the first and famous name comes in front of the searcher is MOOC courses. So here in this article our focus is to collect the comments by enrolled users for the specified MOOC course and apply sentiment analysis over that data. The significance of our article is to introduce a proficient sentiment analysis algorithm with high perceptive execution in MOOC courses, by seeking after the standards of gathering various supervised learning methods where the performance of various supervised machine learning algorithms in performing sentiment analysis of MOOC data. Some research questions have been addressed on sentiment analysis of MOOC data. For the assessment task, we have investigated a large no of MOOC courses, with the different Supervised Learning methods and calculated accuracy of the data by using parameters such as Precision, Recall and F1 Score. From the results we can conclude that when the bigram model was applied to the logistic regression, the Multilayer Perceptron (MLP) overcomes the accuracy by other algorithms as SVM, Naive Bayes and achieved an accuracy of 92.44 percent. To determine the sentiment polarity of a sentence, the suggested method use term frequency (No of Positive, Negative terms in the text) to calculate the sentiment polarity of the text. We use a logistic regression Function to predict the sentiment classification accuracy of positive and negative comments from the data.

## 1. INTRODUCTION

Sentiment Analysis is a classification process through which we can easily determine or classify the data that focuses on Polarity (Positive, negative, neutral), Feelings (Happy, sad, fear), Intentions (Important, aloof, urgency) by using Natural languages techniques. These parameters present inside any textual text, feedback, comments etc. are very beneficial for the common persons for various different field. Massive open online courses (MOOC) are available as an online learning platform, which has gotten a lot of attention for the users. When they attend the courses, many students write comments in the commenting section. These comments not only assess the quality of the curriculum, but also provide direct input on some technical issues with the MOOC platform. There are a number of researches on student reviews about courses, but none of them provide a comprehensive study from the standpoint of sentiment analysis. Student reviews are suggested as a data source for sentiment analysis tasks in this article. We offer a method for examining the compatibility of users' attitudes with artificial sentiment-detection techniques. Students can read the cumulative thoughts of millions of people on a variety of subjects [1]. They can likewise take a gander at the ubiquity record, which is determined utilizing a calculation that thinks about client surveys just as other openly accessible sources like manuals and paper articles. This list runs from 1 to the all-out number of courses accessible. Understudies can look for the most well-known courses or the most fascinating courses. The air pocket rating (client rating),

a 1–5 scale with one air pocket addressing a helpless encounter and five air pockets addressing an extraordinary encounter, is linked to this. This scale is to be used by all reviewers to summaries their feedback. Users can also add their comments to this rating, which can cover a course's performance. As a result, reading and evaluating evaluations might aid in the growth of a company. The technique for deciding a message's opinion is known as feeling examination. Individuals leave remarks via web-based media about occasions they've joined in, and they're interested if most of others had a great or negative involvement in a similar occasion. Investigating opinion is the demonstration of deciding a client's sentiments on something particular, which could be a new occasion, issue, or person [2]. Opinion examination can be performed on three levels: sentence, angle, and archive. Twitter is an awesome asset for deciding the nature of any item. To express opinions, the twitter platform employs tweets, which are written in sentence form. Sentiment analysis at the sentence level is thus used to investigate sentiments. Sentiment analysis has been completed. Companies may use twitter to find out what consumers think about their newly released items. The design is to decide how exact sentences taken from tweets' message are as far as feeling. Twitter information can be exposed to opinion investigation, which orders tweets as great or negative. From the tweets, this examination helps concerned associations in deciding individuals' contemplations about their item, occasions, etc. The most difficult component of sentiment analysis is determining the opinion words. Depending on the situation, an opinion term might be

favorable or negative. Traditional text processing methods will not change the meaning of the information if the words are changed slightly. When there are modifications in two words, however, sentiment analysis can alter the content's meaning. The terms "the phone is ringing" and "the phone is not ringing" are not interchangeable. Sentence-by-sentence analysis is carried out. The user can understand the informal statement in twitter, but not the system.

## 2. LITERATURE SURVEY

The paper [1] explain the methods where data is fetched from Twitter for feeling examination and were sorted by the creator into two classifications: AI approaches and expression recurrence-based methodologies. SVM and NB have been analyzed as far as SA execution. At the point when bigram models were utilized, SVM created the most elevated exactness of 85%. Diverse approaches to perceiving feeling in twitter information can measure up to find the ideal approach. The authors [2] proposes techniques for grouping message audits dependent on the opinion communicated in the reviews. For SA, the directed AI calculations SVM and Nave Bayes were analyzed. The exactness of Naive Bayes was 65%, which was higher than that of the help vector machine. To distinguish the assessments of text in a web-based media organization, abstract examination or extremity estimations are utilized. Opinion examination has been considered in the paper [3] utilizing either feeling dictionaries or AI strategies. In dictionary-based opinion examination, the semantic direction of sentences, words, or still up in the air for each sentence. The AI approach, then again, utilizes Naive Bayes and SVM calculations to order the text. W-Lexicon-based WSD's strategy has demonstrated to be more exact. In human-named archives, the creators [4] found that dictionary-based calculations need next to no work. They infer that utilizing the bigram model in SVM conveys more noteworthy opinion exactness when contrasted with different models, and that all the more spotless information prompts more precise outcomes. The creator [5] classifies the film survey as certain or negative utilizing AI calculations like Nave Bayes, support vector machine, and irregular woods. The hyper parameter further develops SVM and irregular woods model precision. The creator [6] found that feeling examination is useful in an assortment of disciplines, including administration, security, and others, where it very well might be used to execute different undertakings at different levels. It's a decent method to sort out how precise feeling examination AI strategies are. Since the idea of the dataset used varies relying upon the undertaking, existing appraisal gauges seldom give an unmistakable image of how compelling every strategy is in contrast with others [7].

The authors [8] devised a basic strategy to solve a big problem by dividing it into small modules and then solve them. As per his research, sentences were divided into different types and the examination was carried out separately on every different domain sentence. Due to the complexity of the words, a neural network-based concept was used to classify the opinions into three groups based on the number of targets in each sentence.

A classification technique [9] based on the Vote algorithm to investigate the benefits of multiple classifier systems (MCS) on the Turkish sentiment classification problem. Three classifiers were used in the algorithm: NB, SVM, and Bagging.

Although the classification performance is excellent, the system is costly and takes a long time to process huge datasets.

The authors [10] has suggested a technique for determine the polarity and negation of statements. On the basis of the System's summary result, the user can make judgements. The AIRC sentiment analyzer system was outperformed by the sentiment orientation methods. Sentiment Analysis is crucial for those who have viewed other's person's opinion about a product or service. Users can use this to organize the benefits and drawback of their products or service. The authors [11] has suggested that Forum messages in MOOCs (Massive Open Online Courses) are the main source of data about the social connections occurring in these courses. Forum messages can be breaking down to distinguish examples and students' practices. Especially, Sentiment Analysis (e.g., order in positive and negative forum messages) can be utilized as an initial step for distinguishing complex emotions. The point of their work is to analyze diverse ML algorithm for sentiment Analysis, utilizing a genuine contextual analysis to check how the outcomes can give data about students' feelings or examples in the MOOC.

## 3. SUGGESTED WORK

Machine learning techniques such as Logistic regression, SVM, and Multinomial NB are used in Sentiment Analysis to classify sentences in tweets. The performance of four machine learning methods, including logistic regression, SVM, NB and Multilayer Perceptron (MLP) Classifier is evaluated. Sentiment analysis procedures are algorithms that have been trained to detect sentiment polarity [4-6] that can recognize sentiments in texts, sentences, or words automatically. A wide range of sentiment analysis procedure algorithm are available to address the various types of texts. However, analyzing emotions isn't a perfect science, especially when it comes to the unstructured writings that dominate social media [7]. Because human language is complicated, teaching a computer to recognize numerous types of grammatically nuances, cultural variances, jargon, and misspellings in network messages is a tough task, and it's even tougher to automatically comprehend how context affects the tone of the message. We can intuitively grasp the intentionality of any writing because humans can use contextual awareness [9]. Computers, on the other hand, have trouble deciphering the context in which a statement is presented and determining whether or not a person is being sarcastic.

### 3.1 Data preprocessing methodology

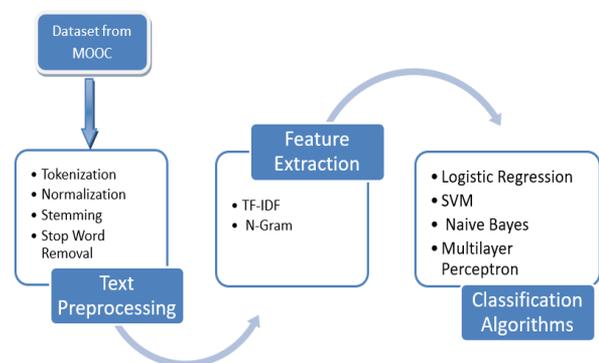


Figure 1. Method of data processing

Various phases are involved in the sentiment analysis of a MOOCS Reviews data set. Data preprocessing is done in several steps as per Figure 1. In step one Dataset has been taken from MOOC, in step second preprocessing is done, in step 3 feature extraction is done and at last in step 4, classification is done by using classification algorithms.

### 3.2 Creating dataset

Dataset creation is required due to the change in the language of the tweets which differs from conventional English. The words may be spelled differently and used in different ways. Because it is a supervised learning technique, it requires a training set to determine the tweet's polarity. It can be done with the help of a Twitter API that collects tweets automatically and manually annotates them as good or negative [11]. There is a slew of options for gathering Twitter data for re-search. Kaggle, Data world, UCI, ASU, Open Knowledge Labs, Kdnuggets, and SNAP are examples of data repositories. Twitter APIs are divided into two categories: search APIs and stream APIs. The search API is used to acquire user Twitter data based on hashtags, and the stream API is used to get real-time Twitter data. Twitter's official archive download, BirdSong Analytics, Cyfe, NodeXL, and TWChat are all tools for downloading Twitter data.

### 3.3 Data preprocessing

Data Preprocessing is required to convert the raw data in a desired format. Preprocessing of tweet is required to be finished before feature extraction. The usage of jargons, URL removal, and misspelt terms, this is done. A dictionary is kept to overcome problems caused by the use of slang words, which substitutes the expression with its corresponding denotation. The tweets are preprocessed before being sent to the classifier.

- If a term contains character repeats, the repeated character is eliminated. LOOOL, for example, is shortened to LOL.
- Altogether whole transcript in the tweets is in lower subscript.
- Arguments like "isn't," "haven't," and "doesn't" are substituted with "is not," "have not," and "does not," and so on.
- The representational language is understood to ordinary fix text, such as @username being replaced with AT USER. When the classifier encounters occurrences of AT USER, it is clear that a user is being referred to.
- Because every hot subject in a tweet begins with a hashtag, the preprocessor strips out the hashtags and maintains the word. #fun, for example, is transformed into fun.
- In the preprocessing step, all white spaces are deleted in order to remove noise that could impact the classifier's performance, and many more processes are conducted.

### 3.4 Feature extraction

There are two processes to feature extraction. The extraction of Twitter-specific features is the initial phase. Emoticons are divided into two categories: positive and negative. Positive emotion has a weight of one, while negative emotion has a weight of one. Positive and negative hashtags

are even used in the hashtags [12]. The positive keyword, negative keyword, and different terms that express negation are all kept on a list. When there are both good and negative keywords, all keywords cannot be treated equally. To fix this, we use a specific term in the tweet that only contains positive or negative keywords. The relevant segment of speech is then identified using a search. Keywords that exhibit greater emotion include adjective, verb, and adverb. If a part of speech is detected for a keyword, it is handled as a special keyword. In addition, it can be chosen at random. If a tweet includes both positive and negative terms, a keyword with a related part of speech is chosen. None is picked if it contains both. A special keyword's weight is allocated to "1" if it is positive, "- 1" if it is negative, and "0" otherwise.

## 4. SENTIMENT CLASSIFICATION

### 4.1 Logistic regression algorithm

Logistic Regression is a classification problem-solving approach that uses predictive analysis. It is based on the probability idea. The Sigmoid function is a more complicated cost function used in Logistic Regression. The logistic regression suggestions suggests that the cost function be incomplete to a value between 0 and 1.

Logistic Regression is a function that predicts the sentiment classification accuracy using phrase frequency integer vectors as input. Logistic regression prediction can be done under condition  $0 \leq h\theta(x) \leq 1$ . Sigmoid function for logistic regression calculated by Eqns. (1) and (2).

$$h\theta(x) = g((\theta^T X)) \tag{1}$$

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{2}$$

### 4.2 Naïve Bayes (NB)

The Nave Bayes Classifier [8], which was invented by Thomas Bayes, is simple to construct and works more efficiently than other machine learning algorithms. It's a supervised classifier that determines if a piece of data is positive or negative. The paper [13] described the researchers' various obstacles and the necessity for greater research to overcome them. For machine learning and data mining, Nave Bayes is the most efficient and effective inductive learning algorithm. It is based on Baye's Theorem, with the assumption of predictor independence. Its competitive classification performance in real-world applications is surprisingly rare. To place it additional way, the Nave Bayes classifier accepts that the presence of preset qualities is disconnected to the presence of some other component. The Nave Bayes Model depends on the Bayes hypothesis and indicates the connection between the probabilities p of two occasions c and Z, which are addressed as  $P(c)$  and  $P(Z)$ , just as the contingent likelihood of occasion c molded by occasion Z as well as the other way around, which is addressed as  $P(c | Z)$  and  $P(Z | c)$ .

$$P(c|z) = \frac{P(c)P(z|c)}{P(z)} \tag{3}$$

A bunch of element esteems ( $t_1, t_2 \dots t_n$ ) is traditionally used to describe a model Z, where  $t_i$  is the worth of trait  $T_i$ .

Allow us to expect to be C, which is an arranging variable, and c, which is the worth of C.

### 4.3 Support Vector Machine (SVM)

Support Vector Machines (SVM) is a non-probabilistic paired straight classifier [12]. This calculation of SVM is embraced best for content classification when the component space is colossal. For a preparation set of points  $(xi, yi)$  where x is the component vector and y are the class, point is to track down the most extreme edge hyperplane that isolates the focuses with  $yi = 1$  and  $yi = -1$ . Support vector machine can be utilized to group the tweets as sure and negative from the information sources vectors and afterward compute the exactness of opinions with the assistance of term recurrence. Linear SVC () is utilized to foresee the precision of the characterization information dependent on the minimal likelihood of the input vectors. The boundary esteems set are punishment: default: 12, C: 1.0, *Class\_weight = none* , *Tol = 0.0001*, *Max\_iter = 1000*.

### 4.4 Multilayer Perceptron (MLP)

As per [14-18], the Multilayer Perceptron (MLP) is a solid and non-direct neural organization model that capacities as a widespread capacity approximate with something like one secret layer and various non-straight units that makes it effective to become familiar with any connection between input variable sets. The information stream in a Multilayer Perceptron (MLP) is unidirectional, like information going from the info layer to the yield layer. The info layer of a multi-facet perceptron (MLP) neural organization starts with every hub going about as an indicator variable. Neurons (input hubs) in the forward streaming and next layer are identified with one another (named as the secret layer). Essentially, covered up layer neurons are connected to other secret layer neurons, etc.; the yield layer design is displayed beneath.

- i) If the expectation is paired, the yield layer is comprised of only one neuron.
- ii) If the expectation is non-paired forecast, the yield layer is comprised of N neurons.

Since the neurons are organized thusly, data streams proficiently from the contribution to the yield layer. As demonstrated in the realistic beneath, a solitary layer perceptron has an information layer and a yield layer, however there is additionally a secret layer network in a similar calculation. The enactment is conveyed from the info layer to the yield layer in the main period of a multi-facet perceptron, which is known as the forward stage. Having the neurons designed in this configuration brings about an effective progression of data from contribution to yield layer. As displayed in the picture underneath there exist an information layer and a yield layer as in a solitary layer perceptron yet in equal there exist a secret layer network in a similar calculation. Multi-facet perceptron has two stages in the main stage the actuation is proliferated from the information layer to the yield layer, it is known as the forward stage. MLP doesn't begin with any presumptions, and it doesn't force any cutoff points on input information; it might likewise assess information in spite of the presence of clamor or bends in the information.

## 5. RESULTS

As per Figure 2, the accuracy for Logistic regression (LR)

is just about 86.23%, Support vector machine (SVM) is 85.69% and naive bayes (NB) is 83.54%.and Multilayer Perceptron (MLP) is 92.44%. Then the Multilayer Perceptron (MLP) as per Table 1, outperforms the LR, SVM and NB with an accuracy of 92.44%.

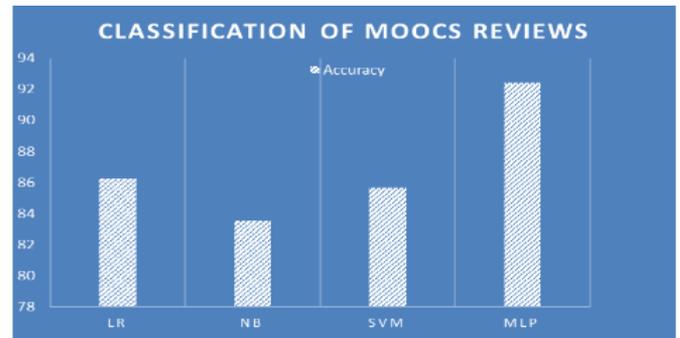


Figure 2. Accuracy analysis of different machine learning algorithms

Table 1. Multilayer Perceptron (MLP)

	Precision	Recall	f1-score	Support
0	0.68	0.81	0.74	687
1	0.97	0.94	0.96	4436
Accuracy			0.92	5123
Macro Avg	0.83	0.87	0.85	5123
Weighted Avg	0.93	0.92	0.93	5123

**Accuracy of MLP Model: 92.44583252000781%**

### 5.1 Confusion matrix for MLP model

A confusion matrix of N x N grid is used for assessing the exhibition of a grouping model as per Figure 3, where N is the quantity of target classes. The network contrasts the genuine objective qualities and those anticipated by the machine learning model. This gives us an inclusive perspective on how well our characterization model is performing and what sorts of mistakes it is making.

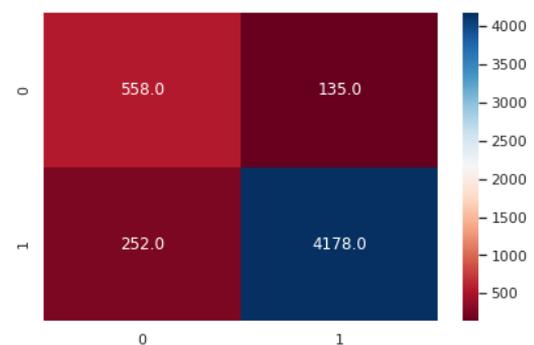
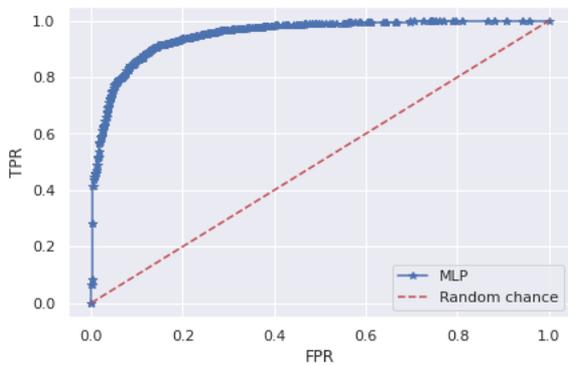


Figure 3. Confusion matrix for MLP model

### 5.2 Receiver operating characteristic

In Figure 4, the receiver operating characteristic (ROC) curve is a graphical representation of binary classifier system where threshold value varies. In ROC, representation on the pairs of true positive rates (y-axis) and false positive rates (x-axis) has been plotted by which result shows that there is lowering of the threshold down from 1, all the way to 0.



**Figure 4.** ROC Curve for MLP classifier

## 6. CONCLUSION AND FUTURE WORK

Sentiment Analysis is a technique for deciding a text's perspective. Individuals leave remarks via web-based media about occasions they've joined in, and they're interested if most of others had a good or negative involvement in a similar occasion. The reason for existing is to decide how exact sentences got from tweet message are as far as sentiments. The sentiment analysis of a tweet can assist with deciding if a tweet's opinion on explicit things, occasions, or different subjects is ideal or negative. The most troublesome test in feeling examination is deciding if an assessment word is ideal or negative. It is feasible to accomplish the best precision by changing the boundaries of the AI calculations. In this work, the presentation of different machine learning approaches includes SVM, Nave Bayes (NB), Logistic Regression (LR), and Multilayer Perceptron (MLP) is explained. The Multilayer Perceptron has accomplished an accuracy of roughly 92.44% when bigram model was utilized. MLP performs better when contrasted with other supervised machine learning algorithms for MOOC data set.

In Future, we can extend our work to evaluate the fluctuation in the concert of sentiment analysis algorithm when various other features are also considered. A further active learning techniques can be designed which can handle and manage expected error reduction, pool-based sampling, uncertainty sampling to detect sentiments on MOOC dataset and to enhance the self-confidence of user.

## REFERENCES

- [1] Araque, O., Zhu, G., Iglesias, C.A. (2019). A semantic similarity-based perspective of affect lexicons for sentiment analysis. *Knowledge-Based Systems*, 165: 346-359. <https://doi.org/10.1016/j.knsys.2018.12.005>
- [2] Kim, S., Eduard, H. (2006). Identifying and analyzing judgment opinions. *Proceedings of the Human Language Technology Conference of the NAACL*. <https://doi.org/10.3115/1220835.1220861>
- [3] Xu, G., Yu, Z., Yao, H., Li, F., Meng, Y., Wu, X. (2019). Chinese text sentiment analysis based on extended sentiment dictionary. *IEEE Access*, 7: 43749-43762. <https://doi.org/10.1109/ACCESS.2019.2907772>
- [4] Hung, C. (2017). Word of mouth quality classification based on contextual sentiment lexicons. *Inf. Process. Manage.*, 53(4): 751-763. <https://doi.org/10.1016/j.ipm.2017.02.007>
- [5] Bakshi, R.K., Kaur, N., Kaur, R., Kaur, G. (2016). Opinion mining and sentiment analysis. *3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 452-455.
- [6] Cai, Y., Yang, K., Huang, D., Zhou, Z., Lei, X., Xie, H.R., Wong, T.L. (2019). A hybrid model for opinion mining based on domain sentiment dictionary. *Int. J. Mach. Learn. & Cyber.*, 10: 2131-2142. <https://doi.org/10.1007/s13042-017-0757-6>
- [7] Wang, J., Yu, L., Lai, K.R., Zhang, X. (2020). Tree-structured regional CNN-LSTM model for dimensional sentiment analysis. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28: 581-591. <https://doi.org/10.1109/TASLP.2019.2959251>
- [8] Kong, L., Li, C.Y., Ge, J.D., Zhang, F.F., Feng, Y., Li, Z.J., Luo, B. (2018). Leveraging multiple features for document sentiment classification. *Inf. Sci.*, 518: 39-55. <https://doi.org/10.1016/j.ins.2020.01.012>
- [9] Valakunde, N.D., Patwardhan, M.S. (2018). Multi-aspect and multi-class based document sentiment analysis of educational data catering accreditation process. *Proc. Int. Conf. Cloud Ubiquitous Comput. Emerg. Technol.*, pp. 188-192. <https://doi.org/10.1109/CUBE.2013.42>
- [10] Chakraborty, K., Bhattacharyya, S., Bag, R. (2020). A survey of sentiment analysis from social media data. *IEEE Transactions on Computational Social Systems*, 7(2): 450-464. <https://doi.org/10.1109/TCSS.2019.2956957>
- [11] Schneider, S.L., Council, M.L. (2021). Distance learning in the era of COVID-19. *Archives of Dermatological Research*, 313(5): 389-390. <https://doi.org/10.1007/s00403-020-02088-9>
- [12] Jiang, M., Zhang, W., Zhang, M., Wu, J., Wen, T. (2019). An LSTM-CNN attention approach for aspect-level sentiment classification. *Journal of Computational Methods in Sciences and Engineering*, 19(4): 859-868. <https://doi.org/10.3233/JCM-190022>
- [13] Moreno-Marcos, P.M., Alario-Hoyos, C., Muñoz-Merino, P.J., Estévez-Ayres, I., Kloos, C.D. (2018). Sentiment analysis in MOOCs: A case study. *IEEE Global Engineering Education Conference (EDUCON)*, pp. 1489-1496. <https://doi.org/10.1109/EDUCON.2018.8363409>
- [14] Tang, D., Qin, B., Feng, X., Liu, T. (2015). Effective LSTMs for target-dependent sentiment classification. *arXiv preprint arXiv:1512.01100*. <https://aclanthology.org/C16-1311>.
- [15] Long, F., Zhou, K., Ou, W. (2019). Sentiment analysis of text based on bidirectional LSTM with multi-head attention. *IEEE Access*, 7: 141960-141969. <https://doi.org/10.1109/ACCESS.2019.2942614>
- [16] Le, Q., Mikolov, T. (2014). Distributed representations of sentences and documents. In *International Conference on Machine Learning*. PMLR, pp. 1188-1196.
- [17] Xu, Q., Zhu, L., Dai, T., Yan, C. (2020). Aspect-based sentiment classification with multi-attention network. *Neurocomputing*, 388: 135-143. <https://doi.org/10.1016/j.neucom.2020.01.024>
- [18] Devlin, J., Chang, M.W., Lee, K., Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. <https://arxiv.org/abs/1810.04805>.

## NOMENCLATURE

NB Naïve Bayes  
MLP Multilayer Perception

LR Logistic Regression  
MOOC Massive Open Online Courses  
SVM Support Vector Machine  
ROC Receiver Operating Characteristic