AN OVERVIEW OF MACHINE LEARNING TECHNOLOGIES AND THEIR USE IN E LEARNING

[1]Md Abdul Mubeen, [2]Mohammed Abdul Rawoof
[1]PG Scholar, M.Tech, Dept of CSE, Nawab Shah Alam Khan College of Engineering and Technology, Hyderabad
mohdabdulmubeen30@gmail.com
[2]Assistant Professor, Dept of CSE, Nawab Shah Alam Khan College of Engineering and Technology, Hyderabad
mohammedabdulrawoof3@gmail.com

ABSTRACT

Because of new technologies, internet, connected objects we produce a wonderful amount of data. Placing these data in context, organizing them to have the option to see, understand and reflect them is vital. Traditionally, human have analyzed data. Nonetheless, as the volume of data surpasses, human increasingly go to mechanized frameworks that can copy him. Those frameworks ready to gain from the two data and changes in data to take care of issues are called machine learning. Man-made consciousness significantly affects e-learning research and the machine learning based methods can be carried out to further develop Technology Enhanced Learning Environments (TELE). This paper is an outline of the new discoveries in this research field. From the get go, we acquaint the key concepts related with machine learning. Then, we present a few ongoing works utilizing machine learning in e-learning context.

1. INTRODUCTION

Nearly all that we do today leaves a computerized follow that portrays our exercises, determines our location, and gives numerous other data about what we say, what we purchase, and so forth. Because of the two data stockpiling limit and digitalization of society, a large portion of gadgets, machines and all that we use, produce data. We can, as model, separate data from pay stations, stopping, PDA, social networks, recordings, photographs, and so on. It is important to benefit and track down significance to this gathered data.

Analyzing data makes it conceivable to understand peculiarities, to demonstrate ways of behaving and to make expectations. Previously, humans dissect data, composed calculations and the machine applied them to take care of issues. Today, humans present data and permits the machine to advance all alone from these data without being expressly modified. We discuss the force of data. This is the guideline of machine learning.

In actuality, there is an attention to the wealth that data can hold and the significance of esteeming it. As a matter of fact, breakdown of complicated data through machine learning methods has arisen as a significant period in a few logical research spaces like medication, web-based business, industry, education, social networks, economics and finance, and so on.

Figure 1 shows machine learning relationships to a few different concepts of data science and artificial intelligence. As a matter of fact, data mining use insights to remove stowed away data (designs) from crude data. Be that as it may, mama chine learning as a subfield of software engineering and artificial intelligence, gains from examples to foresee. The Deep Learning is one of the primary technologies of machine learning and artificial intelligence. We can say that it is the new age of machine learning which portrayed by learning by layer and on each layer the mama chine needs to become familiar with somewhat more.

2. RELATED WORK

Kabyshev, M. V., & Kovalchuk, S. V. 2019. Healthcare systems should provide technology to store data in different forms: numbers, text, chart or images. These data are growing constantly. As a result, medical databases and environments are increasing year by year. Using information technologies and computer systems in medicine can help to get effectiveness in diagnostic decision making and better risk and knowledge management. The strategy to raise quality of healthcare and still be competitive is to build up strong information systems for these goals. The healthcare system should provide transferring medical parameters to an automated system aimed at monitoring in real time patient's health condition. Because of this requirement, all data is analyzed continuously. Automating health monitoring can help doctors with early diagnosis of complications and improve treatment tactics.

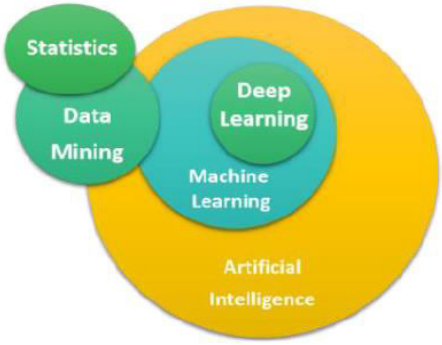

Fig. 1 Machine Learning relationships to other related fields.

The main target of this paper is the development of the client-server system for chronic disease patients, in particular for patients with diabetes mellitus. This system should make predictions of the dynamics of patient characteristics using predictive modeling methods and clinical prediction models, its classification with the automated interpretation of patient conditions with the experts' involvement. For this, it is necessary to integrate predictive models and expert knowledge. Also, the system takes into account the peculiarities of behavior in the framework of personalized and participatory medicine.

Zhu, G., Wu, Z., Wang, Y., Cao, S., & Cao, 2019. In this study, we consider the purchase prediction problem in the context of e-tourism, an emerging and prevailing application in e-commerce. Although a wide array of studies have been taken on purchase prediction, little analysis has been done on the purchasing behaviors towards tourism products. Also, the design of the corresponding purchase prediction model deserves researchers' full attention. We begin by introducing a real-life e-tourism dataset and constructing a suite of variables based on the detailed current and historical clickstream information. To validate the effectiveness of variables, we then perform a quantitative analysis to address quite a few interesting characteristics of purchase patterns. To predict whether or not a purchase is made for a current visiting session, we present a novel model called co-EM Logistic Regression (co-EM-LR) which combines the semi-supervised learning and the multi-view learning into its procedure. The co-EM-LR model has at least two outstanding merits: (1) it inherits the ease interpretation of the logistic modeling; and (2) it fully exploits both unlabeled data and the compatibility of multiple views to improve the prediction accuracy. Comprehensive experiments demonstrate the proposed co-EM-LR model yields significant prediction performance advantages over five competitive methods. Furthermore, two complementary views can mutually improve the performance with each other and finally offer fast convergence.

Hmedna, B., El Mezouary, A., & Baz, O. 2019. Different learners have quite varied learning styles, which have been influenced by their personalities, backgrounds knowledge, and skills. In this paper, we propose an automatic approach to extract learners' features from traces of learners when interacting with learning materials in MOOCs. The Felder–Silverman Learning Style Model (FSLSM) is used since it is one of the most commonly used models in technology-enhanced learning. To carry out this study we used traces of 5482 learners enrolled in the edX course "Statistical Learning (Stat, Winter 2015)" administered via Stanford's Logunita platform. By applying unsupervised clustering method learners are grouped according to their degree of preferences for active/reflective learning styles. The findings of this study reveal that the majority of learners prefer an active learning style.

Kumari, K. V., & Kavitha, C. R, 2019. Social Network (SN) is an online platform broadly used as communication tool by millions of users in order to build social relationships with others for knowledge point of view, career purposes and many more. Social Networks such as Twitter, Facebook, and LinkedIn have become the most leading tools on the web. Spam, floods the Internet with many copies of the same message and it can be manifest in numerous ways, it includes bulk messages, malicious links, fake friends, fraudulent reviews and personally identifiable information. The aim of this paper is to classify the tweets into spam and non-spam using Machine Learning and which will give the best results.

Portugal, I., Alencar, P., & Cowan, D, 2018. Recommender systems use algorithms to provide users with product or service recommendations. Recently, these systems have been using machine learning algorithms from the field of artificial intelligence. However, choosing a suitable machine learning algorithm for a recommender system is difficult because of the number of algorithms described in the literature. Researchers and practitioners developing recommender systems are left with little information about the current approaches in algorithm usage. Moreover, the development of recommender systems using machine learning algorithms often faces problems and raises questions that must be resolved. This paper presents a systematic review of the literature that analyzes the use of machine learning algorithms in recommender systems and identifies new research opportunities. The goals of this study are to (i) identify trends in the use or research of machine learning algorithms in recommender systems; (ii) identify open questions in the use or research of machine learning algorithms; and (iii) assist new researchers to position new research activity in this domain appropriately. The results of this study identify existing classes of recommender systems, characterize adopted machine learning approaches, discuss the use of big data technologies, identify types of machine learning algorithms and their application domains, and analyzes both main and alternative performance metrics.

3. MACHINE LEARNING

In machine learning, a computer learns from example data how to perform tasks. We know that if we give more experiences (E) with a defined task (T) to a machine, its performance (P) improves [12]. For example, let suppose that we want an email client to classify emails as spam or not. The experience E in this case should be a set of emails already classified as spam or not. The Task T perform is to classify automatically

new emails. The performance P that should increase is the accuracy rate of the classification made by the machine on a set of new emails.

A. Machine learning process

The conventional machine learning process comprises of seven steps as described next [13]. The initial step is to collect data. It is a vital task since it will decide how great prescient model can be. In any case, data we assembled are, in many times, unstructured, contain a great deal of clamor or need to take different structures to be valuable for our machine learning. Along these lines, data should be cleaned and pre-handled.

After that we can start constructing our machine learning model. For this, we start by the component designing in which we pick the most significant highlights from data, then, at that point, we attempt to choose the best machine learning calculation for the front and center concern. It is basic in come by the most ideal outcomes.

The following task is preparing. In this step we utilize a piece of our data to gradually further develop machine learning ability to foresee. When preparing is finished, the time has come to test the model and see how it could perform against the other piece of data inconspicuous. The performance assessment is estimated by different boundaries like accuracy, accuracy and review. At times, it is feasible to return and further develop preparing then test once more. The last step is the outcome given by the machine learning. It tends to be an expectation or derivation.

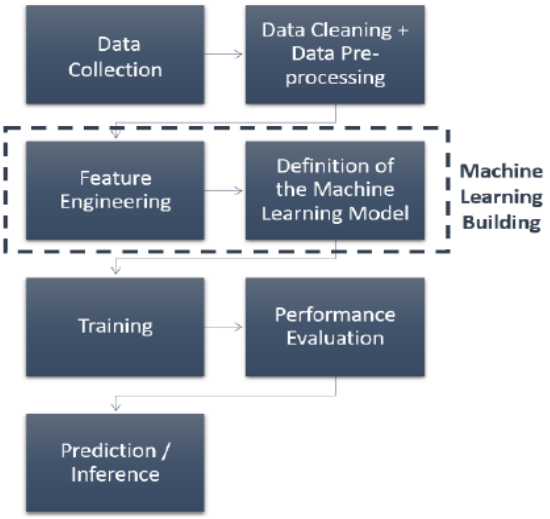

Fig. 2 Components of a Generic machine learning model

B. Machine learning paradigms

Machine learning can be ordered in view of the methodology utilized for the learning system. Four primary classes were recognized: supervised, unsupervised, semi-supervised, and reinforcement learning [12].

In supervised learning, we have a bunch of preparing data or named data in which we know the construction and the result. We take this data and train a machine learning model, so it can figure out patterns in the data. When the model has been prepared, we can utilize it to foresee consequences of data in which results are obscure [14].

Alternately, unsupervised learning methods gain structure from the actual data without the requirement for earlier marking [15]. That is mean we can apply unsupervised machine learning to track down patterns that exist inside marked data.

Be that as it may, full name information isn't accessible consistently. Semi-supervised learning gives a strong structure to utilizing unlabeled data when names are restricted or costly to get [16].

The last machine learning approach, serves when we understand what we are searching for yet we don't have any idea how to get it. The standard is to test a few arrangements and afterward we see which ones make it conceivable to get the ideal outcome. Reinforcement learning issue can be formalized as a specialist that needs to go with choices in an environment. The specialist learns an acceptable conduct. This implies that it alters or obtains new ways of behaving and abilities gradually. In this manner, the reinforcement specialist doesn't need total information or control of the environment, however it just should have the option to collaborate with the environment and collect information [17].

4. PROPOSED MODULES & ALGORITHM

1. **Collect Data**
2. **Data Preprocessing**
3. **Training**
4. **Machine Learning**

1. COLLECT DATA

The first step is to collect data. It is a very important task because it will determine how good predictive model can be. But, data we gathered are, in most times, unstructured, contain a lot of noise or have to take other forms to be useful for our ma-chine learning. So, data need to be cleaned and pre-processed.

2. DATA PREPROCESSING

For achieving better results from the applied model in Machine Learning projects the format of the data has to be in a proper manner. Some specified Machine Learning model needs information in a specified format, for example, Random Forest algorithm does not support null values, therefore to

execute random forest algorithm null values have to be managed from the original raw data set.

Another aspect is that data set should be formatted in such a way that more than one Machine Learning and Deep Learning algorithms are executed in one data set, and best out of them is chosen.

After that we can begin building our machine learning model. For this, we start by the feature engineering in which we choose the most relevant features from data, then we try to select the best machine learning algorithm for the problem at hand. It is imperative in getting the best possible results.

3. TRAINING

In this step we use a part of our data to incrementally improve machine learning ability to predict. Once training is complete, it is time to test the model and observe how it might perform against the other part of data unseen. The performance evaluation is measured by various parameters like accuracy, precision and recall. Sometimes, it is possible to go back and improve training then test again. The last step is the result given by the machine learning. It can be a prediction or inference.

4. MACHINE LEARNING

The last machine learning approach, serves when we know what we are looking for but we do not know how to get it. The principle is to test several solutions and then we see which ones make it possible to obtain the desired result. Reinforcement learning problem can be formalized as an agent that has to make decisions in an environment. The agent learns a good behavior. This means that it modifies or acquires new behaviors and skills incrementally. Thus, the reinforcement agent does not require complete knowledge or control of the environment, but it only needs to be able to interact with the environment and collect information.

5. MACHINE LEARNING E-LEANRING APPLICATIONS

These days, everybody needs to learn and foster his insight in many fields, students, bosses, and so forth. With the spread of long-lasting learning, school systems are seriously facing the modernization and e-learning is be-coming increasingly well known. This prompts a massive development in the quantity of Technology Enhanced Learning Environments (TELE) offering open or private online courses and other various kinds of services. Examining the huge measure of data created by TELE through machine learning methods has arisen. It is valuable to concentrate on the most proficient method to take advantage of this powerful, new technology to upgrade e-learning.

A. Sentiment Analysis

As of late, Massive Open Online Course (MOOC) achievement is considered as the degree of student fulfillment with the course [18]. Opinion investigation can be utilized to distinguish complex feelings [19] holding back nothing student fulfillment. In [19], researchers need to stop mine toss discussion messages in MOOC the extremity of learners' opinions, positive feelings and pessimistic feelings. They look at five supervised machine learning algorithms which have been utilized all the more often in commitments connected with prediction in MOOCs: Logistic Regression, Support Vector Machine, Decision Tree, Random Forest and Naïve Bayes. Results show that the most solid procedure was Random Forest.

Understanding the job of feelings in MOOC students' learning experiences is vital. In one hand, as per [20] the control of accomplishment feelings might further develop learning commitment. In view of SVM, [20] fabricate a supervised machine learning model to sort accomplishment feelings automatically. SVM was taken on as it gives preferable performance results over Naïve Bayes, Logistic Regression and Decision Tree. Then again, [21] track the profound propensities of learners to examine the acknowledgment of the courses utilizing huge data from schoolwork culmination, remarks and gatherings. In light of semantic examination and machine learning, [21] research the relationship be-tween close to home propensities and learning impacts.

B. Student Behaviour Prediction

A fascinating writing audit [22] has resolved the topic of machine learning use in foreseeing student conduct. Tow research objectives were recognized: student classification and dropout prediction.

- **Student Classification:**

Surely, characters, foundations information, abilities and inclinations have a pivotal impact in the learning system. Recommender frameworks give generally reasonable substance to every student. Profiling and classifying learners is an early stage task not exclusively to customize learning yet in addition to distinguish deserting factors and numerous different purposes. We sum up in table 1, a few late works zeroing in on the student classification utilizing machine learning.

- **Dropout Prediction:**

Different machine learning procedures have been applied to investigate intuitive conduct follows left across TELE. As per [27] who centers around learners' clickstream data, Logistic regression (LR) has been the most often utilized procedure to anticipate student dropout in MOOC environment accomplishing 89% as accuracy. SVM and Decision Tree possess the subsequent position, nonetheless, Natural Language Processing Technique come in the third spot.

C. Self-Regulated Learning

With the little outside educator's checking in greater part of TELE, learners are expected to settle on choices connected with their own exercises [28]. All things considered, people areas of strength for with regulated learning (SRL) abilities, portrayed by the ability to design, oversee and control their learning cycle, can learn quicker and better than those with more fragile SRL abilities [29]. As it is one of e-learning stages supporting SRL techniques [30], MOOC points learners to self-assess the quality or the advancement of their work, to lay out objectives and plan and give them the likelihood to rehash notes, logs, tests, or learning materials to get ready for testing, and so on. In spite of that multitude of elements, it stays significant for some re-searchers to upgrade student SRL in light of machine learning approach.

In view of learners' log follows and reactions to an overview, [31] add to upgrade comprehension of how students learn, and how guidance ought to be intended to help SRL in a nonconcurrent online course at a ladies' college in South Kore. In this review, researchers continue to the disclosure of student profiles and the assessment of student SRL process extra time. From the start, they proposed three key SRL ascribes;

time interest in satisfied learning, concentrate on routineness and help-chasing that apply to nonconcurrent online courses to act as the reason for the examination of SRL, and directed the determination of log factors. Second, they distinguished student subpopulations utilizing K-medoids clustering algorithm by outline strategy. Subsequent to finding existing groups and their learning patterns, utilize irregular backwoods classification as a choice tree-based machine learning algorithm to foresee bunch enroll.

5. RESULT

```
Feature column(s):-
['school', 'sex', 'age', 'address', 'famsize', 'Pstatus', 'Medu', 'Fedu', 'Mjob', 'Fjob', 'reason', 'guardian', 'traveltime',
'studytime', 'failures', 'schoolsup', 'famsup', 'paid', 'activities', 'nursery', 'higher', 'internet', 'romantic', 'famrel', 'f
reetime', 'goout', 'Dalc', 'Walc', 'health', 'absences']
Target column: passed

Feature values:-
  school sex  age address famsize Pstatus  Medu  Fedu     Mjob      Fjob  ... \
0     GP   F   18       U     GT3       A     4     4  at_home   teacher  ...
1     GP   F   17       U     GT3       T     1     1  at_home     other  ...
2     GP   F   15       U     LE3       T     1     1  at_home     other  ...
3     GP   F   15       U     GT3       T     4     2   health  services  ...
4     GP   F   16       U     GT3       T     3     3    other     other  ...

  higher internet  romantic  famrel  freetime goout Dalc Walc health absences
0    yes       no        no       4         3     4    1    1      3        6
1    yes      yes        no       5         3     3    1    1      3        4
2    yes      yes        no       4         3     2    2    3      3       10
3    yes      yes       yes       3         2     2    1    1      5        2
4    yes       no        no       4         3     2    1    2      5        4

[5 rows x 30 columns]
```

Fig 3. Data Set

```
-----------------------------------------
Training set size: 100
Training DecisionTreeClassifier...
Done!
Training time (secs): 0.016163825988770
Predicting labels using DecisionTreeClassifier...
Done!
Prediction time (secs): 0.007004976272583
F1 score for training set: 1.0
Predicting labels using DecisionTreeClassifier...
Done!
Prediction time (secs): 0.006105661392212
F1 score for test set: 0.6725663716814159
-----------------------------------------
Training set size: 200
Training DecisionTreeClassifier...
Done!
Training time (secs): 0.007999181747437
Predicting labels using DecisionTreeClassifier...
Done!
Prediction time (secs): 0.008004903793335
F1 score for training set: 1.0
Predicting labels using DecisionTreeClassifier...
Done!
Prediction time (secs): 0.008002519607544
F1 score for test set: 0.703125
-----------------------------------------
```

Fig 4. Training Data1

```
-----------------------------------------
Training set size: 100
Training KNeighborsClassifier...
Done!
Training time (secs): 0.010809421539307
Predicting labels using KNeighborsClassifier...
Done!
Prediction time (secs): 0.154745101928711
F1 score for training set: 0.7832167832167832
Predicting labels using KNeighborsClassifier...
Done!
Prediction time (secs): 0.026201963424683
F1 score for test set: 0.712121212121212
-----------------------------------------
Training set size: 200
Training KNeighborsClassifier...
Done!
Training time (secs): 0.008002281188965
Predicting labels using KNeighborsClassifier...
Done!
Prediction time (secs): 0.024003267288208
F1 score for training set: 0.8611111111111112
Predicting labels using KNeighborsClassifier...
Done!
Prediction time (secs): 0.008006334304810
F1 score for test set: 0.7067669172932329
-----------------------------------------
```

Fig 5. Training Data2

```
------------------------------------------
Training set size: 100
Training SVC...
Done!
Training time (secs): 0.027737379074097
Predicting labels using SVC...
Done!
Prediction time (secs): 0.009079456329346
F1 score for training set: 0.7901234567901235
Predicting labels using SVC...
Done!
Prediction time (secs): 0.007999897003174
F1 score for test set: 0.7581699346405228
------------------------------------------
Training set size: 200
Training SVC...
Done!
Training time (secs): 0.008009672164917
Predicting labels using SVC...
Done!
Prediction time (secs): 0.015995502471924
F1 score for training set: 0.7987987987987988
Predicting labels using SVC...
Done!
Prediction time (secs): 0.007996797561646
F1 score for test set: 0.7662337662337663
------------------------------------------
```

Fig 6. Training Data3

```
Performance for KNN:
     labels  Train_times  Predtime_train  Predtime_test  F1_train  F1_test
0  Size_100         0.01            0.00           0.00      0.80     0.71
1  Size_200         0.00            0.02           0.01      0.86     0.71
2  Size_300         0.01            0.03           0.01      0.87     0.75
Performance for Decision Tree:
     labels  Train_times  Predtime_train  Predtime_test  F1_train  F1_test
0  Size_100         0.00            0.01              0         1     0.69
1  Size_200         0.01            0.00              0         1     0.74
2  Size_300         0.01            0.00              0         1     0.70
Performance for SVM:
     labels  Train_times  Predtime_train  Predtime_test  F1_train  F1_test
0  Size_100         0.01            0.00           0.00     0.859    0.784
1  Size_200         0.01            0.01           0.01     0.869    0.775
2  Size_300         0.02            0.02           0.01     0.869    0.758
```

Fig 7. Performance for Algorithms

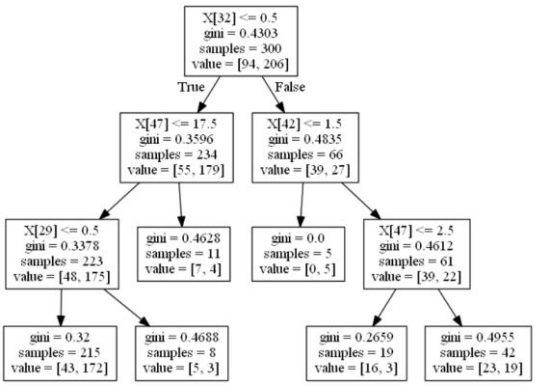

Fig 8. Output Page

6. CONCLUSION AND FUTURE ENHANCEMENT

E-learning researchers have burnt through considerable energy on breaking down learners' data through machine learning methods to improve learning experience. This is by all accounts wise since the student is considered as the principal component in the e-learning circle. In any case, no examination work, the best of our insight does to involve learning data to gauge content quality to further develop it.

In this manner, in our future work will zero in on e-learning con-tent assessment by utilizing machine learning. The primary goal is to assist with flowing designers in the instructive reengineering process in light of machine learning finding and in view of many factors, especially past learners' connections.

REFERENCES

[1] Rakhmetulayeva, S. B., Duisebekova, K. S., Mamyrbekov, A. M., Kozhamzharova, D. K., Astaubayeva, G. N., & Stamkulova, K. (2018). Application of classification algorithm based on SVM for determining the effectiveness of treatment of tuberculosis. Procedia computer science, 130, 231-238.

[2] Kabyshev, M. V., & Kovalchuk, S. V. (2019). Development of personalized mobile assistant for chronic disease patients: diabetes mellitus case study. Procedia Computer Science, 156, 123-133.

[3] Zhu, G., Wu, Z., Wang, Y., Cao, S., & Cao, J. (2019). Online Purchase Decisions for Tourism E-commerce. Electronic Commerce Research and Applications, 100887.

[4] Brik, B., Bettayeb, B., Sahnoun, M. H., & Duval, F. (2019). Towards Predicting System Disruption in Industry 4.0: Machine Learning-Based Approach. Procedia Computer Science, 151, 667-674.

[5] Han, Y., Zeng, Q., Geng, Z., & Zhu, Q. (2018). Energy management and optimization modeling based on a novel fuzzy extreme learning machine: Case study of complex petrochemical industries. Energy conversion and management, 165, 163-171.

[6] Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2019). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. Computers & Education, 103724.

[7] Hmedna, B., El Mezouary, A., & Baz, O. (2019). How Does Learners' Prefer to Process Information in MOOCs? A Data-driven Study. Procedia computer science, 148, 371-379.

[8] Birjali, M., Beni-Hssane, A., & Erritali, M. (2017). Machine learning and semantic sentiment analysis-based algorithms for suicide sentiment prediction in social networks. Procedia Computer Science, 113, 65-72.

[9] Kumari, K. V., & Kavitha, C. R. (2019). Spam Detection Using Machine Learning in R. In International Conference on Computer Networks and Communication Technologies (pp. 55-64). Springer, Singapore.

[10] Ghoddusi, H., Creamer, G. G., & Rafizadeh, N. (2019). Machine learning in energy economics and finance: A review. Energy Economics, 81, 709-727.

[11] Liu, J., Kong, X., Zhou, X., Wang, L., Zhang, D., Lee, I., ... & Xia, F. (2019). Data Mining and Information Retrieval in the 21st century: A

bibliographic review. Computer Science Review, 34, 100193.
[12] Portugal, I., Alencar, P., & Cowan, D. (2018). The use of machine learning algorithms in recommender systems: A systematic review. Expert Systems with Applications, 97, 205-227.
[13] Alzubi, J., Nayyar, A., & Kumar, A. (2018, November). Machine learning from theory to algorithms: an overview. In Journal of Physics: Conference Series (Vol. 1142, No. 1, p. 012012). IOP Publishing.
[14] Schrider, D. R., & Kern, A. D. (2018). Supervised machine learning for population genetics: a new paradigm. Trends in Genetics, 34(4),301-312.
[15] Rodriguez-Nieva, J. F., & Scheurer, M. S. (2019). Identifying topological order through unsupervised machine learning. Nature Physics, 1.
[16] Oliver, A., Odena, A., Raffel, C. A., Cubuk, E. D., & Goodfellow, I. (2018). Realistic evaluation of deep semi-supervised learning algorithms. In Advances in Neural Information Processing Systems (pp. 3235-3246).
[17] François-Lavet, V., Henderson, P., Islam, R., Bellemare, M. G., & Pineau, J. (2018). An introduction to deep reinforcement learning. Foundations and Trends® in Machine Learning, 11(3-4), 219-354.
[18] Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2019). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. Computers & Education, 103724.
[19] Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., Estévez-Ayres, I., & Kloos, C. D. (2018, April). Sentiment Analysis in MOOCs: A case study. In 2018 IEEE Global Engineering Education Conference (EDUCON) (pp. 1489-1496). IEEE.
[20] Xing, W., Tang, H., & Pei, B. (2019). Beyond positive and negative emotions: Looking into the role of achievement emotions in discussion forums of MOOCs. The Internet and Higher Education, 100690.
[21] Wang, L., Hu, G., & Zhou, T. (2018). Semantic analysis of learners' emotional tendencies on online MOOC education. Sustainability, 10(6), 1921.
[22] de Souza, V. F., & Perry, G. (2019). Identifying student behavior in MOOCs using Machine Learning. International Journal of Innovation Education and Research, 7(3), 30-39.
[23] Hamid, S. S. A., Admodisastro, N., Manshor, N., Kamaruddin, A., & Ghani, A. A. A. (2018, February). Dyslexia adaptive learning model: student engagement prediction using machine learning approach.