Machine Learning and Deep Learning applications in E-learning Systems : A Literature Survey using Topic Modeling Approach

1st Chakir FRI
Department of mathematics and computer Science, Moulay Ismail University Of Meknes, Faculty of Sciences
Meknes, Morocco
chakir.fri@gmail.com

2th Rachid ELOUAHBI
Department of mathematics and computer Science, Moulay Ismail University Of Meknes, Faculty of Sciences
Meknes, Morocco
elouahbi@yahoo.fr

2020 6th IEEE Congress on Information Science and Technology (CiSt) | 978-1-7281-6646-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/CIST49399.2021.9357253

***Abstract*— E-learning has been one of the major trends in education and its becoming an attracting topic in the field of artificial intelligence and its subfields like machine learning and deep learning, that are considered the most promising technologies in our era where its application score is almost unlimited. Many researchers are showing interest in the topic with significant research results. The aim of this paper is to extract the applications of machine learning and deep learning in E-learning systems. In this work we collected research papers from five research databases: Springer Link, Science Direct, Scopus, IEEE Digital Library, and Web of Science for a topic modeling application using a machine learning technique known as Latent Dirichlet Allocation (LDA).**

Index Terms*— *E-Learning, Machine Learning, Deep Learning, Topic Modeling, Latent Dirichlet Allocation

I. Introduction

In the recent years, E-Learning Education has presented a significant growth and its attracting an enormous number of participants, providing a great support to learners by taking advantage of the internet and assisting them in acquiring knowledge in a flexible manner. Todays research aims to improve the E-learning experience by making it tailored to learners needs, as E-learning does not commonly consider the diversity of learners in case of their characteristics and abilities which makes adapting courses for them a difficult task, but with the inflation of artificial intelligence and its subfields like machine learning and Deep learning, and with the increasing amount of data becoming available (log files, clickstreams, etc.), many researches are being conducted in this topic showing promising results.

This paper represents a literature survey of machine learning and deep learning applications in E-learning systems. First, we describe some core concepts like E-learning, artificial intelligence, Machine learning and deep learning. Then we give a detailed explanation of our survey methodology followed by results, discussion, and finally a conclusion.

II. E-Learning Overview

E-learning is an approach to learning and teaching, including the application of all or part of the educational models based on the employment of electronic media and devices as tools for improving access to Learning, communication and interaction.[1] There are many ways to categorize E-learning based on some characteristics, one of the important characteristics is "time", that describes the timing when students can access the content, which divides E-learning into two types: [2]

- **Synchronous** :where the studying and teaching is offered in the same time, it requires simultaneous interactions, this style mimics a real-life class that is being held on virtual platform.
- **Asynchronous** : content is available to be accessed at any time by students, interactions can take place at different times.

E-learning can also be divided into five essential components, which are: audience, course structure, page design, content engagement and the usability. [3]

The majority of today's E-learning courses are based on learning management systems (LMS) which are used for administration, tracking, reporting and delivery of courses and training programs. There also the learning content management system (LCMS) which is more focused on the development of educational content and its management, in some advanced cases, this software (LCMS) can automatically build learning contents based on students profiles and characteristics (learning style, knowledge level ,etc.) via an adaptive system approach.[4]

E-learning can be combined with the traditional learning methods that require the presence of the teacher in a classroom with the students. That combination gave is the Blended Learning approach which provides flexibility in terms of time and space to students. One of the famous forms of blended learning is the flipped classroom, its an approach that provides increase interactions time between teacher and students, in a condition, that students take their own learning responsibility, teachers are more like guides and facilitators, motivating student to lead their way, each student taking individual education, classroom time is allocated for discussing difficulties and solving application problems.[5]

978-1-7281-6646-9/20/$31.00 ©2020 IEEE

In advanced researches, there is an interest in analyzing the adaptive e-learning approach, which refers to the functionalities of the system to provide learning content and pedagogic activities and methods for every student based on individual characteristics (knowledge, experience, etc.) that are structured in learners model [6].

One of the newest forms of e-learning that widely appears in the last few years is the massive open online courses (MOOC), it differs from the regular web-based online courses by the following constitutive characteristics: Massive (Address a large number of participants),Open (Open accessibility, no conditions for participation), Online (courses are exclusively conducted via the internet, Digitization), Course (the learning content is structured according to a didactical concept) [7].

III. Machine Learning And Deep Learning brief survey

A. Artificial intelligence

Artificial intelligence or AI (Pronounced AYE-EYE) is a term used to describe computational technologies that allow machines (computers) to act and take decisions imitating human behavior and intelligence[8]. It is a simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and the rules for using it), reasoning (using the rules to reach approximate or definite conclusions), and self-correction[9].Particular applications of Al include expert systems, speech recognition,natural language processing (NLP). etc.

B. Machine learning

Machine learning as defined by Arthur Samuel is a ”field of study that gives computers the ability to learn without being explicitly programmed.” [10]
Machine learning is a subfield of Artificial Intelligence, in which statistical methods and computational algorithms are being used to teach machines through examples and by experimentation with data on how to perform specific tasks [11]. it’s used today in many areas such as, face recognition, risk prediction, Etc.
Machine Learning algorithms are divided into 3 main categories: supervised learning, unsupervised learning and reinforcement learning.

Supervised learning: The supervised learning is the task of learning from training data-sets based on prediction; supervised learning algorithms need external assistance. The input dataset is divided into train and test dataset. The train dataset has output variable which needs to be predicted or classified. All algorithms learn some kind of pattern from training set dataset and apply them to the test dataset for prediction or classification. One of the most famous supervised learning algorithms are : Decision Tree, Nave Bayes, Support Vector Machine [12].

Unsupervised learning: Contrary to supervised learning, unsupervised learning is a machine learning in which a function is inferred based on unlabeled training data, the training data contains of only inputs and unknown outputs. Therefore, unsupervised learning algorithms goal is finding relations and patterns with the training data to make sense of it. Unsupervised learning algorithms can be divided into 3 main categories: classification/clustering, dimensionality reduction, and anomaly detection. Several algorithms within each category, some of the most popular ones are, K-Means, principal component analysis (PCA). [13]

Reinforcement Learning: Reinforcement learning is a type of learning which makes decisions based on which actions to take such that the outcome is more positive. A reinforcement learning algorithm, or the learner, has no knowledge which actions to take, and which actions gives the most reward until its been given a challenging situation. The actions taken by the learner may affect situations and their actions in the future. Reinforcement learning depends on two criteria: trial and error search and delayed outcome. This type of machine learning is widely used in game theory. [14]

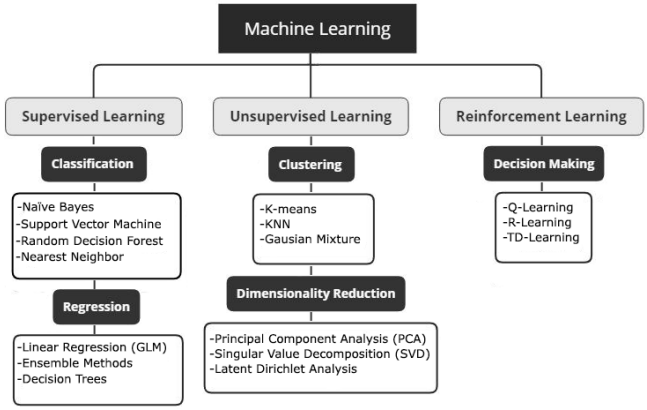

Fig. 1: Machine Learning Types

C. Deep learning

Deep Learning is a subfield of Machine Learning that allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. In order to process complex forms of data (text, sound, video) and solve complex problems such as medical diagnosis, facial recognition, self-driving cars, fraud analysis [15]. There are two types of deep learning, the convolutional neural network (CNN) and the recurrent neural network (RNN):

- **Convolutional neural network (CNN)**: Is a class of deep neural networks, commonly used in analyzing visual imagery and computer vision applications such as face recognition [16], automatic video analysis and classification [17], text classification and natural language processing [18]
- **Recurrent neural network (RNN)** : another type of neural network used in the text/language field when we are using sequential data such as speech recognition [19] and machine translation [20]

IV. SURVEY METHODOLOGY

The main purpose of our paper is to present an exploratory survey of the literature by exploring the most prominent applications of machine learning and deep learning algorithms in E-learning systems. To achieve that, we present a word frequency of preprocessed "keywords", followed by a distribution of research papers by years and journals, and finally we apply a topic modeling approach using the Latent Dirichlet Allocation (LDA) in order to extract topics from our collected papers

A. Papers Extraction

We have extracted a collection of research papers on machine learning and deep learning applications in E-Learning Systems from five research databases. We used the keywords "Machine Learning", "Deep Learning" and "E-learning" combined with the paper type filter option **Journal** papers in order to refine the search results. As described in Table 1,initially 133 papers were retrieved from the five databases, however we excluded 38 papers from the analysis due to redundancy and the out of subject cases. The final corpus consisted of 95 research papers dispersed as described in the same table above.

TABLE I: Total number of Research Papers Extracted from each Database

Research Database	Initial Extraction	Final Selection
Scopus	54	41
Science Direct	24	15
Springer Link	16	13
Web of Science	22	15
IEEE Xplore	17	11
Total	133	95

Every extracted paper is identified by a Metadata containing important informations such as the DOI, title, abstract, keywords, authors, journal, database,etc. in our survey we focused on titles for papers identification , abstracts for the topic modeling analysis, keywords for the word frequency analysis, journals, year of publication and database for paper distribution analysis.

B. Preprocessing

The purpose of preprocessing is to simplify data, eliminating as much as possible language dependent factors. In our work, there are three steps for preprocessing data:

- **Tokenization**: a document is treated as a string, removing all the punctuations, transforming text to lowercase, and then divided into a list of tokens.
- **Removing stop words**: Stop words refer to those words that appear too often in the language but carries no importance in determining the topic of the document such as 'the', 'of', 'about'. etc.
- **the Bag of words model**: a text or document is represented as the bag of its words, disregarding grammar and word order. The purpose of this step is to prepare the data in such a way that it can be processed by the LDA model. First of all, we formed a vocabulary for each word, a kind of their representation. Each word is represented by its number. Each document (abstracts in our case) in the corpus is represented in the form of a list of tuples where the first element is the id (number) of the word and second element the number of times that word occurs in the particular document.

C. Topic Modeling

Topic modeling is a popular method based on clustering approach that is used in finding hidden topical patterns of words in a large collection of documents [21]. There are several approaches for doing topic modeling. In our case, we adopted the Probabilistic approach, where a document is seen as a mixture of topics, which are modeled as distributions over a vocabulary, as every word and document gets probability scores for each topic [22]. In our work we applied the probabilistic Latent Dirichlet Allocation (LDA) to our corpus of preprocessed abstracts.

V. RESULTS AND DISCUSSION

A. Frequency of Keywords

The purpose of frequency analysis is the extract the most frequent keywords in our collection of papers (95) in order to highlight fields and trends involving the use of machine learning and deep learning in E-learning context. As we can notice in figure 2, the most frequent terms are machine learning, neural networks, E-learning, Data mining, Prediction, classification, etc. Frequency analysis is giving us a preliminary idea about the collected papers.

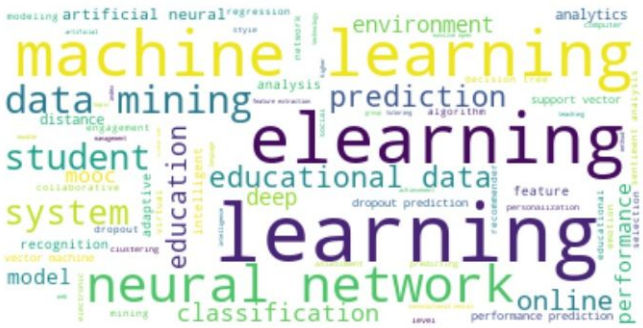

Fig. 2: Wordcloud of Keywords

B. Papers distribution

We notice a significant growth in the number of papers over the years, reaching the peak in 2019 with 23 papers (12 papers from Scopus) and 2020 with 24 papers (10 papers from Scopus). This tendency shows the growing interest in artificial intelligence, machine learning and deep learning that conquered all other fields including E-learning.

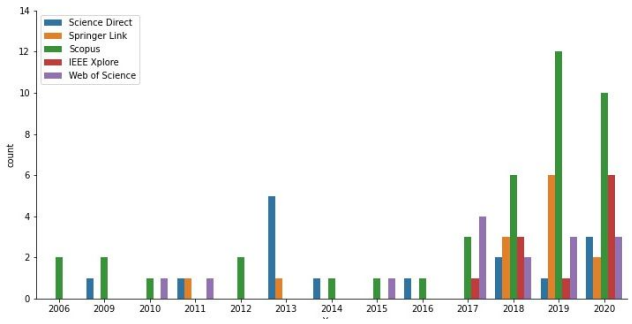

Fig. 3: Published research papers by years

The distribution of research papers across journals (Table 2) shows that most of the extracted journals are specialized in two fields, Education and computer science, as Computers & Education represent the top publishing journal with 6 published papers, in second position we find IEEE Transactions on Learning Technologies with 4 papers.

TABLE II: Number of Research Papers by Journals

Journal	Published Paper
Computers & Education	6
IEEE Access	6
Expert Systems with Applications	4
IEEE Transactions on Learning Technologies	4
International Journal of Advanced Computer Science and Applications	4
Knowledge-Based Systems	3
Computers in Human Behavior	3
International Journal of Emerging Technologies in Learning (iJET)	3
Education and Information Technologies	3
International Journal of Learning Technology	2
Interactive Learning Environments	2
...	2
Journal of Ambient Intelligence and Humanized Computing	1
Journal of Computer Science	1
...	1
Total	**95**

C. Latent Dirichlet Allocation

The used method in our topic modeling process is Latent Dirichlet Allocation, as the number of topics is the main parameter in any LDA model, and can be chosen manually or estimated on the basis of a parameter selection approach [22]. We trained our LDA model by passing three important arguments, Corpus, Number of topics and dictionary. We used the resulted model to evaluate the different LDA topic models that we created with different number of topics (from 2 to 20) using the GENSIM library that provide methods to measure coherence C_V, U_MASS and Perplexity scores for each of these models. The Figures below illustrates the coherence scores for each LDA model so as to select the best appropriate K value for number of topics.

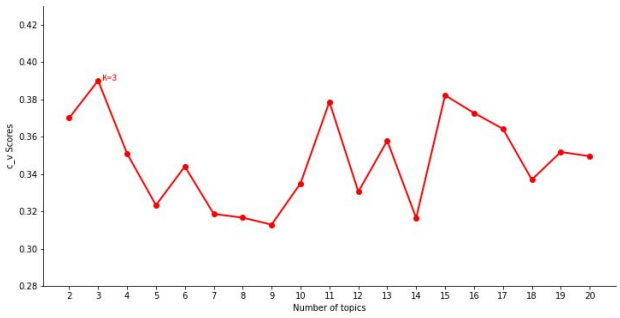

Fig. 4: c_v Scores for k $\epsilon[2-20]$

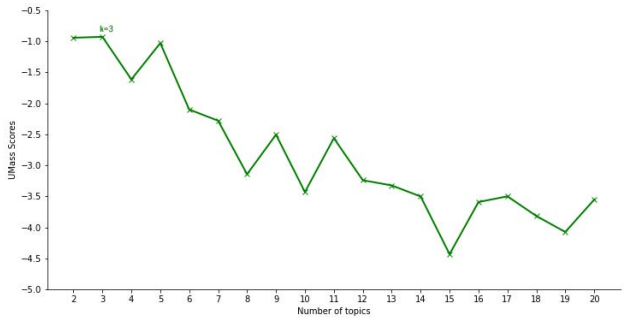

Fig. 5: U_Mass Scores for k $\epsilon[2-20]$

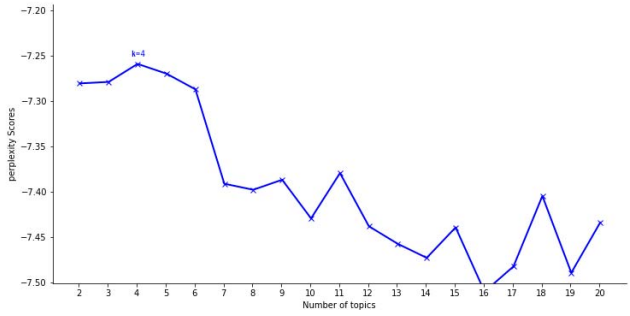

Fig. 6: perplexity Scores for k $\epsilon[2-20]$

Coherence estimation showed that the most coherent LDA model is the one with three topics, as resulted in the C_V and U_Mass scores with k=3, as for the Perplexity, the most coherent LDA model is the four topic model (k=4). In order to confirm the final number of topics we used the interactive topic visualization provided by the PyLDAvis library. The PyLDAvis library shows topics as circles, whereas the surface represent the percentage of words (tokens) under each topic compared to the corpus, and the distance between circles reflect the

coherence of the LDA model (the more distant the better the model).

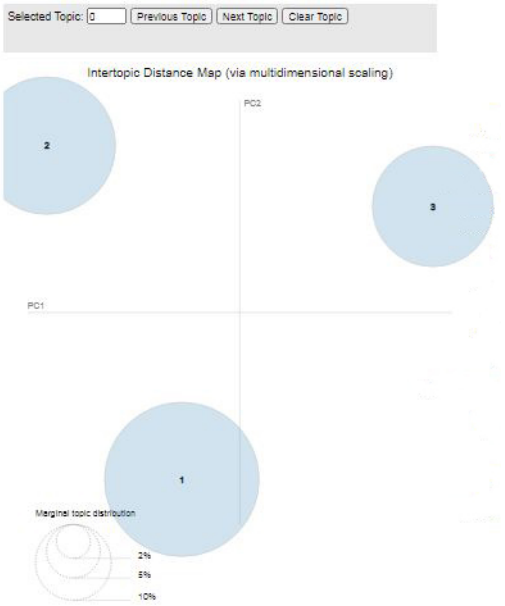

Fig. 7: PyLDAvis visulisation of the three topics LDA model

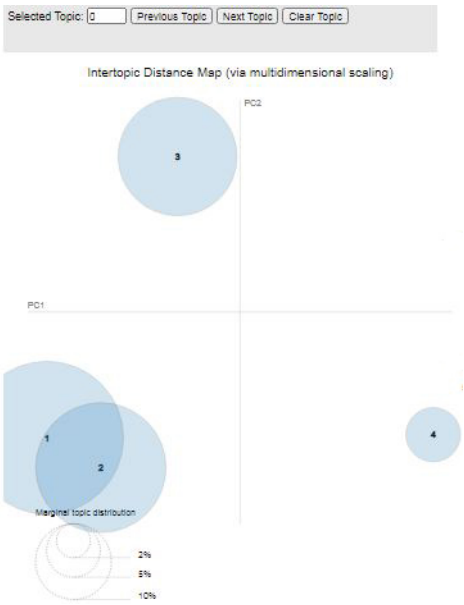

Fig. 8: PyLDAvis visulisation of the four topics LDA model

PyLDAvis visualization shows that the most coherent LDA model is the three topics one (figure 8), as for the four topic model (figure 9), we notice an intersection between topic 1 and 2 which shows less efficiency for this model compared to the three topic one. The PyLDAvis visualization show us the distribution of tokens for each topic which is clearly more balanced in the three topics model:

- Topic 1 : **41%** of tokens.
- Topic 2 : **33%** of tokens.
- Topic 3 : **25%** of tokens.

We used the three topic LDA model in order to highlight the most ten frequent words (weight) under each topic:

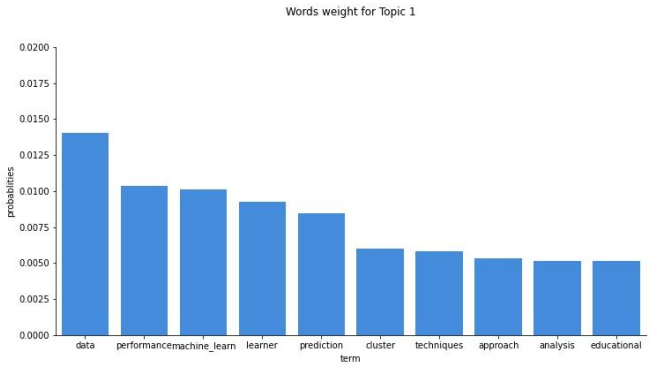

Fig. 9: Words weight under topic 1

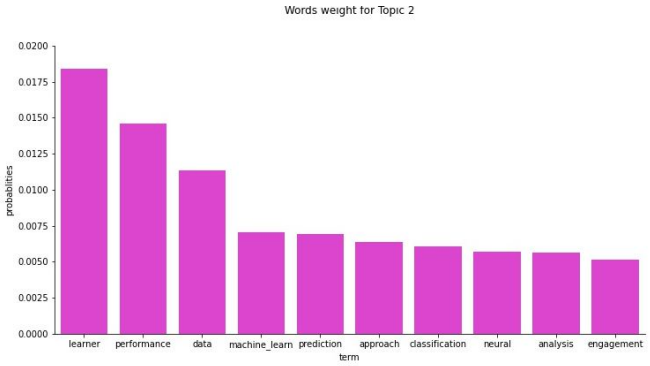

Fig. 10: Words weight under topic 2

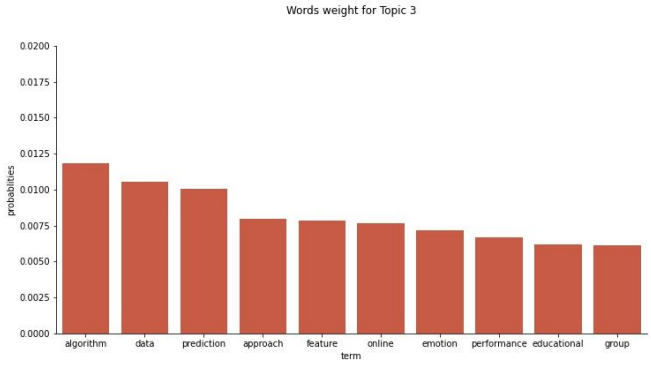

Fig. 11: Words weight under topic 3

The same model was used to extract the top three papers under each topic (papers classification).

TABLE III: Top 3 Research Papers under each Topic

Topics	Top 3 documents (papers)
Topic 1	**1)** Prediction of students procrastination behavior through their submission behavioural pattern in online learning
	2) Predicting students final performance from participation in on-line discussion forums
	3) Near Real-time Comprehension Classification with Artificial Neural Networks: Decoding e-learner Non-Verbal Behavior
Topic 2	**1)** A new e-learning achievement evaluation model based on RBF-NN and similarity filter
	2) Modeling adaptive E-Learning environment using facial expressions and fuzzy logic
	3) Affective database for e-learning and classroom environments using indian students faces, hand gestures and body postures
Topic 3	**1)** Facial emotion recognition with transition detection for students with high-functioning autism in adaptive e-learning
	2) Recognizing and regulating e-learners emotions based on interactive chinese texts in e-learning systems
	3) Using machine learning to predict student difficulties from learning session data

The ranking of papers (Table 3) is measured based on the word weights under each topic, the results are highlighting different aspect of machine learning and deep learning applications in the E-learning context, our last step is suggesting names for the three extracted topics, we will use the top 3 papers titles and the words weights under each topic:

- **Topic 1** : Predicting learners Performance and behavior in on line learning.
- **Topic 2** : Facial expression recognition in E-learning systems
- **Topic 3**: Emotion recognition in E-learning Systems

D. Discussion

During the different phases of our work, papers extraction, preprocessing our corpus and LDA application and its results, we have noticed that the majority of papers are about proposing approaches and models under three main topics, predicting performance and behavior (topic 1), facial recognition (topic 2) and emotion recognition (topic 3), with a distribution of tokens relatively balanced, as we find topic 1 slightly dominant with 41% of tokens, followed by topic 2 and 3 with 33% and 25% for each. The extracted topics doesnt represent all the trends of machine learning and deep learning application in E-learning context, as we only worked on 95 papers, but we believe that we have given a solid overview of the subject which can be useful for any future researches

VI. Conclusion

In this paper, we present a literature survey of machine learning and deep learning applications in E-learning context. This work was achieved by collecting and analyzing 95 papers that we extracted from five research databases: Scopus, Springer Link, Science Direct, Web of Science and IEEE Digital Library. First, we delivered a papers distribution over journals and years, followed by a keywords frequency, and finally we proposed three topics as our LDA topic modeling model suggested. This literature survey might not cover all applications of machine learning and deep learning in E-learning context, considering the limited number of analyzed papers (95) and the journal publication filter used to search for papers in only five indexed scientific databases. Considering all the limitation of our data, results of this paper should be taken with caution.

References

[1] A. Sangr, D. Vlachopoulos, and N. Cabrera, "Building an inclusive definition of e-learning: An approach to the conceptual framework", Int. Rev. Res. Open Distrib. Learn., vol. 13, no. 2, pp.145-159, 2012

[2] T. Ayodele, C. A. Shoniregun, and G. Akmayeva, "Towards E-learning security: A machine learning approach," in International Conference on Information Society (i-Society11), London, UK, June 2011, pp. 490-492.

[3] S. S. Gautam,"Components and benefits of E-learning system," Int. Res. J. Comput. Sci., vol. 3, no. 1, pp. 2014-2017, 2016.

[4] M. L. Sein-echaluce, F. J. Garca-pealvo, and M. A. Condegonzlez,"Learning content management systems for the definition of adaptive learning environments," pp. 105-110, 2014

[5] Bergmann, J., Sams, A. "Flip your classroom: Reach every student in every class every day". Washington, DC: Internal Society for Technology in Education. (2012).

[6] Brusilovsky and C. Peylo, "Adaptive and intelligent Web-based educational systems," Int. J. Artif. Intell. Educ. , vol. 13, no. 2-4, pp. 159-172, 2003.

[7] Wulf, J., Blohm, I., Leimeister, J. M., Brenner, W. "Massive Open Online Courses", Business Information Systems Engineering, vol.6, no 2,pp. 111-114. 2014

[8] McCarthy, J. What Is Artificial Intelligence? Technical report, Stanford University. 2007

[9] Zohuri, B., Moghaddam, M. Business Resilience System (BRS): Driven through Boolean, Fuzzy logics and cloud computation: Real and near real time analysis and decision making system. Business Resilience System (BRS): Driven Through Boolean, Fuzzy Logics and Cloud Computation: Real and Near Real Time Analysis and Decision Making System, 1-425.(2017).

[10] J. McCarthy, "Arthur Samuel: Pioneer in Machine Learning," AI Magazine, vol. 11, no. 3, pp. 10-11, 1990.

[11] Michalski, R. S., Carbonell, J. G., Mitchell, T. M. " Machine Learning: An Artificial Intelligence Approach.", Berlin: Springer Science Business Media, 2013.

[12] Kotsiantis, S.B. "Supervised machine learning: A review of classification techniques.", Informatica, vol. 31, pp. 249-268, 2007.

[13] Ghahramani, Z. "Unsupervised Learning", Chapter in Advanced Lectures on Machine Learning, Vol. 3176, Chapter 5, Springer, pp. 72-112, 2004.

[14] Sutton RS, Barto AG. "Reinforcement learning: an introduction", MIT Press. Cambridge, MA. 1998.

[15] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning", Nature, vol. 521, no. 7553, pp. 436-444, 2015.

[16] L. Yuan, Z. Qu, Y. Zhao, H. Zhang and Q. Nian, "A convolutional neural network based on TensorFlow for face recognition," 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, pp. 525-529, 2017

[17] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, "Large-scale video classification with convolutional neural networks", in IEEE CVPR, pp. 1725-1732, 2014.

[18] X. Zhang, J. Zhao, Y. LeCun, "Character-level convolutional networks for text classification", in:Advances in Neural Information Processing Systems, pp. 649-657, 2015.

[19] Y. Miao, M. Gowayyed, and F. Metze, "EESEN: End-to-end speech recognition using deep RNN models and WFST-based decoding", in Proc. ASRU, pp. 167-174, 2015.

[20] K. Cho, B. van Merrienboer, C. Gulcehre, D.Bahdanau, F. Bougares, H. Schwenk and Y.Bengio, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation", in Proc. 2014 Conf. Empirical Methods in Natural Language Process. (EMNLP), Doha, Qatar, Oct. 2014, pp. 1724-1734.

[21] Rubayyi Alghamdi and Khalid Alfalqi, "A Survey of Topic Modeling in Text Mining", International Journal of Advanced Computer Science and Applications, Vol. 6, No. 1, 2015.
[22] Blei, D. "Introduction to probabilistic topic models". Communications of the ACM, pp. 77-84, 2012
[23] D. OCallaghan, D. Greene, J. Carthy, and P. Cunningham, "An analysis of the coherence of descriptors in topic modeling", Expert Syst. Appl., vol. 42, no. 13, pp. 5645-5657, Aug. 2015