ADICHUNCHANAGIRI UNIVERSITY

B G NAGARA, Nagamangala Taluk, Mandya District – 571448,
Karnataka, India

A TECHNICAL SEMINAR REPORT
ON

"Role of Artificial Intelligence in Online Education : A Systematic Mapping Study"

Submitted in partial fulfillment for the award of degree

BACHELOR OF ENGINEERING
in
ELECTRONICS AND COMMUNICATION ENGINEERING
for the academic year 2023 – 2024

Submitted by:

DIVYA H S **20ECE028**

Under the guidance of:
Dr. Anandaraju M.B, M.E, PhD
Director-HR, ACU,
Professor,Dept of ECE,
BGSIT, B G Nagara.

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

BGS INSTITUTE OF TECHNOLOGY

B G NAGARA, Nagamangala Taluk, Mandya District – 571448
2023-2024

BGS INSTITUTE OF TECHNOLOGY

B G NAGARA, Nagamangala Taluk, Mandya District – 571448

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINERING

CERTIFICATE

This is to certify that the Technical Seminar entitled **"**Role of Artificial Intelligence in Agriculture:An Analysis and Advancements With Focus on Plant Diseases**"**, is a bonafide work carried out by **DIVYA H S (20ECE028)** in partial fulfilment for the award of Bachelor of Engineering in **ELECTRONICS** and **COMMUNICATION ENGINEERING** of the **ADICHUNCHANAGIRI UNIVERSITY**, B G Nagara, during the academic year 2023-2024. It is certified that all the correction/suggestion indicated for internal assessment have been incorporated in the report deposited in the department library. The seminar report has been approved as it satisfies the academic requirement in respect of Technical Seminar work prescribed for the said degree.

Signature of the Guide	**Signature of Technical Seminar Coordinator**	**Signature of the HOD**
Dr. Anandaraju M B,	**Mrs. BALAJI B S**	**Dr. P S PUTTASWAMY**
ME,Ph.D	**Assistant Professor, Dept. of ECE, BGSIT, ACU.**	**Professor & Head, Dept. of ECE, BGSIT, ACU.**
Director -HR, Dept. of ECE, ACU,BGSIT, ACU.		

ACKNOWLEDGEMENT

The completion of any seminar involves the effort of many people. I have been blessed to receive a lot of help and support from all the persons during the making of this seminar. So with gratitude I take this opportunity to acknowledge all those who guided and encouraged me to emerge successful.

It gives me great pleasure in expressing and humble pranams to his holiness **BYRAVAIKYA JAGADGURU PADMA BHUSHAN SRI SRI SRI Dr. BALAGANGADHARANATHA MAHASWAMIJI** and my sincere thanks to the present pontiff his holiness **JAGADGURU SRI SRI SRI Dr. NIRMALANANDANATHA MAHASWAMIJI** and seek their blessings.

It gives me an immense pleasure to write an acknowledgement to this seminar, a contribution of all people who helped me realize it. I would like to convey my heart full thanks to our beloved Principal **Dr. B N SHOBHA**, BGSIT, B G Nagara, for giving support for my academic endeavors.

I am gratefully indebt to **Dr. P S PUTTASWAMY**, Professor and Head, Dept. of ECE, BGSIT, for giving the opportunity to embark upon this topic.

I am also thankful to resourceful guidance, timely assistance and graceful of our Technical Seminar Coordinator **Mr. BALAJI B S**, Assistant Professor, Dept. of ECE, BGSIT, who helped me in every aspect of my seminar work.

I am heartily thankful to **Mrs**. **Aruna B**, Assistant Professor, Dept. of ECE, BGSIT, for her guidance, technical expertise, encouragement and timely help in making this seminar a reality.

I am heartily thankful to **Dr. Anandaraju M B**, Director-HR,ACU, Professor BGSIT, for his guidance, technical expertise, encouragement and timely help in making this seminar a reality.

I am extremely thankful to my parents who have been cooperative and helpful during the entire course of the seminar.

Once again I convey my acknowledgment to all people associated with me for the completion of my seminar. I thank one and all.

DIVYA H S
20ECE028

ABSTRACT

Artificial intelligence (AI) comprises various sub-fields, including machine learning (ML) and deep learning (DL) perform a key role in the transformation of many industries, including education. It changes traditional learning methods by using its Innovative techniques and applications. Using its applications, the teachers may keep track of each student's development, paying close attention to the areas in which they struggle. Many researchers are working with ML and DL to exploit its discoveries and insights. In education, traditional education methods (TEM) are the same for each student, which means each student is taught in the same way as ML and DL, making this process flexible and creative for solving complex problems and enhancing productivity. Nowadays, each institution adopts E-learning methods as the primary way of learning, especially during the pandemic. Despite this evolution of creativity, delivering quality education, making strategies for analyzing performance and future goals, and career counseling for students still pose challenges. Its aim is to provide insights into the applications of these techniques and offer optimal solutions to the research questions at hand. We are convinced that our complete assessment will be a dependable resource for the research group in ascertainment the best approach and information source for their unique needs. Moreover, our findings provide valuable insights on the subject matter that could aid the research community in their future endeavors in the related field.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Nowadays, eLearning is considered the best choice for students and organizations due to its ease in routine life. Modern open education models have made education more accessible to everyone interested in learning about any topic of popular interest . These methods have also enhanced user trust and aided the spread of open education. As a result, there is an increasing tendency toward the dissemination of open educational materials (OER), which enable academic transparency . The benefits of artificial intelligence in education are being recognized by more and more schools, universities, and business entities . In particular, machine learning may improve the effectiveness and fun of the educational process for both students and teachers. ML and DL algorithms make it easier to understand how students acquire knowledge and use it. The education level, interests, present development, and learning style of a student are all taken into consideration. As a result, each student may move forward at their own rate. The probability of leaving someone behind is substantially lower. The teachers may keep track of each student's development, paying close attention to the areas in which they struggle. Such a strategy increases retention and student engage ment rates because learners connect with refined curricula and customized material that is more than standardized ones.

Numerous studies have been undergone on online education for SLR and surveys. For example, in this study , the field of educational data mining (EDM) and the growth of DL applications to EDM were the main topics. With appropriate data anonymization, it has also addressed privacy concerns regarding datasets. Similarly, the study has undergone a comprehensive review of different fields of machine learning. This assessment concentrated on a few of the fields and applications, including social media, network security, banking and finance, and the fields of education and healthcare. Cruz-Jesus et al. represented an innovative strategy for predicting academic success using AI approaches. This study used RF, ANN, SVM, and LR classifiers on various groups of datasets and concluded that AI models perform best when compared to conventional methods. A comprehensive analysis of the literature on current teaching methods and e-learning strategies was presented in the study.

This study compares methods that employ machine learning, crowdsourcing, or even both applications to locate the relevant research that has already been done. According to this study [10], it is possible to measure emotions using the PLS-DA as an emotion classifier, and HRV as a biomarker that discovered that pleasure and melancholy are positively correlated with LF, pNN50, HF, SDNN, RMSSD, LF/HF ratio and pNN50.

As far as we know, earlier research has indicated the use of ML technologies to address online education difficulties, but no complete review has been undertaken to synthesize and reveal the role of machine learning and deep learning in online education. This work employs a systematic mapping methodology to investigate the evolution of ML and DL techniques used to analyze and investigate online education. This review's findings can help with the application of different machine learning and deep learning approaches for the following purposes: • To predict pupil's performance may be at risk in academic institutions. • To predict and determine students' dropout from ongoing classes. • To evaluate static/dynamic based data for students' performance.

Fig 1.1: Role of Artificial Intelligence in Online Education.

CHAPTER 2

LITERATURE SURVEY

Role of artificial intelligence for analysis of COVID-19vaccination related tweets: Opportunities, challenges, and future trends.

- **Authors:** W.Aljedaani, E. Saad, F. Rustam, I. de la Torre Díez, and I. Ashraf.
- **Published Year**: 2022.
- **Proposed Idea**: "Artificial intelligence (AI) has emerged as a powerful tool for analyzing the vast amount of information shared on social media platforms regarding COVID-19 vaccination. This paper explores the opportunities, challenges, and future trends associated with leveraging AI for this purpose. Opportunities include real-time monitoring, sentiment analysis, misinformation detection, targeted communication, and early warning systems. However, challenges such as data quality and volume, bias, privacy concerns, adaptability, and information verification must be addressed. Future trends may involve advancements in natural language processing (NLP) models, integration with public health systems, multimodal analysis, human-AI collaboration, and personalized communication strategies. Understanding these aspects is crucial for harnessing the full potential of AI in facilitating effective public health communication and combating vaccine-related misinformation."

Electroencephalogram signals for detecting confused students in online education platforms with probability-based features

- **Authors :-** T. Daghriri, F. Rustam, W. Aljedaani, A. H. Bashiri, and I. Ashraf.
- **Published Year:** 2022.
- **Proposed Idea**: Detecting and addressing student confusion in online education platforms is crucial for facilitating effective learning experiences. This paper proposes a novel approach that harnesses electroencephalogram (EEG) signals to identify confused students. By utilizing probability-based features extracted from EEG data, the methodology offers a nuanced understanding of cognitive states, enhancing the accuracy of confusion detection. Through real-time analysis of EEG signals, the system aims to provide timely interventions to support student comprehension and engagement. This paper explores the opportunities, challenges, and future trends associated with leveraging EEG signals for detecting confusion in online education platforms, ultimately aiming to enhance learning outcomes and instructions .

Sentimental analysis on online education using machine learning models.

- **Authors:-** S. T. Mathew and L. Jacob.
- **Published Year:** 2022.
- **Proposed Idea: "**This study explores the sentiment analysis of online education using machine learning models to gauge student attitudes and perceptions towards virtual learning platforms. Leveraging natural language processing (NLP) techniques, sentiment analysis aims to extract insights from textual data such as student reviews, forum posts, and social media discussions. Various machine learning algorithms, including support vector machines (SVM), recurrent neural networks (RNN), and transformers, are employed to classify sentiments as positive, negative, or neutral. The research investigates the efficacy of different feature representations, pre-processing methods, and model architectures in accurately capturing the sentiment of online learners. Additionally, the study examines the implications of sentiment analysis for improving online education platforms, enhancing student engagement, and informing instructional design decisions. Through a comprehensive analysis, this research contributes to the understanding of sentiment dynamics in online education and highlights the potential of machine learning models in shaping the future of virtual learning environments."

EXISTING REVIEWS ON ONLINE EDUCATION

Numerous studies have been undergone on online education for SLR and surveys. For example, in this study , the field of educational data mining (EDM) and the growth of DL applications to EDM were the main topics. With appropriate data anonymization, it has also addressed privacy concerns regarding datasets. Similarly, the study has undergone a comprehensive review of different fields of machine learning. This assessment concentrated on a few of the fields and applications, including social media, network security, banking and finance, and the fields of education and healthcare. Cruz-Jesus et al. represented an innovative strategy for predicting academic success using AI approaches. This study used RF, ANN, SVM, and LR classifiers on various groups of datasets and concluded that AI models perform best when compared to conventional methods. A comprehensive analysis of the literature on current teaching methods and e-learning strategies was presented in the study . This study compares methods that employ machine learning, crowdsourcing, or even both applications to locate the relevant research that has already been done. According to this study , it is possible to measure emotions using the PLS-Da as an emotion classifier, and HRV as a biomarker that discovered that pleasure and melancholy are positively correlated with LF, pNN50, HF, SDNN, RMSSD, LF/HF ratio and pNN50.

CHAPTER 3

IMPLIMENTATION

The literature on education utilizing ML and DL approaches is thoroughly reviewed as part of our research employing a systematic mapping methodology. We adhered to the rules of a systematic mapping study (SMS), a technique for impartially compiling and summarizing educational data about our RQs that offers readers a stronger comprehension of insights into the primary issue.To do this, we used a three-step process:

• **Planning:** Publications from digital libraries are located, f iltered, and verified depending on criteria for inclusion and exclusion.

• **Execution:** Journals are examined to weed out pointless studies.

• **Synthesis:** Sorting and analyzing the retrieved data in order to respond to the intended RQs.

Table I :The online repositories utilized to identify relevant articles in this study.

No.	Digital Library	URL
1	ACM Digital Library	https://dl.acm.org/
2	IEEE Xplore	https://ieeexplore.ieee.org/
3	Science Direct	https://www.sciencedirect.com/
4	Scopus	https://www.scopus.com/
5	Web of Science	https://webofknowledge.com/
6	Springer	https://link.springer.com/

Table II. Inclusion and exclusion criteria followed in this study.

No.	Inclusion criteria
1	Published prior before October 2022.
2	Written in English.
3	Available in digital format.
4	Proposed or used machine learning or deep learning.
No.	**Exclusion criteria**
1	Websites, leaflets, reviews, and survey literature.
2	The complete text is not accessible electronically.
3	Paper has not undergone a peer-review process..
4	Retracted papers and thesis papers.
5	Papers only rely on survey-based experiments.
6	A paper solely focused on online education.

PLANNING

The literature search of this study included publications from journals that were indexed in 7 prominent digital libraries, as presented in Table 1. These digital libraries were selected based on their strong scientific foundation and their acceptance and relevance for the study.

i. **CRITERIA FOR INCLUSION AND EXCLUSION:** Incorporating inclusion and exclusion criteria is crucial as it helps in reducing bias, restricting the scope of the search, identifying relevant publications, and eliminating studies that do not align with the study's objectives. The peer-reviewed articles that have been manually screened to fulfill these criteria aid reviewers in determining if the implemented or suggested AI strategies for education are acceptable for the research. We also used backward and forwarded snowballing with the first pool of publications. Table 2 outlines the inclusion and exclusion criteria utilized in this research. The beginning date was not restricted, but the end date was set to October 3, 2022, to ensure that only articles published before that date were included in the study. This approach helped to minimize bias, focus the search, and identify relevant publications while excluding studies that did not align with the research objectives.

ii. **SEARCH STRING :**To ensure the search for relevant articles was comprehensive and precise, a test search was carried out on two prominent digital libraries, namely the Institute of Electrical and Electronics Engineers (IEEE) and the Association for Computing Machinery (ACM). The ACM was selected to determine the appropriate search string for generalizability purposes. The search string was formulated using phrases related to the research questions (RQs) objectives, with the aim of identifying relevant phrases or synonyms used in articles on sentiment analysis of popular perceptions of education. The pilot search was conducted multiple times, with adjustments made to the search keywords as necessary. The search was limited to the title and abstract of articles to increase accuracy and avoid false positives, and only the metadata of publications was included in the search.

EXECUTION

This section discusses the procedure for processing and filtering the publications retrieved from digital library searches. Initially, a pool of 4,079 records was obtained, consisting of article information such as title, abstract, publication year, and digital library name. The digital

library ‘‘Springer’’ had the highest number of publications, with2,262 records. To filter out publications that did not meet our inclusion criteria, a four-phase quality assessment procedure was implemented. Figure 1 illustrates the number of articles filtered at each step. The quality assessment procedure involved a manual review by three authors, and it began with the removal of 186 retracted and duplicate publications in Stage (1). In Stage (2), the remaining 3,893 publications were screened based on their title and abstract, resulting in the exclusion of 3,760 publications that did not use machine learning or deep learning techniques or were not peer-reviewed. In Stage (3), the 28 publications that passed the inclusion and exclusion criteria were subjected to full-text scanning. Here, 105 publications were excluded, resulting in 28 publications for the snowball sampling in Stage (4). This sampling process led to the identification of 7 additional relevant publications. Overall, a total of 35 relevant publications were identified, including 28 publications in the first three phases and 7 in the snowball sampling.

SYNTHESIS

This section examined the data we extracted in this stage to answer the study questions offered (RQs). To begin, we categorized the original collection of papers based on their usage of machine learning and deep learning in online education to answer RQ1. Based on approach analysis, we divided our research into several areas, namely the Covid19 Pandemic, Education System, Learning Results, Student Satisfaction and Feedback, and others. We assigned the task of topic analysis research to the remaining articles that utilized topic extraction techniques such as the Latent Dirichlet Allocation (LDA) algorithm. Moreover, we cataloged the publications by year and location of publication to provide an overview of their features. Furthermore, we examined the parameters of the dataset used in the chosen set of publications to address RQ2, which includes characteristics such as the source of data, the number of records, and the availability status. Lastly, all data relating to the study topics gathered during the publication review process was peer-reviewed, and any disagreements were addressed through talks. To support cooperation efforts throughout the author-review process, we captured the manually extracted data in an Excel spreadsheet.

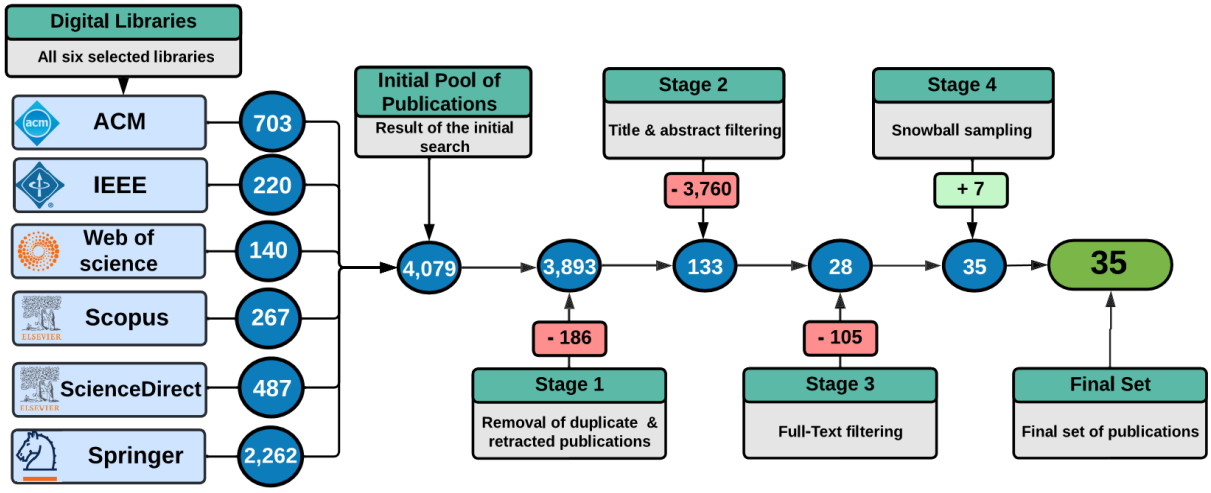

Fig 3.1: A summary of the number of articles received following the filtering procedure.

Machine learning (ML) is quite valuable since it is a highly practical science that offers numerous answers to issues that arise in our daily lives. This approach can help to access students' background, learning speed, feedback, aptitude, and cumulative performance based on various factors recommended by the teacher. Many researchers used this approach in their primary study for example used an ML-based approach for mining students' opinions and TF-IDF for feature extraction.TF-IDF is a measure that determines the importance of certain keywords within specific documents by calculating the frequency of the words and comparing it to the frequency of those words in a larger corpus. According to , K-Nearest Neighbor (KNN) is a fundamental machine learning (ML) model that may be used for regression and classification problems. This approach uses a distance function to classify fresh data based on its nearest neighbors.

- Decision Trees (DT)are tree-like structures that are used to develop models, as described, Because of their ease of use and short execution time, these frameworks are frequently used in the analysis of educational and medical data

- .As detailed in Xtreme Gradient Boosting (XG Boost) is a fast and open-source version of the technique of the gradient-boosted tree, which is used for supervised learning. In order to create precise predictions about a target variable, this strategy integrates the predictions of numerous simpler, weaker models.

- LightGBM (LGBM) is a quick, distributed, high performance gradient-boosting framework that may be used for many different machine learning applications,

including classification and ranking. It is based on the decision tree method .

- Recurrent Neural Network (RNN) deals with modeling, time-dependent, and sequential data issues, such as stock market forecasting, machine translation, and text synthesis. RNN, however, is challenging to train because of the gradient issue. Vanishing gradients are a difficulty for RNNs .

- An Artificial Neural Network (ANN) is a computational model that is designed to mimic the structure and function of the human brain, A comprises many processing components, accepts inputs, and produces results in accordance with predetermined activation functions .

- MLP classifier, in contrast to other classification algorithms like Support Vectors or Naive Bayes Classifier, uses an underlying Neural Network to carry out the classification process .

- Gradient Boosting Machine (GBM) is one of the most often used forward learning ensemble techniques in machine learning. It is an effective method for creating predictive models for problems involving regression and classification .

Table III. List of datasets used in reviewed articles.

ID	Study	Source of Info	Dataset	Dataset size	Dataset Availability
PS1	[19]	Google Form	Questionnaire	5,005	No
PS2	[20]	Google Form	Questionnaire	1,264	No
PS3	[21]	Case study	Interviewees	30	No
PS4	[22]	Google Docs	Questionnaire	2,553	No
PS5	[23]	Google Docs	Questionnaire	263	No
PS6	[24]	Audio Edited for Synchronous Tracks and Organization	Audio	1,513	Request Required
PS7	[25]	Survey	Questionnaire	1,205	No
PS8	[26]	Unknown	Feedback	25,417	No
PS9	[27]	Public opinion	MOOC	460,000	[54]
PS10	[28]	Offline Form	Questionnaire	1,627	No
PS11	[29]	StarC	Questionnaire	32,000	Request Required
PS12	[30]	Google Form	Questionnaire	1,001	No
PS13	[31]	Survey	Questionnaire	647	Request Required
PS14	[32]	Google Form	Questionnaire	170	No
PS15	[33]	Class camera	Real-Time	170	No
PS16	[34]	Class materials	Course	249	No
PS17	[35]	Twitter	Tweets	500	No
PS18	[36]	Google Scholar	Articles	67	No
PS19	[37]	Google Form	Poll	6,500	No
PS20	[38]	K-12 online learning platform	Courses	3,922	No
PS21	[39]	Facebook posts	Questionnaire	55,829	[55]
PS22	[40]	Course Talk	Posts and reviews	5,884	Open Source
PS23	[41]	Google Form	Stats	3,687	No
PS24	[42]	singular events	Multi dataset	63	Request Required
PS25	[43]	Kaggle	OULAD	7	[56]
PS26	[44]	Classroom	Questionnaire	35	No
PS27	[45]	Online LMS	LMS	150	No
PS28	[46]	Survey	Questionnaire	583	No
PS29	[47]	Google Form	Questionnaire	N/A	No
PS30	[48]	PASCAL Visual Object Classes	Images	9963	[57]
PS31	[49]	Live	Emoticon	40	No
PS32	[50]	Kaggle	Public domain dataset	N/A	[58]
PS33	[51]	GitHub	Book ratings	N/A	[59]
PS34	[52]	Kaggle	OULAD	N/A	[58]
PS35	[53]	Google Form	University of Vermont	33	Request Required

The quantitative research methodology known as structural equation modeling (SEM) can also use qualitative techniques. The causal associations between variables are displayed using SEM. The correlations shown by SEM correspond to the researchers' hypotheses. A researcher may be interested in the strength of the relationships between variables in a hypothesis, and SEM is a way to examine those variables . Overall, when combining the statistics, machine learning had the highest implementation rate at 65% (n=23) in the selected studies. This suggests that many researchers were interested in exploring the main topics discussed in this field.

Table IV. Summary of approaches used in education analysis.

ID	Study	Features	Approach	ML	DL	SEM	Hybrid
PS1	[19]	TF-IDF	ML	SVM, RF, MNB, LR	-	NO	-
PS2	[20]	-	ML	KNN, NB, DT, RF, XGB, SVM	-	NO	-
PS3	[21]	-	SEM	-	-	YES	-
PS4	[22]	-	SEM	-	-	YES	-
PS5	[23]	-	ML	SVM, LR, RF,LGBM, DT	-	NO	-
PS6	[24]	-	DL	-	DNN	NO	-
PS7	[25]	-	ML & DL	DT, RF, NB, SVM,KNN	ANN	NO	-
PS8	[26]	-	ML	DT	-	NO	-
PS9	[27]	Heuristics1, Heuristics2, Heuristics3	Hybrid	K-means	-	NO	DUCK, DUCK/BPC DUCK/LDS, DUCK/BPC/LDS
PS10	[28]	Anova	ML	DT,RF	-		-
PS11	[29]	Anova	DL	-	BP, BN	NO	DBNLS
PS12	[30]	EDA, TF-IDF	ML	NB, SVM, LR, RF, DT	-	NO	-
PS13	[31]	-	ML	RF, KNN ,SVM, LR, NB, DT	-	NO	-
PS14	[32]	-	ML	RF, LR, NB, DT	-	NO	-
PS15	[33]	-	ML	SVM	-	NO	-
PS16	[34]	-	ML	DT	-	NO	-
PS17	[35]	-	ML	SVC, LR, NB	-	NO	-
PS19	[37]	-	ML & DL	KNN	RNN ANN	NO	-
PS20	[38]	Periodic, Linguistic	ML	LR, DT, RF	-	NO	-
PS21	[39]	DF, IG, DTM	ML & DL	RF, SVM (Liner), SVM(Radical)	ANN	NO	-
PS22	[40]	DF, IG, DTM	ML	KNN, GBT, SVM, LR, NB	-	NO	-
PS23	[41]	DF, IG, DTM	ML & DL	DT, SVM	DNN	NO	-
PS24	[42]	-	ML & DL	DT,SVM	DNN,ANFIS, ANN, FS	NO	-
PS25	[43]	IMD	ML & DL	RF	ANN	NO	-
PS26	[44]	-	ML & DL	Linear Regression, LR, SVM	ANN	NO	-
PS27	[45]	-	ML	Real, Gentle, Modest, Adaboost, TrMAdaBoost	-	NO	-
PS28	[46]	-	ML	-	-	NO	-
PS29	[47]	-	ML	SVM, NB, KNN, RF	-	NO	-
PS30	[48]	-	DL	-	R-CNN, RPN	NO	-
PS31	[49]	-	SEM	-	-	NO	method1, method2
PS32	[50]	Age, Name	ML	KNN, LR, SVM, DT	-	NO	-
PS33	[51]	-	Hybrid	-	-	NO	EC, MTSVM
PS34	[52]	-	Hybrid	MLP, RF, GBM	-	NO	Rpart,Nnet
PS35	[53]	Statistical Analysis	SEM	-	-	NO	UVM, ERT

Table IV. shows that the most frequently used deep learning classifiers are Artificial Neural Networks (ANN)(n=6),Deep Neural Networks (DNN) (n=3), Recurrent Neural Networks (RNN) (n=1), Region-Based Convolutional Neural Network (R-CNN) (n=1), Region Proposal Network (RPN) (n=1), Back Propagation (BP) (n=1), and Batch Normalization (BN) (n=1), which are commonly selected by researchers. Figure 3 illustrates the distribution of machine learning classifiers.

while illustrates the distribution of deep learning classifiers utilized in the selected pool of publications. From the publications (n = 35), the best-performing machine learning algorithms are DT SVM and LR stood exceptional performances. According to when comparing the performance of machine learning and deep learning models, as well as comparing machine learning and deep learning models to each other, the K-Nearest Neighbor (KNN) algorithm was found to have the highest accuracy.

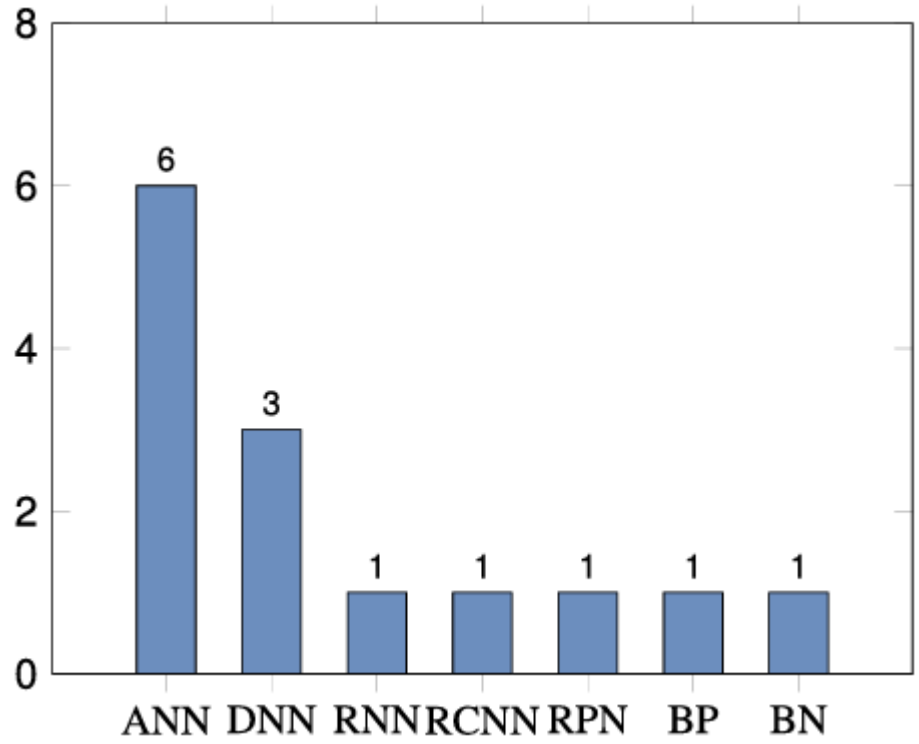

Fig 3.2. Distribution of deep learning-based classifiers used in learning-based approaches.

The purpose of this section is to investigate the feature engineering strategies used in the research included in this study. This section of RQ1 was inspired by the fact that feature engineering can extract predicted data or features for successfully training learning algorithms. As indicated in Table 4, we discovered six feature engineering strategies that are often employed in the chosen articles. We give a quick rundown of the six characteristics of engineering methodologies that have been used for the particular study. The TF-IDF is a statistical measure used in natural language processing (NLP) to assess the importance of a word in a collection of texts. It is determined by dividing the overall inverse document frequency of the word by the number of times it appears in a given document.

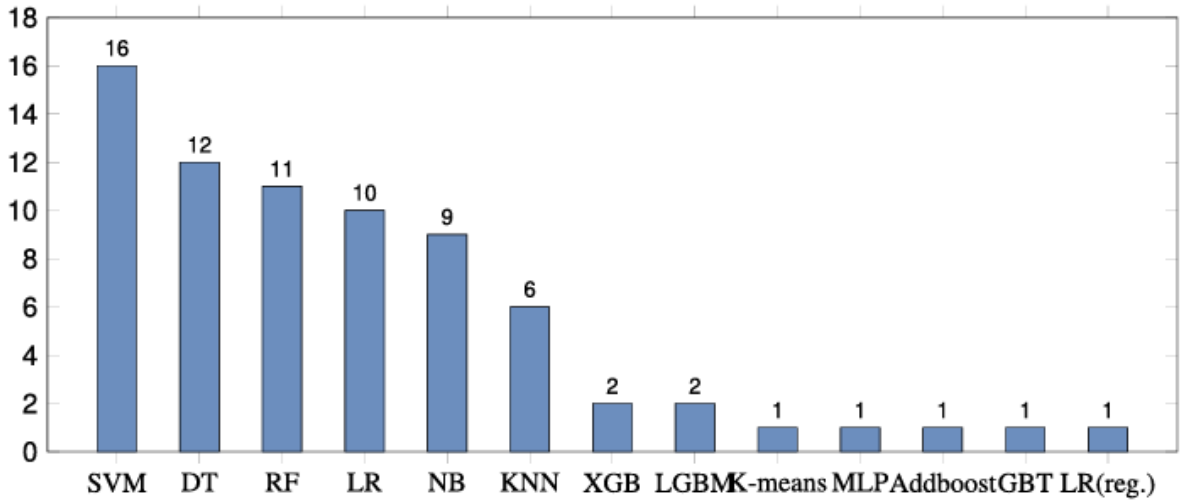

Fig 3.3. Machine learning-based classifier distribution in learning-based techniques.

Table V. An overview of the techniques used for feature engineering in the studies analyzed**.**

No.	Feature Engineering Technique	Total	Publications
1	TF-IDF	2	[19], [30]
2	ANOVA	2	[28], [29]
3	Document frequency(DF)	3	[39]–[41]
4	Information gain (IG)	3	[39]–[41]
5	Document Term Matrix(DTM)	3	[39]–[41]

The technique of information gain (IG) is used to determine the most useful qualities or attributes that give the most information about a specific class. IG is based on the entropy concept and operates by lowering entropy from the root node to the leaf nodes. The objective is to identify the qualities that result in the largest reduction in entropy, hence providing the most useful information for categorization. ANOVA is a statistical approach used to evaluate whether the means of two or more groups of data (typically three or more) are generated from the same distribution. Parametric statistics are used in this exam. A mathematical matrix called a document-term matrix (DTM) reveals the frequency of words used in a group of documents. In a document-term matrix, columns represent terms in the collection, and rows represent documents in the collection.

The most popular feature engineering methods across all four categories of education analysis were found to be Term Frequency-Inverse Document Frequency (TFIDF), Document Term Matrix (DTM), Document Frequency (DF), and Information Gain (IG). Machine learning models are the main users of the feature engineering method. Additionally, deep learning models automatically create features thoroughly explained . In research, feature engineering

approaches were not used in the research that used deep learning models as sentiment classifiers. This research question seeks information on the data sources utilized in the education analysis. The majority of the datasets used were gathered through Google services, such as Google Forms and Google Documents. Surprisingly, 11 of the 35 papers examined used Twitter data. With your Google account, you may now use the free, fully functional forms application known as Google Forms. Standard question kinds may be added, questions can be dragged and dropped into the desired sequence, the form can be customized with basic images or color schemes, replies can be gathered in Forms, or they can be saved to a Google Sheets spreadsheet. The second most important source of data collecting was offline survey forms. With the help of a survey questionnaire, students' opinion is collected. It also includes teachers' options about the system and possible outcomes after going through the change in the current education system. Almost 8 publications out of 35 were utilized to collect data via survey forms. Other resources amount of contributions to the collection of data that resources involve physical activity in classroom (n=2) studies, camera monitoring for getting data of a class (n=1) study, some benchmark publicly available datasets like Kaggle(n=1) study, GitHub (n=1) study, and Twitter (n=1) study used. Table 4 presents a detailed description of the dataset sources .

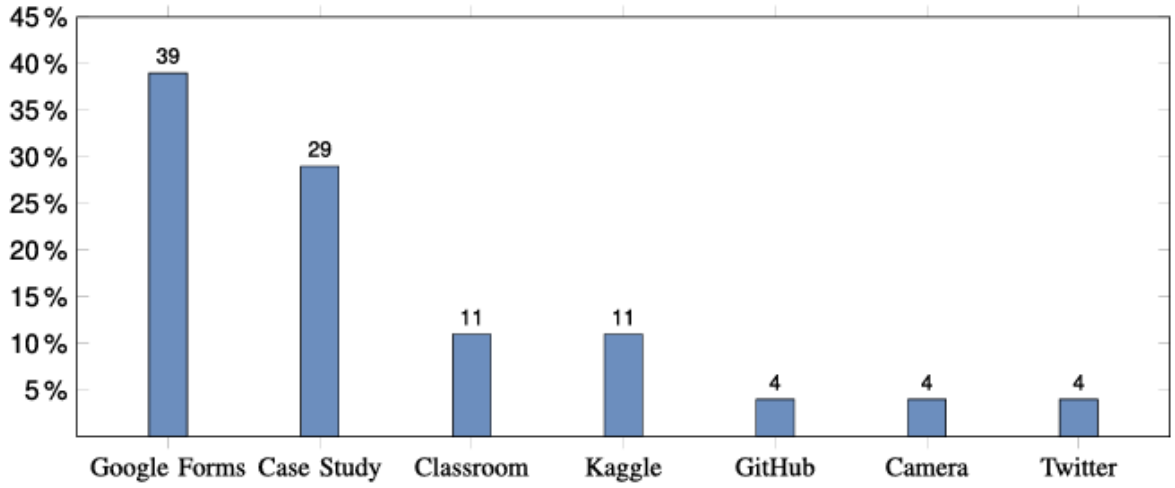

Fig 3.4. Distribution of deep learning classifiers used in machine learning-based approaches.

- Additionally, it was observed that the dropout and prediction of students at risk studies for on-campus students only covered a relatively small percentage of the dataset. This was a point of contention among the researchers. Machine learning algorithms might not yield satisfactory results when trained on relatively small datasets. In addition, the pre-processing data method has the potential.

- In some instances, machine learning (ML) and deep learning (DL) algorithms provided probabilistic predictions without an actual value, which was a critical task for the evaluation of students, necessitating random test human grading to assure fair and reliable ratings.

- There hasn't been enough research on the temporal nature of the indicators used to predict at-risk and dropout pupils. Due to these properties' dynamic nature, their values fluctuate over time. The performance of the predictors can be improved by using temporal data in the classification process. For example, article computes temporal features.

- The majority of research papers approached the issue as a categorization challenge. The methods used to identify the classes of students in the dataset were the topic of very few investigations. Additionally, the aforementioned issues are classified as binary issues, but numerous other classifications would be included to assist management in creating more successful intervention strategies.

- Tasks involving feature engineering, where the features used might affect how well a predictor performs, have received less attention. In the experiments, features like student demographics, TF, IDF, Heuristics and e-learning interaction session logs were predominantly utilized. Additionally, it was found that only a small number of researchers examined the potential of deep learning algorithms like ANN, DNN, RNN, and CNN, instead of the more common classical machine learning algorithms like SVM, DT, NB, and KNN.

- The dynamic character of student performance is not considered in the existing literature. The performance of the pupils is a dynamic process that either improves or degrades with time. Predictor performance on real-time dynamic data is still unknown.

- There have been several different kinds of studies that have focused on the class, or data balance, Maintaining a good class balance is widely regarded as the most significant component in achieving high classification success.

POTENTIAL THREATS TO VALIDITY:

- The study comprises online lectures that were gathered from a variety of digital databases. Thus, we may have overlooked some important ones since they weren't adequately indexed there or weren't indexed in other digital libraries.

- The Study may not have found all pertinent lectures or assignments because the search method was created to look for lectures using phrases occurring in keywords, titles, and abstracts.

- The analysis eliminates scientific studies that are not peer-reviewed, such as book chapters and novels, and instead depends on conferences and peer-reviewed publications. A few studies that carried out systematic literature reviews were also disregarded since they wouldn't have produced accurate data for our research study.

- Stage 2 on our approach, screening was done to include the pertinent research based on the articles' titles, abstracts, and keywords. In a few instances, the title, abstract, and keyword screening alone cannot determine if an article is relevant; instead, a complete paper screening is required. As a result, it is likely that we may have missed certain articles with relevant information because of this problem.

CHAPTER 4

APPLICATIONS

- **Personalized Learning Paths:** AI algorithms can analyze individual student data to create personalized learning paths. These paths can adapt based on the student's progress, preferences, and learning style, ensuring that each student receives customized instruction tailored to their needs.

- **Adaptive Learning Platforms:** AI-powered adaptive learning platforms adjust the difficulty and pace of instruction in real-time based on the student's performance. This ensures that students are appropriately challenged and can progress at their own pace, leading to improved learning outcomes.

- **Intelligent Tutoring Systems:** AI-driven tutoring systems provide personalized support and feedback to students. These systems can answer questions, provide explanations, and offer additional resources based on the student's needs, helping to reinforce learning and address misconceptions.

- **Automated Grading and Feedback:** AI algorithms can automatically grade assignments, quizzes, and exams in online courses. This saves instructors time and allows students to receive immediate feedback on their work, facilitating faster learning and improvement.

- **Content Recommendations:** AI can analyze student data and learning preferences to recommend relevant educational content. This can include articles, videos, simulations, and other resources that supplement course materials and enhance the learning experience.

- **Natural Language Processing (NLP) for Language Learning:** NLP algorithms can analyze written or spoken language to provide language learners with feedback on grammar, pronunciation, and vocabulary usage. These tools can be integrated into language learning platforms to provide immersive and interactive experiences.

- **Virtual Assistants and Chatbots:** AI-powered virtual assistants and chatbots can provide 24/7 support to students. They can answer questions, provide guidance on course navigation, and offer assistance with technical issues.

- **Learning Analytics:** AI can analyze large datasets of student activity to provide insights into learning patterns and trends. Learning analytics can help instructors identify struggling students, monitor engagement levels, and assess the effectiveness of teaching strategies, allowing for data-driven decision-making in course design and instruction.

- **Predictive Analytics for Student Success:** AI algorithms can analyze student data to predict which students are at risk of dropping out or failing. This allows instructors to intervene early with targeted support and interventions, helping to improve student retention and success rates.

- **Accessibility Tools:** AI-powered accessibility tools, such as screen readers and speech recognition software, can help students with disabilities access online course materials. These tools ensure that online education is accessible to all students, regardless of their individual needs or abilities.

CONCLUSION

Recent developments in data-collecting techniques and system performance indicators have made studying educational systems easier and more effective. The analysis and monitoring of massive data using advanced data mining and machine learning techniques have given rise to a new field of big data analytics. Machine learning, along with data mining techniques, has been used to predict student performance and provide feedback to teachers and educators. However, limited research has been conducted on implementing corrective measures promptly.

In the future, research should focus on developing effective ensemble methods to implement machine learning-based performance prediction techniques and find dynamic ways to predict student performance and offer timely remedial actions. Our goal is to apply some of the existing remarkable works and emphasize the dynamic role of machine learning in online education, providing teachers with additional insights to design effective interventions and achieve precise educational objectives.

REFERENCES

[1] D. Zhang, J. L. Zhao, L. Zhou, and J. F. Nunamaker, ‘‘Can e-learning replace classroom learning?’’ Commun. ACM, vol. 47, no. 5, pp. 75–79, May 2004.

[2] J.S.BrownandR.Adler,‘‘Openeducation,thelongtail,andlearning2.0,’’ Educause Rev., vol. 43, no. 1, pp. 16–20, 2008.

[3] S.Baldi, H. Heier, and A. Mehler-Bicher, ‘‘Open courseware and open source software,’’ Commun. ACM, vol. 46, no. 9, pp. 105–107, Sep. 2003.

[4] E.Scanlon, ‘‘Scholarship in the digital age: Open educational resources, publication and public engagement,’’ Brit. J. Educ. Technol., vol. 45, no. 1, pp. 12–23, Jan. 2014.

[5] B.du Boulay, ‘‘Artificial intelligence as an effective classroom assistant,’’ IEEE Intell. Syst., vol. 31, no. 6, pp. 76–81, Nov. 2016.

[6] M.Injadat, A. Moubayed, A. B. Nassif, and A. Shami, ‘‘Machine learning towards intelligent systems: Applications, challenges, and opportunities,’’ Artif. Intell. Rev., vol. 54, no. 5, pp. 3299–3348, Jun. 2021.

[7] F.Cruz-Jesus, M. Castelli, T. Oliveira, R. Mendes, C. Nunes, M. Sa Velho, and A. Rosa-Louro, ‘‘Using artificial intelligence methods to assess academic achievement in public high schools of a European union country,’’ Heliyon, vol. 6, no. 6, Jun. 2020, Art. no. e04081.

[8] H .S. Alenezi and M. H. Faisal, ‘‘Utilizing crowdsourcing and machine learning in education: Literature review,’’ Educ. Inf. Technol., vol. 25, no. 4, pp. 2971–2986, Jul. 2020.

[9] J.W. Y. Chung, H. C. F. So, M. M. T. Choi, V. C. M. Yan, and T. K. S. Wong, ‘‘Artificial intelligence in education: Using heart rate variability (HRV) as a biomarker to assess emotions objectively,’’ Comput. Educ., Artif. Intell., vol. 2, 2021, Art. no. 100011

[10] W.Aljedaani, R. Krasniqi, S. Aljedaani, M. W. Mkaouer, S. Ludi, and K. Al-Raddah, ‘‘If online learning works for you, what about deaf students? Emerging challenges of online learning for deaf and hearing impaired students during COVID-19: A literature review,’’ Universal Access Inf. Soc., pp. 1–20, Jul. 2022.

[11] W.Aljedaani, E. Saad, F. Rustam, I. de la Torre Díez, and I. Ashraf, ‘‘Role of artificial intelligence for analysis of COVID-19 vaccination related tweets: Opportunities, challenges, and future trends,’’ Mathematics, vol. 10, no. 17, p. 3199, Sep. 2022.