Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

MULTI-AGENT E-LEARNING SYSTEM BASED ON ENHANCED UNIVARIATE AND PREDICTIVE EXTRA TREE TECHNIQUES FOR SUSTAINABLE DEVELOPMENT

Viswanath K[a*], Ayshwarya Lakshmi S[b], Anandavelu K[c]
[a,c]Department of Mechanical Engineering, MRK Institute of Technology, Cuddalore, Tamilnadu, India.
[b]Department of Computer Science and Engineering, University College of Engineering Panruti, Tamilnadu, India.
Email: vishwamrk73@gmail.com; ayshwaryalakshmi@ymail.com; anandavelu@mrkit.in

Abstract: The multi-agent system is an effective system to inculcate knowledge through online mode. In this research work, two feature selection techniques, namely enhanced univariate and predictive extra tree have been proposed. These feature selection techniques are used to communicate between the multiple agents. The feature selection method proposed in this work is to predict the students' result. Machine learning algorithms have been employed to produce better results by selecting the relevant features from the database. The parameters evaluated are accuracy, precision, recall, and the F measure. The random forest algorithm has produced better results during the parameter analysis and the naïve bayes algorithm has produced comparatively poor results. Thus, the random forest is the optimized one for the proposed e learning multi agent system.

Keywords: Multi-agent, Univariate, Extra tree, Machine learning, Accuracy

I. INTRODUCTION

The occurrence of global disasters within the past two years has compelled individuals to rely on technology to access services remotely. By utilizing appropriate tools, individuals have the potential to complete tasks more efficiently and at a reduced cost [1]. There is a significant level of enthusiasm surrounding the field of artificial intelligence (AI), thereby presenting opportunities to leverage machine learning techniques to enhance productivity and efficacy. Proficiency in effective communication can yield significant advantages across various domains, including education, healthcare, business, and culture. Machine learning strategies, a widely recognized artificial intelligence component, are increasingly employed within the expanding educational landscape [2]. There is a significant emphasis on enhancing learning systems within the field of education, with particular attention being placed on online education platforms.

The rationale behind this assertion is rooted in the observation that contemporary educational settings exhibit notable disparities when compared to their historical counterparts. One of our objectives is to establish an intelligent system that facilitates the administration of online courses. Due to the continuous evolution, novel approaches to utilizing e-learning have emerged, effectively addressing students' contemporary requirements [3]. Utilizing an online learning platform would allow students to access the tool at their convenience, regardless of time or location. Utilizing the multiagent model is necessary for users of the e-learning system to compensate for the absence of an optimal learning environment [4]. A multi-agent approach is deemed the most optimal strategy for effectively managing the e-learning system due to the diverse attributes involved in the e-learning process. The potential for enhanced collaboration among agents within a multiagent system arises from their ability to communicate within a shared environment. The utilization of a multiagent system facilitates the exchange of information among e-learning tools, enabling them to ascertain their interdependencies. The e-learning platform provided students with diverse resources that facilitated their overall academic performance. The suggested multiagent system can potentially enhance various categories of student assignments on e-learning platforms. By employing effective feature selection strategies, artificial intelligence (AI)-enabled systems can potentially enhance academic performance among students within educational settings [5]. Machine learning refers to using algorithmic models designed to analyze data to identify patterns, make predictions, and determine optimal courses of action. Machine learning algorithms are increasingly gaining significance across various domains, with a particular emphasis on education. Machine learning algorithms significantly impact both the process of learning and the selection of features [6]. Feature selection algorithms possess the capability to efficiently analyze extensive datasets and identify the most significant features, thereby enabling accurate predictions. Numerous feature selection algorithms exist, each possessing the capacity to enhance productivity and reveal latent features. This paper presents a novel approach in the form of a multiagent-based educational system to investigate the impact of interacting agents on online education. Multiple agents, such as courses, students, and activities, were integrated, and various feature selection techniques were employed to refine the data, retaining only the attributes that would contribute most significantly to enhancing the e-learning experience [7].

II. RELATED WORKS

Machine learning, alternatively referred to as ML, is a specialized domain within the realm of artificial intelligence (AI) that enables machines to acquire knowledge from previously gathered data, thereby enhancing their overall proficiency in accomplishing a designated objective [8]. The incorporation of artificial intelligence techniques plays a pivotal role in enhancing predictive accuracy and optimizing operational efficiency. Before utilizing machine learning algorithms, it is imperative to possess knowledge regarding feature selection, commonly referred to as feature selection (FS). Using feature selection techniques makes it possible to selectively identify and retain the most significant and valuable components within a dataset while simultaneously discarding those with lesser importance or utility [9]. Many researchers have begun employing feature selection strategies and machine learning algorithms to enhance student's education quality. The authors propose a fuzzy methodology for determining the academic outcome of a student, specifically whether they have failed or not. Students learning outcomes can be influenced by their prior academic background, level of engagement within the classroom, and pre-existing knowledge about the subject matter. A fuzzy algorithm was employed, incorporating multiple criteria, in order to rank the students. This ranking could subsequently be utilized to infer the potential academic performance of each student. The dataset consisted of 131 students from three distinct schools, with each student possessing 22 distinct characteristics.

The researchers employed machine learning algorithms to forecast the academic performance of novice computer science and information technology students. Supervised machine learning algorithms were devised to predict the outcomes of tests through a two-step process [10]. The data must undergo thorough cleansing and meticulous preparation before being processed by machine learning algorithms for performance prediction. A series of supervised algorithms were examined, and it was determined that the logistic regression classifier yielded the most favourable results for a sample size of 498 students. Instead of employing alternative algorithms, the authors propose using a decision tree. This algorithm exhibits dissimilar characteristics compared to the other three algorithms. The data collected were analysed using Weka tools, which provided insights into the project's potential success. The researchers examined the variables that impact precision by evaluating the model's efficacy in identifying pertinent features [11].

III. PROPOSED WORK

The block diagram of the proposed multi-agent learning system is shown in Figure 1. The various steps involved are data collection, pre-processing of the collected data, integration of the data sets, extraction of features, segmenting of the datasets, optimization, and evaluation of the machine learning algorithms.

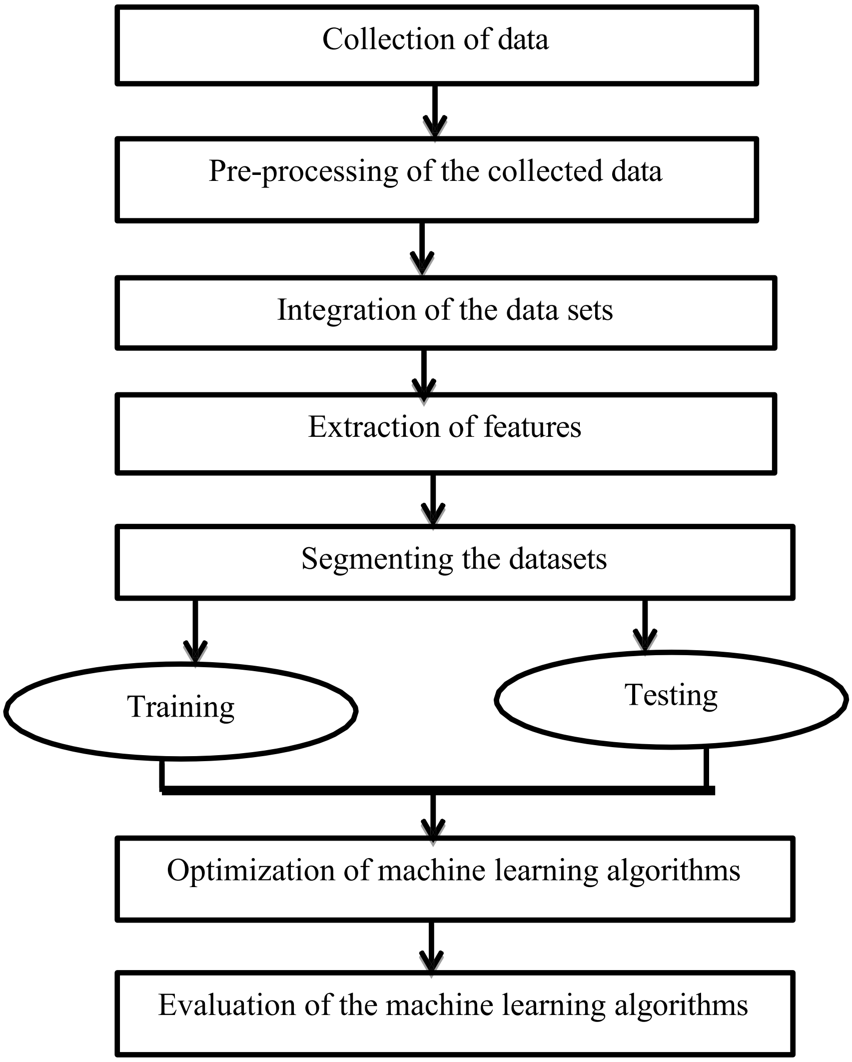

Figure 1. Block diagram of the proposed multi-agent learning system

A. Evaluation of the Machine Learning Algorithms

The evaluation of the machine learning algorithms is based on the four parameters: accuracy, precision, recall, and F measure.

$$accuracy = \frac{True\ positive + True\ negative}{True\ positive + False\ positive + True\ negative + False negative}$$

$$precision = \frac{True\ positive}{True\ positive + False\ postive}$$

$$recall = \frac{True\ positive}{True\ positive + False negative}$$

$$F\ measure = \frac{2 \times precision \times recall}{precision + recall}$$

Table 1: The features used to measure the performance along with their ranks

S.No	Feature	Rank	S.No	Feature	Rank
1.	Home-site	12.74	11.	page	2.52
2.	questions	12.32	12.	survey	2.32
3.	Content	11.72	13.	glossary	1.92
4.	Sub-site	9.34	14.	Presentation of the code	1.82
5.	forumng	7.76	15.	External quiz	1.57
6.	resource	6.59	16.	Data plus	1.03
7.	URL	6.14	17.	illuminate	0.62
8.	Module of code	5.15	18.	Html actions	0.47
9.	wikki	4.23	19.	Shared site	0.12
10.	cooperate	2.76	20	Repeat actions	0.01

Table 1 gives the features used to measure the performance and their ranks. The features used for the classification are listed in the table. The top 20 features are listed along with the ranking. The top 10 features are home site, questions, content, subsite, forum, resource, URL, code module, wiki, and cooperate. The following 10 features are page, survey, glossary, presentation of the code, external quiz, data plus, illuminate, HTML actions, shared site, and repeat actions. These features are used for the analysis and classification of the different types of algorithms

IV. RESULTS AND DISCUSSION

Five algorithms' accuracy, precision, recall, and F measures have been evaluated. Figure 2 shows the accuracy, precision, recall, and the F measure of the decision tree algorithm and the K nearest neighbour algorithm are 82.35, 82.54, 82.47, 82.54, and 83.52, 84.63, 83.45, and 83.62, respectively, through cross-validation. The accuracy, precision, recall, and the F measure of the Naïve Bayes algorithm and the linear regression algorithm are 69.12, 71.03, 69.52, 68.54 and 80.57, 80.52, 80.46, 80.35, respectively, through cross-validation. The accuracy, precision, recall, and the F measure of the random forest algorithm are 88.32, 88.24, 88.23, and 88.52, respectively, through cross-validation. The accuracy, precision, recall, and the F measure of the decision tree algorithm and the K nearest neighbour algorithm are 80.52, 80.18, 80.82, 80.82, and 82.15, 82.53, 82.56, and 82.64, respectively, through testing.

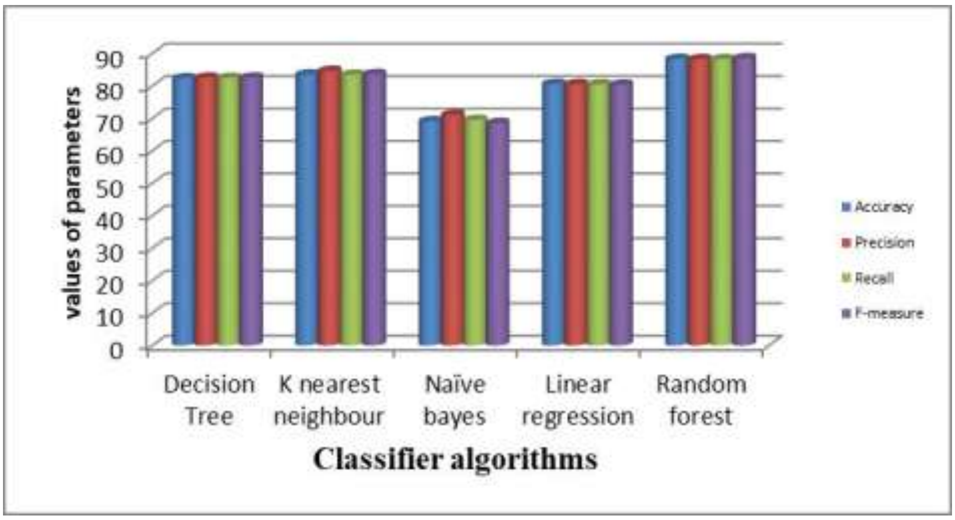

Figure 2: Performance analysis by the application of classifier to all the features through cross-validation.

Figure 3 shows the accuracy, precision, recall, and the F measure of the Naïve Bayes algorithm and the Linear regression algorithm are 69.52, 70.74, 69.57, 68.35 and 80.52, 80.36, 80.36, 80.28, respectively, through testing. The accuracy, precision, recall, and the F measure of the random forest algorithm are 86.34, 87.25, 86.47, and 86.89, respectively, through testing.

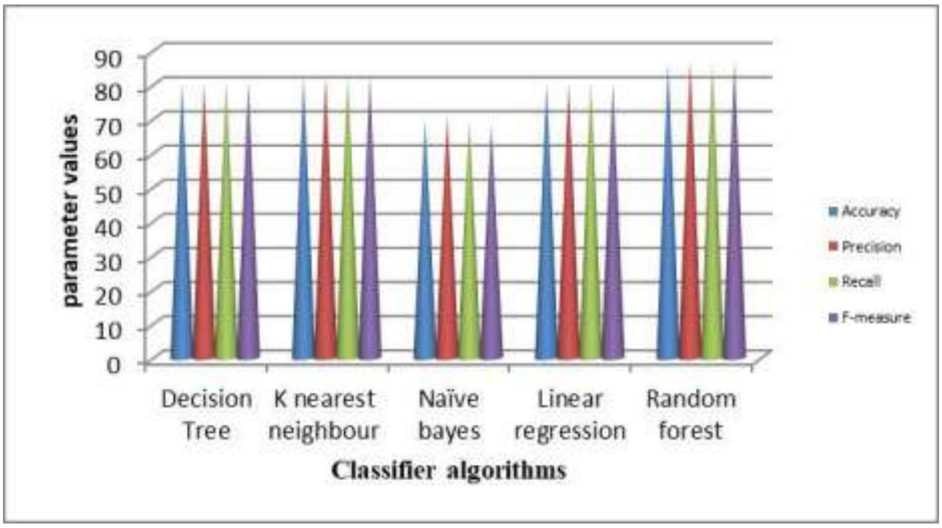

Figure 3: Performance analysis by the application of classifier to all the features through Testing

Figure 4 shows the accuracy, precision, recall, and the F measure of the decision tree algorithm and the K nearest neighbour algorithm are 80.21, 80.25, 80.52, 80.52, 83.65, 84.63, 83.65, and 83.65, respectively, through cross-validation. The accuracy, precision, recall, and the F measure of the Naïve Bayes algorithm and the Linear regression algorithm are 66.58, 66.57, 66.18, 65.59, and 79.23, 78.35, 78.34, 79.36, respectively, through cross-validation. The accuracy, precision, recall, and the F measure of the random forest algorithm are 86.87, 86.95, 86.39, and 86.94, respectively, through cross-validation.

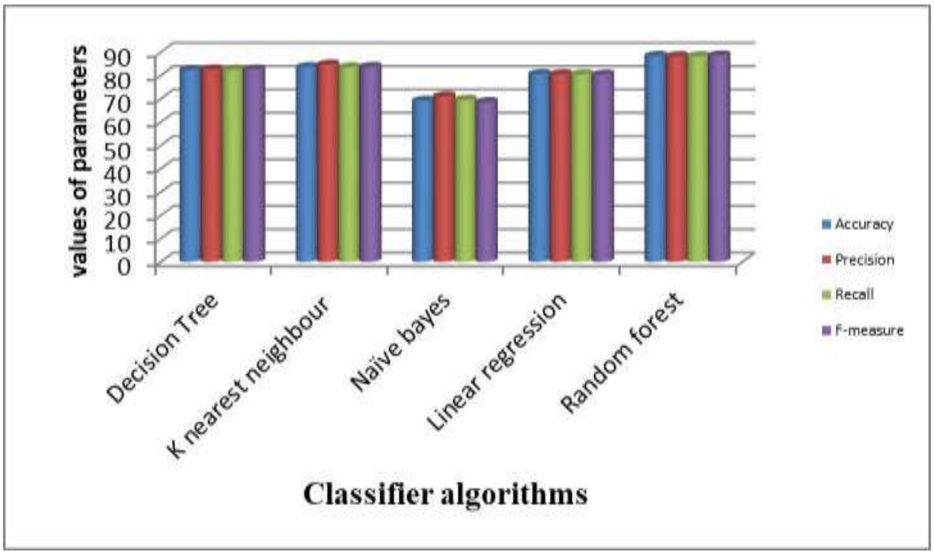

Figure 4: Cross-validation performance analysis of KNN and RF algorithms by applying a classifier to 13 features through a univariate feature selection technique.

Figure 5 shows the cross-validation performance analysis of KNN and RF algorithms by applying a classifier to 13 features through a univariate feature selection technique. The accuracy, precision, recall, and the F measure of the decision tree algorithm and the K nearest neighbour algorithm are 78.36, 78.56, 78.69, 78.34, and 82.65, 82.64, 82.36, and 82.69, respectively, through testing.

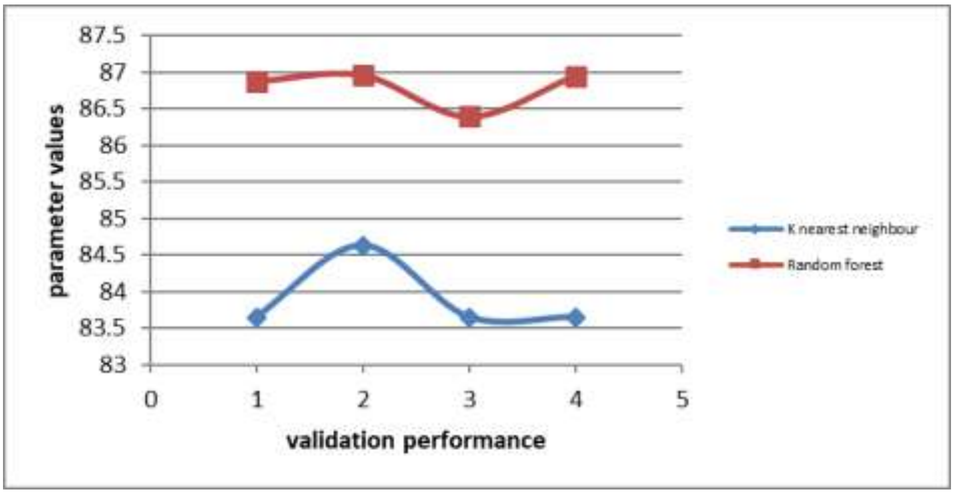

Figure 5: Testing performance analysis of KNN and RF algorithms by applying the classifier to 13 features through a univariate feature selection technique

Figure 6 shows the accuracy, precision, recall, and the F measure of the Naïve Bayes algorithm and the Linear regression algorithm are 65.14, 65.23, 65.24, 64.58, and 79.36, 79.65, 79.44, 79.67, respectively, through testing. The accuracy, precision, recall, and the F measure of the random forest algorithm are 85.92, 85.34, 85.36, and 85.31, respectively, through testing.

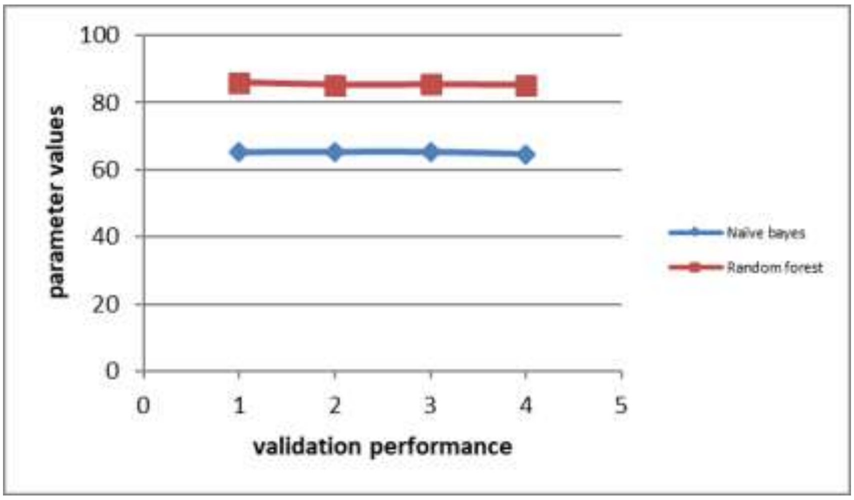

Figure 6: Cross-validation performance analysis by the application of classifier to 13 features through extra trees feature selection technique

Figure 7 shows the accuracy, precision, recall, and the F measure of the decision tree algorithm and the K nearest neighbour algorithm are 82.58, 82.64, 82.67, 82.36, and 83.64, 84.69, 83.69, and 83.94, respectively, through cross-validation. The accuracy, precision, recall, and the F measure of the Naïve Bayes algorithm and the Linear regression algorithm are 68.94, 69.58, 68.45, 67.25 and 80.36, 80.31, 80.21, 80.22, respectively, through cross-validation. The accuracy, precision, recall, and the F measure of the random forest algorithm are 87.59, 88.33, 87.36, and 87.36, respectively, through cross-validation.

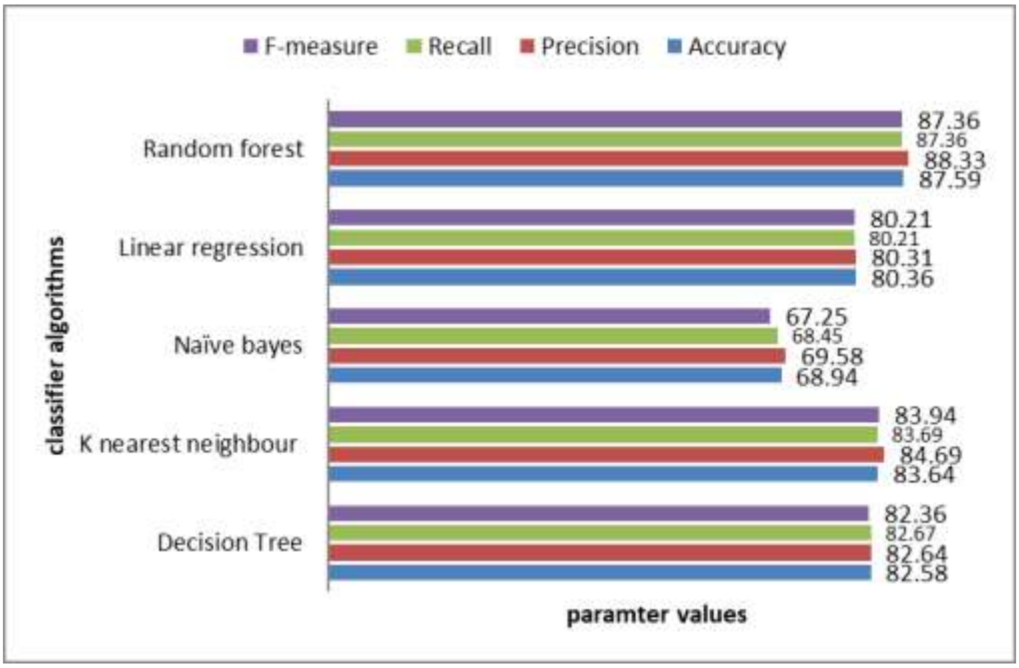

Figure 7: Cross-validation performance analysis by applying the classifier to 13 features through the extra trees feature selection technique.

Figure 8 shows the accuracy, precision, recall, and the F measure of the decision tree algorithm and the K nearest neighbour algorithm are 82.26, 82.35, 82.36, 82.35, and 82.36, 82.64, 82.65, and 82.69, respectively, through testing. The accuracy, precision, recall, and the F measure of the Naïve Bayes algorithm and the Linear regression algorithm are 68.51, 69.34, 68.95, 67.59 and 80.24, 80.34, 80.64, 80.65, respectively, through testing. The accuracy, precision, recall, and the F measure of the random forest algorithm are 86.91, 87.69, 86.57, and 86.71, respectively, through testing

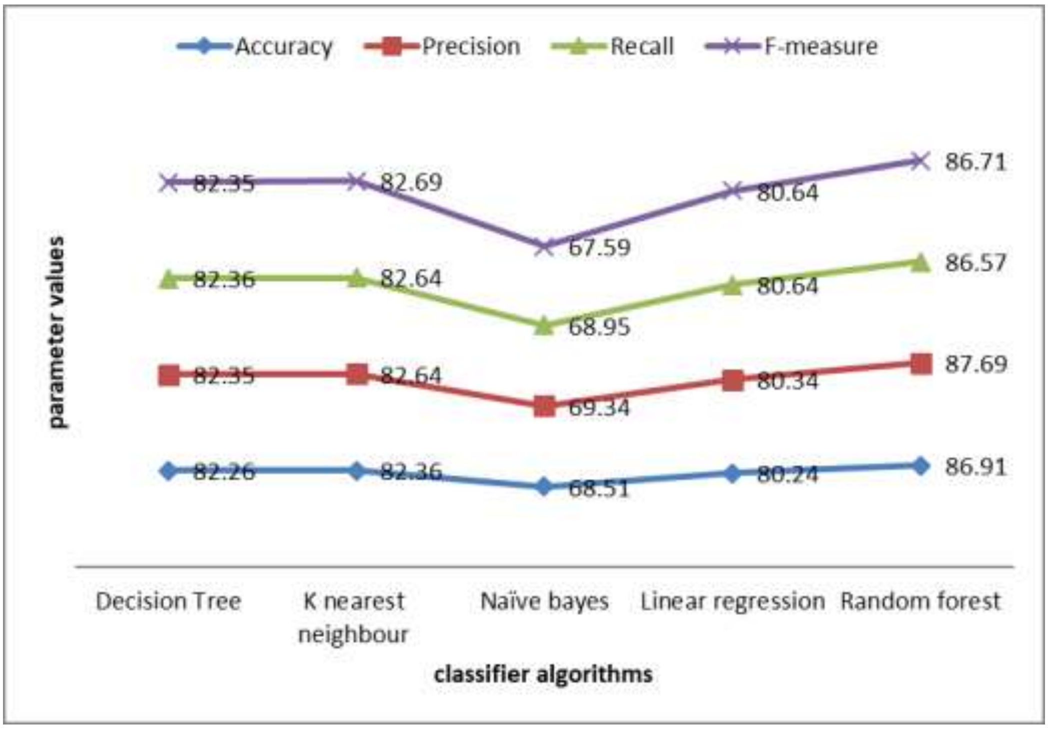

Figure 8: Testing performance analysis by the application of classifier to 13 features through extra trees feature selection technique

CONCLUSION

The performance of the multi-agent e-learning system has been analysed in this research work using feature selection techniques and machine learning algorithms. The feature selection techniques are the extra tree and the univariate methods. The machine learning algorithms used here are the random forest algorithm, decision tree algorithm, logistic regression algorithm, KNN algorithm, and the naïve Bayes algorithm. The parameters considered for the analysis are accuracy, precision, recall, and the F measure. The results have been tabulated and depicted by graphical representations. Among the evaluated algorithms, the random forest algorithm has produced better results in terms of accuracy, precision, recall and F measure. The K means algorithms has produced good results but not more than the random forest algorithm. Naïve bayes has produced the lowest result among the chosen algorithms. The future work could involve the enhancement of the random forest algorithm to improve the performance with the same data sets.

REFERENCES

1. Lokmic-Tomkins, Zerina, Dinesh Bhandari, Chris Bain, Ann Borda, Timothy Charles Kariotis, David Reser: Lessons learned from natural disasters around digital health technologies and delivering quality healthcare. International Journal of environmental research and public health, **20** (5), 4542 (2023).
2. Tominc, Polona, Maja Rožman: Artificial Intelligence and Business Studies: Study Cycle Differences Regarding the Perceptions of the Key Future Competences. Education Sciences, **13** (6), 580 (2023).
3. Mikić, Vladimir, Miloš Ilić, Lazar Kopanja, Boban Vesin: Personalisation methods in e-learning: A literature review. Computer Applications in Engineering Education, **30** (6), 2022, 1931-1958.
4. Apoki, ufuoma chima, aqeel m. Ali hussein, humam k. Majeed al-chalabi, costin badica, mihai l. Mocanu. The role of pedagogical agents in personalized adaptive learning: A review. Sustainability, **14** (11), 6442 (2022).
5. Ouyang, Fan, Mian Wu, Luyi Zheng, Liyin Zhang, Pengcheng Jiao: Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. International Journal of Educational Technology in Higher Education, 20 (1), (2023), 1-23.
6. K. Kumar, A.S. Suresh, S. Radhamani, T. Sundaresan, Ananth kumar: Medical image classification and manifold disease identification through convolutional neural networks: a research perspective. Handbook of Deep Learning in Biomedical Engineering and Health Informatics, (2021), 203-225.
7. Shokouhifar, Mohammad, Nazanin Pilevari: Combined adaptive neuro-fuzzy inference system and genetic algorithm for E-learning resilience assessment during COVID-19 Pandemic. Concurrency and Computation: Practice and Experience, 34 (10) e6791 (2022).
8. S. Janani, R. Dilip, Suryansh Bhaskar Talukdar, Veera Bhaskar Talukdar, Krishna Nand Mishra, Dharmesh Dhabliya: IoT and Machine Learning in Smart City Healthcare Systems. In Handbook of Research on Data-Driven Mathematical Modeling in Smart Cities, IGI Global, (2023), 262-279.
9. França, Tiago Jacob Fernandes, Henrique São Mamede, João Manuel Pereira Barroso, Vítor Manuel Pereira Duarte Dos Santos: Artificial intelligence applied to potential assessment and talent identification in an organizational context. Heliyon, 9 (4), 2023.
10. S. Arunmozhiselvi, T. Ananth kumar, P. Manju bala, S. Usharani, G. Glorindal: A Systematic Approach to Agricultural Drones Using a Machine Learning Model. In Machine Learning Approaches and Applications in Applied Intelligence for Healthcare Data Analytics, CRC Press, (2022), 41-60.
11. Konstantakos, Vasileios, Anastasios Nentidis, Anastasia Krithara, Georgios Paliouras: Crispr–Cas9 gRNA efficiency prediction: an overview of predictive tools and the role of deep learning. Nucleic Acids Research, **50** (7), (2022), 3616-3637.