

Natural Inspired Learning Path Recommendation System for Students in E-learning Platforms

Gokul Kottilapurath Surendran¹, P. Thiyagarajan²

¹Department of Computer Science, Central University of Tamil Nadu, Thiruvavur 610 005, India

²Department of Computer Science, Rajiv Gandhi National Institute of Youth Development 602 105, India

¹gokulkscientist@gmail.com

²thiyagu.phd@gmail.com

Abstract: E-learning has emerged as one of the advanced and efficient methods of education, especially during the ongoing pandemic. While many learning platforms have integrated recommender systems to aid learners in understanding and improving their skills, it has been observed that this approach alone may not be sufficient for optimal learning outcomes. Therefore, this paper proposes a unique method and algorithm inspired by bird swarms to recommend an efficient learning path for learners. The effectiveness of the proposed model was evaluated using nearly three hundred and ninety-five student records, and qualitative evaluation metrics were used to assess its performance.

Keywords: E-learning, Nature-inspired algorithms, Recommendation systems, Migrating birds optimization.

1. Introduction

Education is widely regarded as one of the most powerful tools available to individuals. It provides learners unique insights and helps them develop new ideas and personal improvements. Additionally, learning is a lifelong process, with no limitations or particular criteria for acquiring knowledge-it can occur anywhere and at any time.

It is widely recognized that individuals learn from their experiences, which may be either positive or negative, with the latter referred to as reinforcement learning [21]. As per behavior psychology, it is believed that humans and animals tend to repeat behavior that leads to positive rewards and avoid those that lead to negative rewards [2].

In formal education systems, learners are typically taught by educators who follow a pre-planned curriculum. Most education systems focus on on-campus learning; it provides more information and real-time doubt clarification. According to Marc Rosenberg [14], E-learning is a form of distance learning that can take place in either synchronous or asynchronous modes. In other words, E-learning is a mode of learning that relies on electronic devices such as computers, smartphones, and the internet to facilitate the acquisition of knowledge and skills.

P. Thiyagarajan

Department of Computer Science, Rajiv Gandhi National
Institute of Youth Development 602 105, India
thiyagu.phd@gmail.com

One of the key advantages of E-learning is its ability to provide educational content to learners in virtually any geographic location where they can access reliable electricity and internet connectivity. This concept changed the traditional educational scenario into a fruitful learning method. Moreover, most of the available learning methods are Asynchronous; that is, the learners can study and take exams at any time based on their convenience. This feature attracts more learners to E-learning. Universities like Harvard University and MOOC providers like Coursera and Udemy provide different online courses.

E-learning platforms are shifting away from traditional marks-based educational systems and towards more result-oriented approaches. It has been observed that many developed countries have begun to shift towards result-oriented educational approaches in their on-campus learning systems. However, in developing countries such as Indonesia and India, traditional marks-based education continues to be the norm [17]. It is recognized that implementing a truly result-oriented education platform can be a complex and challenging endeavor. As such, online courses have increasingly turned to real-time project works to promote practical, applicable learning outcomes for their learners. At the same time, exams remain an important means of evaluating learners and assessing their progress. However, the vast amount of data generated by these exams can pose a significant challenge in managing and analyzing it effectively.

Technological advancements have enabled significant reductions in the workforce required to perform many tasks, resulting in the generation of vast data. In recent years, the sheer volume of data produced has grown exponentially, leading to new challenges and opportunities for researchers and practitioners [20]. Data has become a critical resource for all industries, including health care, business, and many others. However, working with such data can be a complex and challenging task, as it often contains a vast amount of information that may be interconnected or intra-connected. Each type of data serves specific needs and purposes. Therefore, simply discarding data is not a viable option. Instead, the focus should be on maximizing its utilization and extracting insights and knowledge that can drive positive outcomes and impacts. However, manual analysis of large-scale data sets can be time-consuming and impractical, which is why many

analytical tools have been developed to facilitate this process [10]. Data analytical tools are designed to assist with initial data analysis, removing partial data and other types of errors similar to what a human analyst might do in real-time.

The adoption of the Internet of Things (IoT) and its associated services has significantly increased the volume of data stored in the cloud [12]. The majority of contemporary technological devices and services incorporate IoT in order to enhance their efficiency. As a result, it has become increasingly important to employ data mining and knowledge extraction techniques across all fields associated with the internet.

The proliferation of data has given rise to challenges in the realm of real-time applications and other related architectures. While data pre-processing and analysis can help address these challenges, it is important to ensure that these processes are utilized effectively. Having access to flawless data and patterns can certainly simplify business tasks, but what about opportunities for improvement? This is where recommendation systems [7] come into play. By effectively leveraging pre-processed data, recommendation systems can help optimize business strategies. It is worth noting that the impact of recommendation systems is particularly evident in the online video streaming and marketing domains, where they are becoming increasingly prevalent. This is partly because approximately 90% of recommended items tend to achieve success. Take E-commerce giants like Amazon and Walmart, for instance; a substantial proportion of their recommended items are sold with remarkable efficacy. This highlights the potency of recommendation systems in capturing the attention of consumers. Beyond the retail space, online streaming platforms such as Netflix and YouTube have greatly leveraged recommendation systems.

Technology is advancing at an unprecedented pace, with numerous fields benefiting from these advancements. Despite the growing prevalence of recommendation systems across various domains, their effective integration into E-learning platforms has yet to be widely explored. Many E-learning platforms have implemented conventional recommendation systems, which rely on analysing user data and generating recommendations based on historical usage patterns. However, given the fundamental differences between business and

education, it is crucial to develop tailored scenarios that are more suitable for educational purposes. Designing an E-learning course or platform entails extensive groundwork and requires in-depth research to ensure seamless deployment of the service. Academic learning is especially complex, as it caters to a wide range and is not limited by language barriers in modern education. As such, the course structure and delivery medium carry significant weight and should be approached with great care and attention to detail.

Sofiadin [18], studied how e-learning can sustain education and foster a humanising learning environment. Authors have sought to understand students' perceptions, challenges, and potential benefits of e-learning about supporting education and humanising learning experiences. The study included conducting focus group sessions with a sample of students. The participants engaged in discussions centred around the advantages and disadvantages of e-learning and how it could promote a humanising educational experience. The authors collected data by recording and transcribing these focus group sessions.

The research conducted by Dahan [5], presents a case study on the metaverses Framework in an e-learning environment (ELEM) context. Authors have explored the application of the Metaverse Framework as a potential solution for enhancing the e-learning experience. Case studies focus on implementing and evaluating the Metaverse Framework in an e-learning setting. The authors have described designing and developing the virtual environment using the framework. They discuss various features and functionalities incorporated into the e-learning platform, such as virtual classrooms, interactive content, and social interaction elements.

Calaguas NP [3] introduces a structural equation model (SEM) to predict adults' online learning self-efficacy. Authors have explored the factors influencing adults' belief in their ability to succeed in online learning environments and proposed a structural equation model to examine the factors that contribute to adults' online learning self-efficacy. The model includes three key components: technological self-efficacy, task value, and perceived social support.

Sohaib et al., [15] examines digital students' satisfaction with an intention to use online teaching modes, focusing on the role of the Big Five personality traits. Authors have studied how personality traits influence students' perceptions of

online teaching modes and their intention to continue using them.

The process of teaching and learning is inherently challenging. To enhance the quality of education through the integration of modern technologies, it is imperative that we adopt innovative strategies and approaches. The primary focus of this paper is to explore the potential of utilizing technological advancements in both the science and business fields to develop a more effective method for E-learning systems.

This paper proposes a new and effective method, referred to as 'learning paths,' which aims to offer learners on E-learning platforms a more systematic approach to their learning. The proposed system utilizes an algorithm called the "Natural-Inspired Migrating Bird's Learning Path" to access the progress of learners and guide them towards a more productive and efficient learning experience, catering to learners of all levels.

This paper is divided into four main sections. The first section provides an introduction to the study, while the second section outlines the methods used. The third section presents the experimental results, and the fourth section concludes with a discussion on future work.

2. Materials And Methods

A. Data Collection and Pre-processing

In every machine learning project, it is crucial to identify and pre-process relevant datasets. In the proposed work, we utilised secondary data, as creating an E-learning platform based on the proposed architecture (as shown in figure 3) would be time-consuming. Our primary challenge was finding an appropriate secondary dataset for the system. As traditional E-learning datasets do not provide extensive information on student performance and resource utilization, we opted to collect the required data from the UCI repository [19]. The dataset we utilized comprises the academic achievements and progress of students in two Portuguese schools, with data segregated based on two subjects: mathematics and Portuguese. For our proposed research, we only used the mathematics subject data, which included 395 records and approximately 32 attributes. These attributes included information such as sex, age, address, family size, internal scores, and more. We

conducted a thorough analysis of all the features in the dataset and selected a few based on our intuition for the research. The chosen features include study time, internet usage, free time, go out, weekday and weekend alcohol usage, health status, absences, and internal scores. However, we decided to exclude other features such as parent status, sex, age, school name, address, and more.

In the study, we have considered all the information provided in the dataset as features. The attribute columns such as students' addresses, school names, sex, age, and parent's status were excluded from the study as they are considered to be general information that would not add significant value to the research. However, it is important to note that the absence of evidence does not necessarily mean that these variables do not play a role in evaluating students' performance. While there may not be existing evidence to support the inclusion of these variables, it is still possible that they could impact a student's academic performance and should not be discounted. Therefore, it may be beneficial for future research to consider including these variables in the evaluation process and explore their potential impact on student performance.

As a result, the feature selection process was manually carried out based on the authors' experience and intuition. It is crucial to remove any features that are not considered for the study to prevent irrelevant correlations that could mislead the research and reduce its efficiency and accuracy. As all the features in this dataset are in numerical form, including details such as age or school name could create a high correlation between individuals of the same age or attending the same school. Hence, it is imperative to exclude such features from the study.

In order to analyze students' improvement, it is important to gather information about their hobbies and habits, especially their internet usage. While the UCI student data were directly collected from the students, the proposed E-learning platform suggests the implementation of chatbots to gather such information. Usually, students may not provide accurate details when completing surveys or questionnaires. However, implementing the chatbot system can collect more accurate results directly from the learners. Nonetheless, the available dataset contains only a few relevant features, and the system utilized those resources to their maximum potential.

The dataset we are using for this study contains three test scores: first period grade, second period grade, and final grade. All test scores are out of 20. It is important to note that the final grade has a strong correlation with the other two grades since it is based on them. Additionally, travel time is represented using numerical values, where 1 means less than 2 hours, 2 is 15 to 30 minutes, 3 means 30 to 1 hour, and 4 is greater than 1 hour. Similarly, the feature 'study time' is coded with values 1 for less than 2 hours, 2 for 2 to 5 hours, 3 for 5 to 10 hours, and 4 for greater than 10 hours. For features such as 'free time,' 'going out (go out),' 'alcohol consumption, and 'health,' a scale from 1 to 5 is used to represent low to high values. However, for absences, the values are numeric, ranging from 0 to 93.

The Exploratory Data Analysis (EDA) was conducted to investigate the correlations among all the features. The correlation heatmap and basic correlogram were used for this purpose, as shown in figure 1. The EDA process also helped identify the more important features of this research. In EDA, all the data features are analyzed and different comparisons are made.

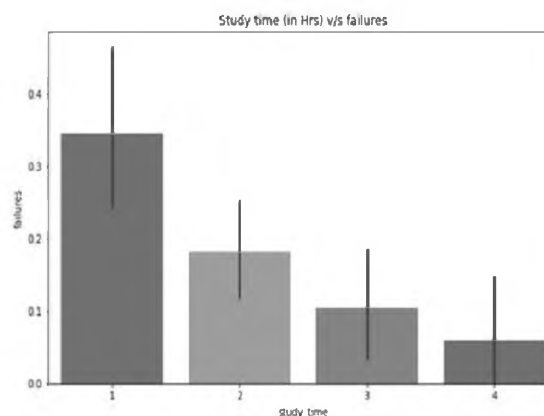


Fig. 1 : Correlation graph of study time (in Hrs.)

Figure 1 depicts the correlation between study time and failure rates, where study time is measured in hours. The figure displays the distribution of study time, ranging from 1 to 4 on the scale, and the corresponding failure rates. Our analysis indicates that there is a strong negative correlation between study time and failure rates. The highest failure rate is observed at a study time scale of 1, whereas an increase in study time results in a decrease in failure rates. On a study scale of 4, the failure rate of students is at its lowest. This suggests that study time and failure rates are inversely proportional to each other.

During the exploratory data analysis, all the feature values were carefully examined to assess the effectiveness of each feature. To support this evaluation, please refer to the graph image provided below.

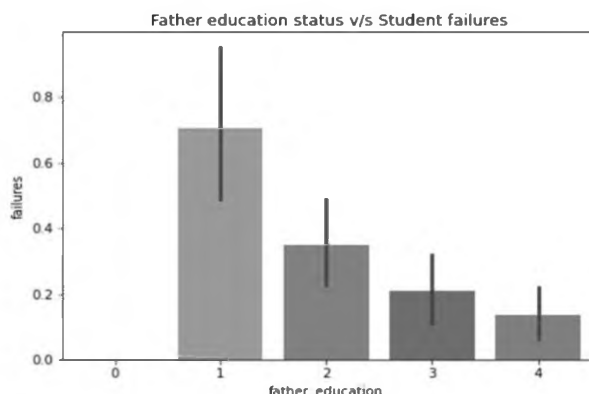


Fig. 1.1: Correlation graph of parental education and student failures

Following the analysis phase, the significant features were identified and sorted based on their correlation. In the subsequent phase, the authors manually analyzed these features, taking into consideration factors such as study time, going out and so on. However, certain features such as parental education and parents' occupation were not included in this study, despite their potentially significant influence. The focus of this study is primary on the students themselves. As an online recommender system, the authors recognized the importance of considering these features, as teachers could utilize them to provide enhanced attention to specific students. However, it should be noted that when it comes to student recommendations within the student population, these features may not yield optimal results. As a result, the authors made informed decisions regarding the selection of features for the final process.

B. Proposed Method: Natural-Inspired Migrating Bird's Learning Path

Research studies have shown that natural-inspired algorithms can provide efficient solutions for real-time problem-solving. Among the most well-known natural-inspired algorithms are Ant Colony Optimisation (ACO) [6] and Migrating Birds Optimisation (MBO) (Eduardo et al., 2015). The ACO method is inspired by real ants' foraging behavior and has been used as a solution for the Traveling Salesman Problem (TSP) [8]. The Particle Swarm Optimisation

(PSO) [7] methods are inspired by the group work behavior of bird swarms in travel patterns, and the algorithm optimizes the problem iteratively to improve candidate solutions for complex real-world problems [16] [9].

The proposed method aims to develop a new model for efficient recommendation systems, which is inspired by the behavior of migrating birds. In this model, each particle has a fitness [7] that is calculated based on various distance measures. However, our proposed method utilises an improvement calculation method instead of the conventional fitness calculation method, and it does not rely on existing distance measures. In addition, we have implemented a unique version of the recommendation system on the student dataset, which is inspired by how migrating birds choose their leader in the migration process. As shown in figure 2, the proposed method takes advantages of the unique flying patterns of migrating bird swarms to optimize the recommendation system more efficiently.

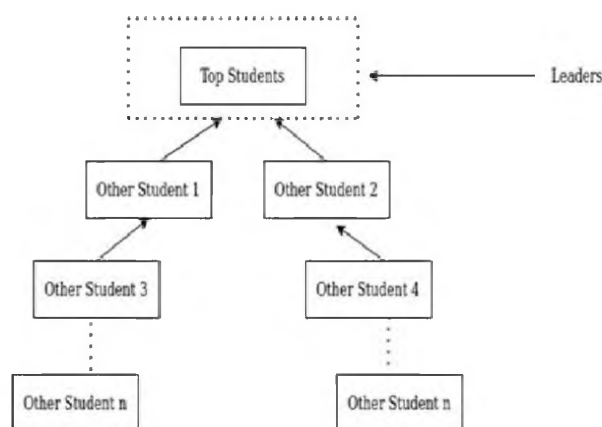


Fig. 2 : Proposed recommendation.

In the natural world, bird swarms change their leaders after a random time quantum. Similarly, in our proposed method, the student leaders (Top students), recommended resources, and the learning paths will change based on each student's improvements in each term. This ensures that the system never offers the same set of resources and paths repeatedly; instead, the system provides new resources and a new set of top performers in each term, allowing for greater diversity and effectiveness.

Proposed Algorithm: Natural-Inspired Migrating Bird's Learning Path

```

Initialize K-means
  Find cluster heads
  Initial recommendation
If (more instance of internal scores available):
  Begin: Migrating Bird's criteria:
    Initialize learners
    For each learner Calculate improvement:
      Improvement = Final internal - Previous internal
      Sort improvement by decreasing order
      Intersection of 'n' improved student features
if (more instance of internals available):
  go to: improvement calculation
repeat until course overs
  
```

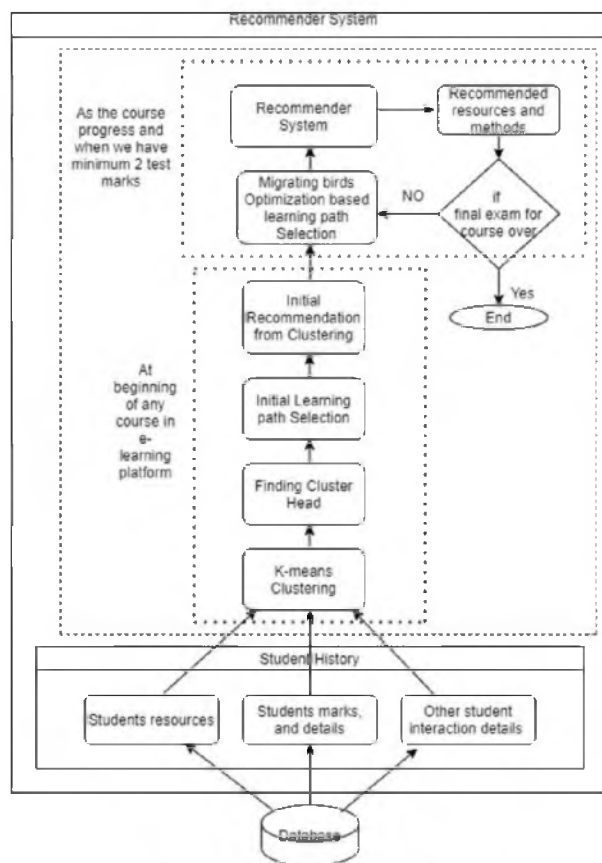


Fig 3 : Proposed system

Consider the Natural-Inspired Learning Path algorithm and proposed system architecture (Figure 3.). In the initial phase, the system may encounter a cold start problem if there are no internal scores available for the course. This means that the system may not be able to make personalized recommendations for learners until sufficient data has been collected. The proposed model consists of combination of two algorithms which are K-means

and MBO. Here, the K-means algorithm to eliminate the cold start problem [21] and provides the initial recommendations. By leveraging student's previous academic scores and other relevant parameters, the K-means method can provide more personalized recommendations. This algorithm identifies patterns in the input data and groups similar data points on their features.

K-means algorithm working scenario:

1. Initialize the K clusters.
2. Choosing initial centroids randomly.
3. Iterations repeat until no changes in centroid values:

- Compute the squared distance between data points and centroids.
- Creating clusters
- Computing centroids for clusters.

The algorithm utilizes various distance measures to determine similarity and dissimilarity. In K-means, the most commonly used distance measures are Cosine distance, Euclidean distance, and Manhattan distances [10].

We employed Euclidean distance as the distance measure in our study and selected the distance metrics based on the properties of the dataset.

$$d(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (1)$$

p, q , are two points in Euclidean n -space.

q_i, p_i is Euclidean vectors, starting from the initial point.

$n = n$ -space.

The Euclidean distance calculates the squared difference between data values, which serves as the basis for the clustering process. This approach ignores the extremes of the data, such as the smallest and

largest differences, and instead clusters the data based on the remaining values.

The K-means algorithm categorizes students into good, fair, and low clusters, based on their performance. The system offers initial recommendations based on the details of the good cluster. As the E-learning course progresses and at least two internal scores are available, the system enters its second recommendation phase, where it evaluates the student's performance based on different feature instances. The system computes the performance improvement using the formula provided below.

$$\text{Performance} = \frac{\Delta \text{feature}}{\text{feature}(\text{max})} \quad (2)$$

In this equation, the term 'feature' refers to the various attributes of the student data, and 'max' corresponds to the highest value among these attributes.

To compute the 'Performance', we require a minimum of two internal scores in online learning. For instance, if a student's internal scores are 14 and 15 out of 20, we can apply the above formula to assess their performance.

Feature value initial (feature 1) = 14

Feature value final (feature 1) = 15

$$= \frac{15 - 14}{20} = 0.05$$

If the dataset comprises only two internal scores or other features, the system recommends the best option based on the learners' performance improvement. When we have more than two internal scores for a particular course, we can consider that there is 3 internal scores test1, test2, and test3. This will provide two feature sets, A and B (based on performance formula), which consist of the test1, test2 (test2-test1), and test2, test3 (test3-test2) scores, respectively, for the first and second iterations. The system identifies the intersection between the A and B feature sets and provides recommendations based on this shared information. These are the fundamental principles behind the proposed recommendation system, which prioritizes students' improvements over their maximum scores in each term.

Table 1: K-means recommendations

Student id	Study time(Hrs)	Abse nces	Go out	Period1 Score(20)	Period2 Score(20)
47	4	4	2	19	19
42	2	2	3	19	18
374	3	0	4	19	18
293	4	6	2	18	18

High-performing learners in each iteration often share common characteristics, such as study time, internet use, and other factors. By identifying and suggesting these features to other students in the same classroom, we can help weaker students to improve. Merely increasing study time may not always lead to better performance; other factors, such as the quality and relevance of study materials, can also play a crucial role. Providing students with more comfortable and standardized resources can facilitate their learning experience. Therefore, recommending such resources to other students in the classroom can be highly beneficial.

3. Experimental Results

The proposed system is implemented using Python and various libraries, such as scikit-learn and matplotlib. Initially, the K-means algorithm is utilized to find similarities between neighboring learners and form clusters. The system clusters the dataset based on the several features, such as period scores, internet usage, and study time, before providing recommendations based on the cluster. The algorithm identifies the most similar students by analysing the similarities and dissimilarities in the data. For instance, in Table 1 and figure 4 provided below, the K-means method yields the most similar categories.

	period1_score	period2_score	period1_status	Cluster
47	19	19	good	2
42	19	18	good	2
374	19	18	good	2
245	18	18	good	2
293	18	18	good	2
..
161	5	9	poor	8
164	5	8	poor	8

Fig. 4: K-means results

The results of the k-means recommendation are presented in Table 1, corresponding to Figure 4. Based on these findings, it is observed that the study times, going out frequency, and other factors of the four students are not significantly similar, except for their student scores. Utilizing these similarities, the algorithm successfully identified similar students. The Euclidean distance metric was employed in the

algorithm to calculate both similarities and dissimilarities, leading to the discovery of more students with similar characteristics in the recommendations. The feature values derived from these results, such as study time, absence rate, and going out frequency, can be suggested to other students as an average to follow.

The system first performs an initial clustering process to identify the top-performing students and their resources, which are then suggested to other students. Subsequently, the system checks for additional internal scores, such as internal 2,3 and so on, and enters the main phase of the algorithm. In this phase, the system evaluates the student's progress in learning by applying the performance formula (equation.2). Based on this, the system makes recommendations that take into account the learners' improvements, rather than just their maximum marks.

Table 2: Recommendations from demo dataset

Student id	Period 1 score	Period 2 score
100	11	11
668	13	10

The proposed system also conducted experiments on dummy data that fulfills the method's requirements. The experiment results are presented in Table II, where the system primarily emphasizes student resources. The recommended resources and learning habits are determined based on student's performance improvements. Table III displays the cluster of weak students, as the system also aims to identify weak students and offer better guidance. Therefore, if the system were to be deployed in real classroom, the recommendation would be for students to follow the resources and learning paths of the best-

Table 3: Weak students cluster

Student id	Books referred	Sites accessed	Free time	Perid1 Score(20)
449	Expert C Programming: Deep C series by Peter	C and C++ programming resources	2	18
444	Advanced programming in Unix environment	C and C++ in resources	4	18
213	Advanced programming in Unix environment	C and C++ in resources	5	18

performing students, based on their attribute values such as study time, go-out rate, etc. This would serve as a resource recommendation and learning path for weaker students to improve their performance. However, currently the best students attributes values are considered as recommendation.

A. Model Evaluation

The proposed system is evaluated using a qualitative technique because it is unique and distinct from existing concepts. The system recommends learning paths, resources based on performance improvement, so all recommendations are in increasing order of performance values, making them better. Therefore, qualitative methods are appropriate for evaluating the system's outputs by comparing them with other methods. However, the final evaluation can only be conducted once the recommended learning paths are implemented in a real-world setting.

B. Comparison of K-means and Natural Inspired Method

When examining the recommendations suggested by Natural Inspired learning systems and K-means in Tables IV & V, it becomes apparent that there are discrepancies in the feature values. For example, when looking at period scores, K-means clustering results in more similar scores and therefore, clusters are created based on these internal scores. In contrast, other feature values such as study time, absences, and going out are only slightly similar to the other recommended values. On the other hand, the Natural Inspired system recommendations have distinct internal values, with different valued scores recommended. Moreover, upon evaluating the recommended students based on their improvements, it can be inferred that if implemented in the real world, this approach would recommend the most exceptional students.

Table 4: Proposed system recommendations (MBO)

Student id	Study time (Hrs)	Absences	Go out	Period 1 Score (20)	Period 2 Score (20)
47	4	4	2	19	19
42	2	2	3	19	18
374	3	0	4	19	18
293	4	6	2	18	18

Additionally, other feature values such as study time and absences are more similar to other students. Consequently, the Natural Inspired learning system provides more effective recommendations based on performance improvement.

Upon comparing the two recommendation tables, it is evident that the Natural Inspired system generates more comparable recommendations. The system recommends students who exhibit similar performance metrics, with the primary feature being the students' marks. It is noteworthy that this feature can be replaced with any other feature, but for better comprehension, we have utilized internal marks. The most significant advantage of the proposed method is that it yields recommendations that are more similar to one another. In contrast, when examining the K-means results in Table IV, the recommendations are not as similar as those generated by the Natural Inspired system.

C. A Common Issue in Evaluation

In this research, the proposed system architecture employs a distinct approach. Initially, the system identifies a suitable secondary dataset and then applies the proposed algorithm. The system employs a machine learning technique that determines the correlation between the research elements or problems. Based on these correlation values, the system generates the output. By utilizing secondary data, the system can proceed with the subsequent phases of the research based on the accuracy of the output.

Consequently, it is imperative to conduct the second phase of the research with a real-time dataset. The results obtained from this phase will enable us to draw a final conclusion regarding the overall performance of the model. This will help determine the model's applicability for real-time or traditional teaching-learning scenarios. The proposed method is entirely distinct from existing recommendation systems, rendering other methods unsuitable for comparison. The system's performance is based on the

improvement of the students, and therefore, the recommended results can only be compared with other methods. In the context of evaluating E-learning or traditional learning, instructors or course providers often consider only the highest scores in examinations. However, this method of evaluation does not accurately reflect the learning process since it only focuses on marks, while learning is a holistic process that can not be evaluated based solely on scores. The proposed method allows for the evaluation of different feature sets, not just internal scores, as the objective of education is not solely to acquire more marks but to improve students in all areas. Unfortunately, we currently rely on a marks-based evaluation system. In the proposed work, we analyze the students' marks, but not in the same manner as existing methods. The method prioritizes improvements in learning, so the learners are chosen based on their learning improvements rather than solely on the maximum marks method.

Providing recommendations for resources and learning habits on E-learning platforms is significantly more challenging than other existing recommender systems like movie recommendations [4] and product recommendations [1]. The primary reason for this is that those recommendation systems are based on the conventional approach of finding similar users and habits using distance measures. In contrast, the implemented system focuses on student improvements and recommends resources and lifestyles accordingly. The key difference between the proposed system and existing method is that the recommendations provided by the proposed system will differ in each phase. Only the improved students and their patterns are recommended, which sets it apart from other methods.

4. Conclusion and Future Works

The Natural Inspired learning system introduced a novel approach to E-learning recommendations using machine learning techniques. The system first clusters student groups with the K-means algorithm and provides initial recommendations based on the cluster results. The Natural Inspired learning method only considers performance improvement and aims to address the cold start problem by incorporating K-means recommendations. The algorithm provides equal priority to all learners on the platform, using a democratic approach. The system is designed to promote better and more efficient learning activities among students.

Table 5: Initial K-means System Recommendations

Student id	Study time (Hrs)	Absences	Go out	Period 1 Score (20)	Period 2 Score (20)
254	4	5	4	8	12
161	4	3	4	5	9
4	5	5	3	6	10

In this study, a recommendation system utilizing the MBO method was developed, which provides recommendations based on student improvement. As future work, it is necessary to implement this method on a real E-learning platform to evaluate its performance. Additionally, the algorithm can be applied to different E-commerce platforms to offer improved recommendations for products and services. Since many E-commerce platforms include ratings for their products and services, this algorithm can compare these ratings and suggest new products to different customers based on their previous purchasing history. This method can also be utilized in movie recommendations by calculating the improvements and accomplishments of directors and actors based on their previous work.

References

- [1] Aggrwal, C.C., (2016). Recommender Systems: The Textbook, 1st ed. 2016. ed. Springer International Publishing : Imprint: Springer, Cham. <https://doi.org/10.1007/978-3-319-29659-3>
- [2] B.F. Skinner (2014). Science and Human Behavior, Pearson
- [3] Calaguas NP, Consunji PMP(2022). A structural equation model predicting adults' online learning self-efficacy. Educ Inf Technol (Dordr). 2022;27(5):6233-6249. doi: 10.1007/s10639-021-10871-y. Epub 2022 Jan 5. PMID: 35002467; PMCID: PMC8727476.
- [4] Costing-Gabriel Chiru, Vladimir-Nicolae Dinu, Catalina Perda, Matci Macri (2015). Recommender System Using the User's Psychological Profile. IEEE International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, pp. 93-99 <https://doi.org/10.1109/ICCP.2015.73.12611>
- [5] Dahan NA, Al-Razgan M, Al-Laith A, Alsoufi MA, Al-Asaly MS, Alfakih T (2022). Metaverse Framework: A Case Study on E-Learning Environment (ELEM). Electronics. 2022; 11 (1 0) : 1 6 1 6 . <https://doi.org/10.3390/electronics11101616>.
- [6] Dorigo, Macro Birattari, Mauro Stutzle, Thomas. (2006). Ant Colony Optimization, Computational Intelligence Magazine, IEEE.1, 28-39. <https://doi.org/10.1109/MCI.2006.329691>
- [7] Ekrem Duman, Mitat Uysal, Ali Fuat Alkaya (2012), Migrating Birds Optimization: A New Metaheuristic Approach and its performance on quadratic assignment Problem, Information Sciences, 217, pp. 65 - 77 . <https://doi.org/10.1016/j.ins.2012.06.032>
- [8] Eldem, H., Ülker, E., (2017). The application of ant colony optimization in the solution of 3D traveling salesman problem on a sphere, Engineering Science and Technology, an International Journal, 20(4), 1242-1248, <https://doi.org/10.1016/j.jestch.2017.08.005>
- [9] Givi, H., Hosseini, E., Farjah, E., (2021). Estimation of Batteries Voltages and Resistances in Modular Multilevel Converter With Half-Bridge Modules Using Modified PSO Algorithm, 2021 12th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC), Tabriz, Iran, 2021, pp. 1-7, <https://doi.org/10.1109/PEDSTC52094.2021.9405967>
- [10] Giudici, P., (2003). Applied Data Mining, John Wiley & Sons Ltd, ISBN 0-470-84679-8.
- [11] Lalla-Ruiz, E., Exposito-Izquierdo, C., de Armas, J., Belen Melian-Batista, B., Moreno-Vega, J.M., (2015). Migrating Birds Optimization for the Seaside Problems at Maritime Container Terminals, Journal of Applied Mathematics 2015, pp. 1-12. <https://doi.org/10.1155/2015/781907>
- [12] Lawrence Miller (2017). Internet of Things for Dummies, John Wiley & Sons, Inc., ISBN 978-1-119-34989-1.
- [13] Marini, F., Walczak, B., (2015), Particle swarm optimization (PSO). A tutorial, Chemometrics and Intelligent Laboratory Systems, 149, 153-165, <https://doi.org/10.1016/j.chemolab.2015.08.020>
- [14] Marc Rosenberg, Kevin Moore, Feank Hanfland, Patti Shank, Lisa Young, Lance Dublin, Ryan Watkins, Michael Corry (2007). The eLearning Guild's Handbook of e-learning Strategy, The eLearning Guild.

- [15] Mustafa Sohaib, Qiao Yu, Yan Xin, Anwar Aliya, Hao Tengyue, Rana Sehrish (2022), Digital Students' Satisfaction With and Intention to Use Online Teaching Modes, Role of Big Five Personality Traits , Frontiers in Psychology , Vol:13 , <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.956281>, 10.3389/fpsyg.2022.956281 ISSN: 1664-1078
- [16] Rezaee Jordehi, A., Jasni, J., Abdul Wahab, N.I., Abd Kadir, M.Z.A.,(2013). Particle swarm optimisation applications in FACTS optimisation problem. 2013 IEEE 7th International Power Engineering and Optimization Conference (PEOCO), Langkawi, Malaysia, pp. 193 - 198 . doi : 10.1109/PEOCO.2013.6564541
- [17] Setiawan, Wahyu. (2018). Differences of Education Systems in Developed and Developing Countries Curriculum, Educators and Financing in Indonesia and Finland. Didaktika Religia, 6. 139-152.
- [18] Sofiadin, Aidrina. (2022). Students' Perspective on Sustaining Education and Promoting Humanising Education through e-Learning A Qualitative Focus Group Study. International Journal of Advanced Computer Science and Applications . 13 . 2022 . 10.14569/IJACSA.2022.0131168
- [19] UCI data repository (2020), <https://archive.ics.uci.edu/ml/datasets/Student+Performance>.
- [20] Vishnu Pendyala (2018). Veracity of Big Data, a press publisher, ISBN-13(pbk):978-1-4842-3632-1.
- [21] Watt, J., Borhani, R., Katsaggelos, A.K., (2016). Machine Learning Refined: Foundations, Algorithms, and Applications, Cambridge University Press, ISBN 978-1-107-12352-6.