[1] Prakash Srivastava

[2] Mohammad Mahyoob Albuhairy

[3] Jeehan Algaraady

[4] Samridhi Srivastava

[5] Mohammad Zubair Khan*

[5] Abdulaziz Alblwi

Selection of Education Modes during Pandemics using MCDM and Machine Learning Techniques

Journal of Electrical Systems

Abstract: **-** Online education is rapidly evolving worldwide and is increasingly gaining favour over the traditional method of learning and teaching. This paradigm shifts in learning and managing using the Internet, or e-learning, is directly linked to the evolution of digital technology. E-learning systems depend on different factors for their success, which need to be thoroughly analysed to help us decide on the mode of education; therefore, proper management is required. In the case of a pandemic situation, for example, the condition of COVID–19, where physical interaction has stopped completely, the mode of education needs to be changed to continue education unhindered. In such unprecedented circumstances, a quick decision must be taken. This study aims to find out the optimum mode of education as per the inputs received from students and teachers by suggesting the Technique for Order of Preference by Similarity to the Ideal Solution (TOPSIS) approach based on Multi-Criteria Decision Making (MCDM) integrated with Machine Learning (ML) algorithm, K-Nearest Neighbors (KNN). The results demonstrate the method's effectiveness in managing the modes more accurately

Keywords: COVID-19, K-Nearest Neighbors (KNN), Multi-Criteria Decision Making, Online Education Management, Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

I. INTRODUCTION

In this digital era, the traditional education format requires a significant change. While some have already adopted it, others are perplexed about whether to adopt the changes. The conventional education format of face-to-face education has served us far too long, but now it is time for this mode to change. There is no doubt that the traditional mode is a complete package of knowledge as there is continuous interaction between students, allowing them to gain knowledge from teachers and exchange information. However, online education has still been accepted as a significant education component in this era. Online Education is expanding rapidly, but its existence is somehow scarce in engineering courses as students prefer face-to-face education for difficult and essential courses that require lab work and practical knowledge [1]. For shifting from face-to-face to online education, a proper, robust infrastructure is needed, which could be tested in all types of education scenarios and provide us with exceptional outcomes [2]. Portal education is increasing daily as technology enhances with new variants of embedded systems, microprocessors, and other data acquisition tools. Nevertheless, for this change to occur, many resources are required, and various challenges like website availability, scalability, the need to be fault-tolerant, and other needs are required for the smooth execution of online teaching. Online practical assessment for various

[1] Department of Computer Science & Engineering, Graphic Era, (Deemed to be University), Dehradun, India prakash2418@gmail.com

[2] Language and Translation Dept, Taibah University, Saudi Arabia Mqassem@taibahu.edu.sa

[3] Languages and Translation Dept, Taiz University, Taiz, Yemen Jihan.amu@gmail.com

[4] Mahindra Bristlecone, Noida, India Samridhi0507@gmail.com

[5]Department of Computer Science and Information, Taibah University, Medina 42353, Saudi Arabia *zubair.762001@gmail.com

Ablwi@taibahu.edu.sa

courses is possible using microprocessors and embedded systems, but the courses for which it is not possible should also be considered. Hence, the blended mode of education can be adapted to deal with these scenarios.

One of the most important aspects to be considered while choosing a mode of education is student awareness, as students lack interaction in online and blended modes. Student evaluation is also a significant issue for faculty. The evaluation process should be as rigorous as possible, requiring various authenticity and integrity implementations to prevent cheating and identify possible fraud during the session . Hence, the implementation of online education requires a proper risk assessment by the university to identify all sorts of possible threats to the proposed multimedia platform being used. Most students prefer to opt for blended mode, which allows both objectives to be fulfilled as maximum education can be provided online. Still, some courses are efficient only when given offline. The whole world is suffering from a pandemic, causing tremendous effects on the economy of various countries. Further, the youth's education has been affected too. This pandemic has left the world in such a state that people have no other way left than living in isolation, fighting against the virus, and trying their best to save themselves by following various government-suggested safety norms. The government had to shut down places with heavy crowds, leading to the shutdown of schools and colleges [3]. This shutdown is having a minacious effect on students' education, and the institutes are left with a choice of either staying closed or taking steps towards the new era and moving to online education. Many institutions have started education online, but many are still deciding whether to take this step. These institutions have many questions regarding the availability of resources, the choice of platform to adopt, the number of students attending these sessions, how to connect teachers with students, whether to opt for fully online or blended mode, and many other such questions. We use the Technique for Order Preference by Similarity to the Ideal Solution (TOPSIS) method to test various criteria available and decide whether to move online or blended or wait for the situation to recover and continue with traditional face-to-face methodology only.

The TOPSIS method helps to find the ideal solution. We can appropriately rank the closest ideal alternative and the most negative alternative. The TOPSIS method can be used for multi-criteria decision-making based on other options like Online, Blended, and Offline Education Mode, which can be ranked appropriately by evaluating numerous criteria [4].

Contributions

- Designing a framework to identify the mode of education using TOPSIS.
- Designing a framework to identify the mode of education using machine learning techniques.
- Comparison of outcomes based on TOPSIS and machine learning techniques.

The paper outline is as follows: Section 2 contains a brief explanation of related works in this field; Section 3 contains the criteria analysis, where we define the criteria used and alternatives considered with the used method, along with a practical application shown as an example; Section 4 presents the result analysis and suggestions for future work; and Section 5 contains the conclusions of our work.

II. Related Work

[1] stated that online education grows with time and directly depends on various parameters. The paper has described four critical methodologies on which online education depends and is essential for education in this form. These are as follows: first is the diverse population attending the session. Educators, researchers, etc., can use descriptors to measure the population relying on online rather than offline education. Second, it provides a new conceptual model that integrates descriptors with identifiers for better accuracy in the population examination based on a content-based literature model.

[4] proposed a view on how the new era of microprocessors, embedded systems, etc., has made the practical assessment of education in online mode better, further allowing students to design, program, and test circuitry in the online platform rather than waiting for offline sessions, so it increases the access and sharing of information in a massive way. This invention has helped students learn at their own pace and become more acquainted with the various toolsets in a more interactive way with full hands-on assistance. According to the latest survey, the strength of two-year retention of students has increased excellently.

[5] proposed that online education has expanded to master's programs, but a maximum of engineering degrees is not followed. However, maximum online education lies in the postgraduate and master's levels. This shows how the master's courses have been transformed from face-to-face mode to online mode or blended mode, depending upon the necessity of the course. This online education is integrated with various measures to keep all the students active throughout the session in their institution. After every five hours of a course, a test is conducted to familiarize the students with the ongoing session. This paper mainly shows how the learning and organizational changes are occurring with the implementation of this model.

The multi-criteria decision-making (MCDM) process is used in various fields to identify the best alternative based on the criteria. It uses matrix formation, normalization, and ranking to compare the best choices based on the criteria taken. [6] showed that we are surrounded by numerous intelligent devices that integrate the Internet of Things (IoT) to work for our lives efficiently. However, in the background, these devices operate with Wireless Sensors to disseminate the data to the base station for storage necessary for working these devices properly.

[7] proposed that Delay-Tolerant Networks (DTNs) are network devices that address technical communication issues due to a lack of network connectivity. DTNs are used in mobile communication and mostly in space communication, which is an essential need nowadays, as communication needs to be as efficient as possible.

Multi-Criteria Decision-Making deals with computational and mathematical tools developed to choose the best alternative based on various criteria. It provides a suite of multiple methods available for decision-making, for example, the Analytic Hierarchy Process (AHP), Weighted Sum Model (WSM), Simple Additive Weighting (SAW), Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), etc. However, all these methods are calculated manually, causing a burden for many complex scenarios. [8] proposed using the TOPSIS methodology, which is widely used, as it is considered one of the best and easiest methods to rule out the best alternative and rank the other options based on the calculations.

Machine Learning has been adopted to resolve various problems the nation faces due to COVID–19. Artificial Intelligence (AI) has helped us a lot, and [9] proposed that it works in replace humans with machines or robots which does not get infected with diseases and can do most of the critical tasks required. Artificial Intelligence can help for early detection as it can be used to compare specific common symptoms, which can help to test only the persons that have been tested positive by AI technology and can save many resources and further if the person has been tested for infected by AI can help to keep the person in quarantine such that the disease is not spread further.

[10] proposed that healthcare sectors have a vast amount of data because they are a vibrant sector worldwide. Data compliance is also the circumstance that led to Big Data Analytics for effectively processing these data. In this paper, many machine learning algorithms were applied, and Logistic Reasoning was the most accurate, with an accuracy of 96%. Using pipelines, the AdaBoost classifier came out to be 98.8% as the best model.

[11] aimed to implement a model that determines whether a student deserves a scholarship based on specific criteria: semester, family income, and many family members, amongst others. The dataset used contains 24 students with scholarships out of 2028 students considered. They used Euclidean distance to calculate the distances of data points and a confusion matrix to calculate the various parameters of the model. Results show that an accuracy of 95.83% can be reached when different parameters and values are considered.

III. METHODOLOGY

A. Criteria Analysis

Online Education is vital today as it can change millions of dreams worldwide. Blended mode integrates online and offline modes, which is required because certain things cannot be conducted online. While some concepts can be taught better offline, online mode allows students to continue their education through different means depending upon the prevalent scenario [5]

However, it is essential to consider both faculty and students' difficulties with the implementation of online education, and only after considering their suggestions should changes in education mode be considered. There are different situations available that can be chosen as criteria. Eleven of the most important criteria are used in this paper for decision-making. These criteria include Net Connectivity, Multimedia availability, Level of Stress, Session Interactiveness, Health Issues, Online Practical Session, Assignment completion and evaluation, Mode of

Examination, Overall preference, Level of Knowledge achieved and Session attendance. Internet connectivity is a critical need in online education, making this parameter important in choosing an appropriate alternative, as the unavailability of the Internet hinders online education. Here, offline education is weightier than online education. This parameter is considered from both faculty and students' point of view. Multimedia availability is a must to attend online sessions as without appropriate devices and software, it is difficult to attend online sessions. Here, offline sessions will be beneficial. Stress is an essential factor when considering the mode of education, and it entirely depends on the students and faculties as to which mode they consider less or more stressful. Further, this paper uses TOPSIS and the Machine Learning-based algorithm K-Nearest Neighbors (KNN) to identify whether students prefer online, offline, or blended modes for their education. The dataset was collected from the students and faculty using a questionnaire with 11 questions, with which we performed the analysis mentioned in this paper.

B Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)

TOPSIS is one of the most widely used multi-criteria decision-making techniques for choosing the most appropriate alternative for any particular study based on enormous criteria. However, only numeric values of alternatives based on appropriate criteria are considered in this method. TOPSIS ranks the alternatives based on their closeness to the ideal solution. In this model, the distance of the alternative form, both the ideal and anti-ideal solution, is considered [12]. This model works in a total of 6 steps to give us the alternatives ranked appropriately.
These steps are:

- Formation of Matrix:

A matrix is formed of order m X n, where m is the set of alternatives and n is the criteria considered (Zavadskas et al., 2016).

$$\begin{bmatrix} x_{11} & x_{12} & x_{23} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{bmatrix}$$

X_{ij} is the numeric value for i^{th} alternative based on j^{th} criteria where {i=1,2,3….m},{j=1,2,3…n}

- Normalization of Matrix:

As these values are based on various criteria, so normalization to an appropriate scale is necessary. In this paper, vector normalization is used where the normalized value is the ratio of the original value and the square root of the sum of the squares of all its alternatives. The formula is:

$R_{ij}= \frac{X_{ij}}{\sqrt{\sum_{i=1}^{m} X_{ij}^2}}$ where i is the i^{th} alternative and j is the j^{th} criteria.

Determination of ideal (A^+) Solution and anti-ideal (A^-) Solution:

$$A^+ = \{(\max_i r_{ij} \mid j \in J), (\min_i r_{ij} \mid j \in J') \mid i = 1,2,3..m\} = \{A_1^+, A_2^+, A_3^+, \ldots\ldots\ldots A_K^+\}$$

$$A^- = \{(\min_i r_{ij} \mid j \in J), (\max_i r_{ij} \mid j \in J') \mid i = 1,2,3..m\} = \{A_1^-, A_2^-, A_3^-, \ldots\ldots\ldots A_K^-\}$$

$$J = \{j = 1,2,3\ldots k \mid k \text{ belongs to benefit criteria}\}$$

$$J' = \{j = 1,2,3\ldots k \mid k \text{ belongs to cost criteria}\}$$

Both benefit and cost criteria correspond to high performance but differ as the benefit criteria have a high identical value, whereas cost criteria correspond to a small identical value [12]. It corresponds as A^+ corresponds to the best alternative, whereas A^- corresponds to the worst alternative. Calculation of separation measure, which is based on Euclidean distance, based on the positive ideal and negative ideal solution separately:

$$S_i^+ = \sqrt{\sum_{j=1}^{n}(r_{ij} - A_j^+)^2}$$

$$S_i^- = \sqrt{\sum_{j=1}^{n}(r_{ij} - A_j^-)^2}$$

Where i = 1,2,3,….m

Calculation of relative closeness of alternatives to the appropriate ideal solution: This is done to bring all the alternatives to a single base for comparing purposes with other alternatives.
$C_i^* = \frac{S_i^-}{S_i^+ + S_i^-}$. Then using C*, all the alternatives are correctly ranked to give us the correct order to choose the best alternative for our purpose. This ranking is done based on the decreasing order of C*.

C. *K-Nearest Neighbors Classifier*

KNN is one of the most widely used algorithms for its simplicity and non-parametric calculations. This model works for both regression and classification. The input to the KNN model is a data point, and the output is a class based on the proximity calculations of all other data points in the feature space. The feature space is an n-dimensional space where n is the number of attributes in the dataset. For a given input data point d, we calculate the distance from d to every other data point. The consensus of the classes of K nearest neighbors of d determines the class to be assigned [13]. There are different methods to calculate the distance between the points, all having their advantages [14]. In this paper, we have used the Euclidean formula to calculate the distance between the points.

Let $\{(x_1, y_1), \ldots, (x_N, y_N)\}$ be a set of observations in a q-dimensional space, having $X=\{x\}^N_{i=1}$ and $Y=\{y\}^N_{i=1}$. For an unknown pattern x', we employ the Minkowski metric,

$$||x' - x_j||^p = (\sum_{i=1}^{q} |(x_i)' - (x_i)_j|^p)^{1/p} \quad (1)$$

where p=2 corresponds to Euclidean distance.

We can also apply KNN for multi-class classification. For predicting an unknown pattern x' in multi- class classification, we use,

$$f_{KNN}(x') = \arg\max_{y \in Y} \sum_{i \in N_K(x')} I(y_i = y) \quad (2)$$

Where function $I(.)$ gives a true value when the label y_i of pattern x_i is y and false otherwise.

The KNN algorithm poses some limitations [15], one of which is that the model results depend on the value of K chosen. When the value of K is small, we get an overfitted model, whereas when the value is significant, we get an underfitted model. This model is a lazy learner, i.e., it runs slowly and has high computation complexity.

IV. Results and Discussions

TOPSIS model has been chosen for appropriate decision-making in this paper to opt out of the best alternative to adopting online education. In this paper, suggestions from various faculties and students have been taken using a survey. Using the survey, various suggestions were opted based on 11 criteria for ex.: Internet Connectivity, Multimedia availability, Level of Stress, Session Interactiveness, Health Issues, Online Practical Session, Assignment completion and evaluation, Mode of Examination, Overall preference, Level of Knowledge achieved and Session attendance as (A1, A2, A3….A11 respectively). The students' perspective suggests the level of knowledge achieved, and session attendance is the faculty's perspective; their weights are considered as 0.05 each, whereas all other criteria are evaluated with the same weight [12]. Further, the total weight of all alternatives must be 1. So, the weight of the other nine criteria is determined to be 0.1. So, the values from the survey are taken in numeric form, and a decision matrix is formed with all the appropriate values. The decision matrix is 3 x 11 in order, as shown in Table 1.

Table 1: Data Decision Matrix

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
Offline Mode	11	25	36	85	40	44	48	52	49	102	80
Blended Mode	102	90	69	35	60	40	35	35	20	28	23
Online Mode	**72**	**70**	**80**	**65**	**85**	**101**	**102**	**98**	**116**	**55**	**82**

After the formation of the original decision matrix, then the matrix is normalized using step 2. The normalized matrix is shown in Table 2. Then the matrix values are integrated with their weights according to the criteria weightage matrix and Table 3 represents the weighted normalization matrix. The criteria weights for the matrix are as follows: {0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.05, 0.05} for {A1, A2, A3……., A11 respectively.

Table 2: Normalised Data Matrix

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
Offline Mode	0.087	0.214	0.322	0.712	0.346	0.362	0.406	0.447	0.384	0.855	0.692
Blended Mode	0.814	0.771	0.618	0.332	0.520	0.329	0.296	0.300	0.156	0.234	0.199
Online Mode	0.574	0.599	0.716	0.617	0.780	0.872	0.864	0.842	0.909	0.461	0.710

Table 3: Weighted Normalized Matrix

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
Offline Mode	0.0087	0.0214	0.0322	0.0712	0.0346	0.0362	0.0406	0.0447	0.0384	0.0427	0.0346
Blended Mode	0.0814	0.0771	0.0618	0.0332	0.0520	0.0329	0.0296	0.0300	0.0156	0.0117	0.0099
Online Mode	0.0574	0.0599	0.0716	0.0617	0.0780	0.0872	0.0864	0.0842	0.0909	0.0230	0.0355

With the help of the Weighted Normalized Matrix, the positive ideal and negative ideal solutions are chosen according to step 3. The solution for the respective criteria is:

A^+ = {0.0814; 0.0771; 0.0716; 0.0712; 0.0780; 0.0872; 0.0864; 0.0842; 0.0909; 0.0427; 0.3550}

A^- = {0.0087; 0.0214; 0.0322; 0.0332; 0.0346; 0.0329; 0.0296; 0.0300; 0.0156; 0.0117; 0.0099}

Then using this A^+ and A^- matrix, the separation of each alternative based on the criteria is calculated from the positive ideal solution as S^+ and negative ideal solution as S^- as shown in table4.

Table 1: Separation Measures of Existing Alternatives

	S^+	S^-
Offline Mode	0.02084	0.00388
Blended Mode	0.01861	0.00957
Online Mode	0.00135	0.02367

Then using these S^+ and S^- values, the relatives' closeness of each alternative is calculated using step 5. Finally, the corresponding values of relative closeness are stated in Table 5.

Table 2: Modes ranked according to their relative closeness index

	CC*	Rank
Offline Mode	0.15706	**3**
Blended Mode	0.33947	2
Online Mode	0.94603	**1**

The dataset consists of 185 records collected from both teachers and students. The 11 columns, as mentioned before, have been used as input to the model. In addition, a 12th column was created where we added the labels to the already collected data. These labels were added considering an individual's choices in the survey. These labels were used to train and test the model [14]. To obtain effective results, it is imperative to determine the optimal value of K in the KNN model. There are different methods to determine the value of K in this model. We have used the most common method, called the Elbow method, in supervised machine learning.

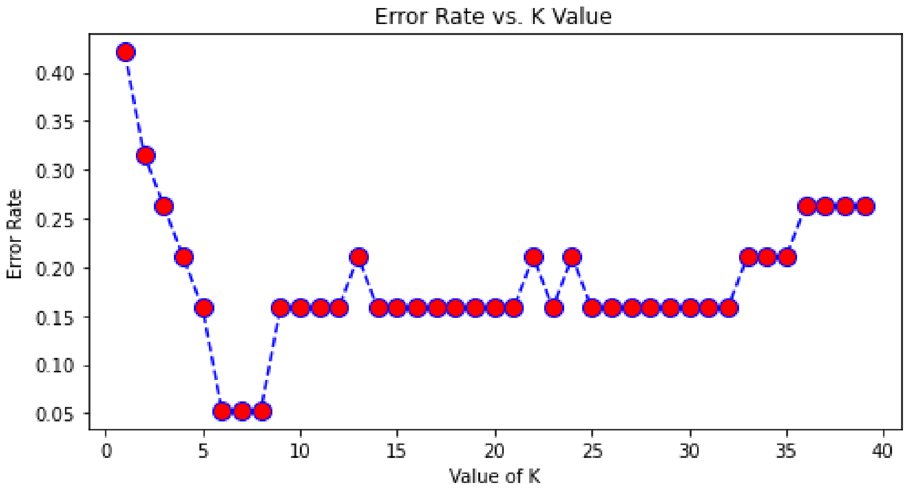

Fig.1 Plotting the Error Rate with the Value of k

The graph Fig.1 shows that the error rate is minimum when the value of K is between 6-8. Therefore, the value of K as 8 was used to train the KNN model.

KNN model performance metrics

To determine the model's performance, we generally consider its precision, recall, accuracy, and f1-score. In our paper, we find the precision, recall, and f1-score of all three classes, viz., Online Mode, Blended Mode, and Offline Mode, and the accuracy score of the model.

Accuracy refers to the ratio of accurately predicted samples to total samples in the dataset.

$$\text{Accuracy } = \frac{\left((TP + TN)\right)}{(TP + TN + FP + FN)}$$

Precision is the ratio of all the correctly predicted positives to the positively predicted samples by the classifier [8].

$$Percision = \frac{TP}{(TP + FP.)}$$

Similarly, recall refers to the ratio of correctly predicted positives to the total actual number of positive samples in the dataset.

$$Recall = \frac{TP}{(TP + FN)}$$

Finally, the f1-score is taken as the Harmonic mean of precision and recall.

$$F1 - score = \frac{(2 * \text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})}$$

These parameters have been calculated using the Confusion Matrix[16]. The KNN model is beneficial when the dataset size is small, as it can still generate high [17].The values of these parameters are given in the Table 6.

Table 6: Performance Measures of the KNN Model

Metrics	Online Education Mode	Blended Education Mode	Offline Education Mode
Accuracy	94.73%	95%	93%
Precision	1.00	1.00	0.93
Recall	1.00	0.75	1.00
F1-Score	1.00	0.86	0.97

The accuracy obtained from the KNN Model is 94.73%. The reason for this model's great performance is that KNN can work with smaller datasets and still have high accuracy. Visualization of the performance measures has been shown as follows Fig.2:

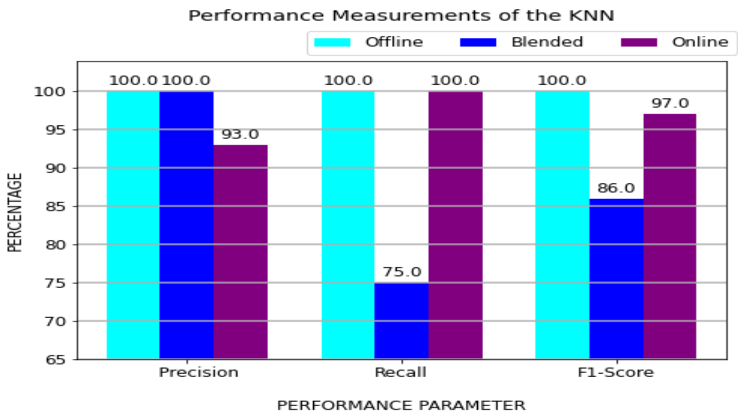

Fig.2 Performance Measurements of KNN

In Figure 3, the confusion matrix shows that the KNN model outperforms in predicting online and Offline education modes, achieving perfect recall and high precision of 100% and 93%, respectively. However, it struggles with the Blended mode, where 25% of the instances are incorrectly classified as Offline, while 0.75% are correctly classified as Blended.

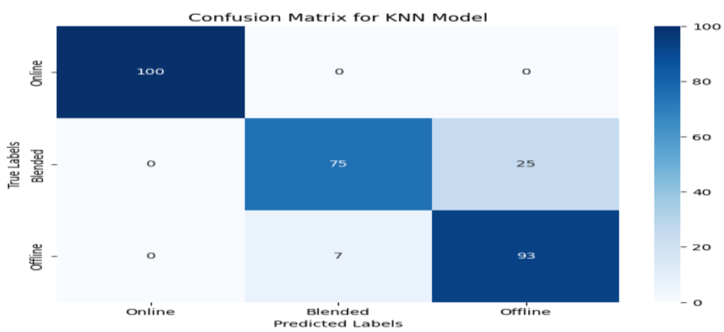

Fig.3 Confusion matrix for the KNN model

V. Conclusion and Future Scope

The MCDM model using the TOPSIS technique is beneficial for criteria analysis based on different alternatives as it gives us the normalized rank of the alternatives based on the criteria. TOPSIS helps us identify the most preferred alternative, whereas the KNN model helps us identify which options lead to which preferred mode. The KNN algorithm produces an accurate model that can predict the overall preference of individuals based on a set of 11 parameters asked. TOPSIS has been chosen for this approach over other multi-criteria decision-making methods because it has ease of application and universality and provides consideration of distances to an ideal solution. Overall preference has been predicted using KNN as it is non-parametric, i.e., it does not make any assumptions on the data, and a lazy learning algorithm, storing all the data before performing any query and then analyzing it when the query is made. It uses the similarity of new data to classify it into a class. It is straightforward to implement and is not significantly affected by noisy data.

In the future, we plan to use other multi-criteria decision-making techniques to calculate the mode with the most optimum technique. Further, we will modify the evaluation criteria and alternatives to make the University/College resilient and scalable in choosing the best education mode for their students.

References

[1] Money, W. H., & Dean, B. P. (2019). Incorporating student population differences for effective online education: A content-based review and integrative model. Computers & Education, 138, 57-82.

[2] Mahyoob, M., Algaraady, J., & Elmahdi, O. E. (2024). Challenges and Sustainability of E-Learning Adoption Among EFL Teachers in the Post-COVID-19 Era. In D. Carbonara & L. Tomei (Eds.), Instructional Technology Theory in the Post-Pandemic Era (pp. 428-454). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-7645-4.ch017

[3] Mahyoob, M. (2021). Online Learning Effectiveness During the COVID-19 Pandemic: A Case Study of Saudi Universities. International Journal of Information and Communication Technology Education (IJICTE), 17(4), 1-14. https://doi.org/10.4018/IJICTE.20211001.oa7.

[4] Bhatia, N. (2010). Survey of nearest neighbor techniques. arXiv preprint arXiv:1007.0085.Carlson, C., Peterson, G., & Day, D. (2019). Utilizing Portable Learning Technologies to Improve Student Engagement and Retention. IEEE Transactions on Education, 63(1), 32-38.

[5] Martínez, P. J., Aguilar, F. J., & Ortiz, M. (2019). Transitioning from face-to-face to blended and full online learning engineering master's program. IEEE Transactions on Education, 63(1), 2-9.

[6] Jain, B., Brar, G., Malhotra, J., & Rani, S. (2017). A novel approach for smart cities in convergence to wireless sensor networks. Sustainable cities and society, 35, 440-448.

[7] Paul, A. B., Biswas, S., Nandi, S., & Chakraborty, S. (2018). MATEM: A unified framework based on trust and MCDM for assuring security, reliability and QoS in DTN routing. Journal of Network and Computer Applications, 104, 1-20.

[8] Wu, H.C. Luk, R.W.P. Wong, K.F. Kwok, and K. Lam. (2008) "Interpreting tf-idf term weights as making relevance decisions," ACM Transactions on Information Systems (TOIS), 26(3), 1-37, Yadav, V., Karmakar, S., Kalbar, P. P., & Dikshit, A. K. (2019). PyTOPS: A Python based tool for TOPSIS. SoftwareX, 9, 217-222.

[9] Ozturk, M. Talco, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," Computers in Biology and Medicine, 2020, pp. 103792.

[10] Mujumdar, A., &Vaidehi, V. (2019). Diabetes prediction using machine learning algorithms. Procedia Computer Science, vol. 165, pp.292-299.

[11] Kurniadi, D., Abdurachman, E., Warnars, H. L. H. S., &Suparta, W. (2018, November). The prediction of scholarship recipients in higher education using k-Nearest neighbor algorithm. In IOP Conference Series: Materials Science and Engineering (Vol. 434, No. 1, p. 012039). IOP Publishing

[12] Zavadskas, E. K., Mardani, A., Turskis, Z., Jusoh, A., & Nor, K. M. (2016). Development of TOPSIS method to solve complicated decision-making problems—An overview on developments from 2000 to 2015. International Journal of Information Technology & Decision Making, 15(03), 645-682

[13] Sarkar, M., & Leong, T. Y. (2000). Application of K-nearest neighbors algorithm on breast cancer diagnosis problem. In Proceedings of the AMIA Symposium (p. 759). American Medical Informatics Association

[14] Lu, J., Song, E., Ghoneim, A., &Alrashoud, M. (2020). Machine learning for assisting cervical cancer diagnosis: An ensemble approach. Future Generation Computer Systems, vol. 106, pp.199-205.

[15] Wu, H.C. Luk, R.W.P. Wong, K.F. Kwok, and K. Lam. (2008) "Interpreting tf-idf term weights as making relevance decisions," ACM Transactions on Information Systems (TOIS), 26(3), 1-37, Yadav, V., Karmakar, S., Kalbar, P. P., & Dikshit, A. K. (2019). PyTOPS: A Python based tool for TOPSIS. SoftwareX, 9, 217-222.

[16] Visa, S., Ramsay, B., Ralescu, A. L., & Van Der Knaap, E. (2011). Confusion Matrix-based Feature Selection. MAICS, 710, 120-127

[17] Dhanabal, S., &Chandramathi, S. (2011). A review of various k-nearest neighbor query processing techniques. International Journal of Computer Applications, 31(7), 14-22.