International Interdisciplinary Virtual Conference on
'Recent Advancements in Computer Science, Management and Information Technology'
International Journal of Scientific Research in Computer Science,
Engineering and Information Technology| ISSN : 2456-3307 (www.ijsrcseit.com)

Cluster Prediction of Students Data by using K-Means Clustering Algorithm

Gautam Kudale[1], Dr. Kailash Patidar[2]
[1]Research Student, [2]Research Guide
Department of Computer Science, Dr. A.P.J. Abdul Kalam University, Indore, Madhya Pradesh, India

ABSTRACT
Academic performance prediction of students is actually a challenging task in current scenario. To enhance the quality of education system, student performance analysis plays a vital role for decision support. Evaluation of students' performance is an important aspect in every educational institute. Important decisions can be made by the academic leaders with the help of the huge data available to them using various algorithms.Clustering is the grouping of a particular set of objects based on their characteristics and aggregating them according to their similarities. In this paper data clustering is used as k-means clustering to evaluate students' performance.
Keywords- Data mining, Clustering, Classification, K Means Clustering Algorithm

I. INTRODUCTION

The growth in information and statement technologies has changed the way in which large quantities of information are accessed, such that the work of academic leaders is reduced or made easy. Important decisions can be made by the academic leaders with the help of the huge data available to them using various algorithms.Thisresearch paper presents k-means clustering algorithm as a simple and efficient tool to monitor the progression of students' academic performance.

Cluster analysis could be divided into hierarchical clustering and non-hierarchical clustering techniques[5,8,25,28]. Examples of hierarchical techniques are single linkage, complete linkage, average linkage, median, and Ward. Non-hierarchical techniques include k-means, adaptive k-means, k-medoids, and fuzzy clustering. To determine which algorithm is good is a function of the type of data available and the particular purpose of analysis. In more objective way, the stability of clusters can be investigated in simulation studies. The problem of selecting the "best" algorithm/parameter setting is a difficult one. A good clustering algorithm ideally should produce groups with distinct non-overlapping boundaries, although a perfect separation cannot typically be achieved in practice. Figure of merit measures (indices) such as the silhouette width score can be used to evaluate the quality of separation obtained using a clustering algorithm. The concept of stability of a clustering algorithm was considered in. The idea behind this validation approach is that an algorithm should be rewarded for consistency. [5,8,25,28]

The clustering algorithm used in the proposed methodology is K-means algorithm. . K-means is one of the easiest algorithms of unsupervised learning used for clustering [5]. Clustering is the grouping of a particular set of objects based on their characteristics and aggregating them according to their similarities[19,20,23]. With respect to data mining this methodology partitions the data implementing a specific join algorithm, most suitable for the desired information analysis. It allows an object not to be part of a cluster, or strictly belong to it, calling this type of grouping hard partitioning.In the other hand, soft partitioning states that every object belongs to a cluster in a determined degree. More specific divisions can be possible to create objects belonging to multiple clusters, to force an object to participate in only one cluster or even construct hierarchical trees on group relationships. The clustering algorithm used in the proposed methodology is parallel k-Means algorithm. It is one of the most popular and simple clustering algorithms is K-Means which, was first published in 1955. In spite of the fact that K-Means was proposed over 50 years ago and thousands of clustering algorithms have been published since then, K-Means is still widely used.

Clustering is the data mining technique that has attracted a great deal of attention in the information industry and in society as a whole, due to the wide availability of huge amount of data and imminent need for turning such data into useful information and knowledge.In this paper, we used k-means clustering algorithm, elbow method and silhouette score in the analysis of the students' academic performance prediction.

II. LITERATURE SURVEY

Prashantsaxena, Govil M. C. [5], in there paper they apply the k-means clustering technique to analyze the relationship between students behavioral and their success. In this paper an extraction method known as principal component analysis is used for predicting cluster analysis. The primary data is collected from a self-finance university, based at Jaipur, India.Questionnaire method is also used to collect data based on some selected input variables. They conclude that type of school is not influence student performance and parent's occupation plays a major role in predicting performance.

Md. Hedayetul Islam Shovon, MahfuzaHaque [7], in there paper present a hybrid procedure based on decision tree and k means data clustering algorithm to predict students GPA and based on this instructor take decisions to improve students' academic performance. 50 training samples are taken for processing. After applying algorithm on training data students are divided in three classes i.e. High, Medium and Low.

Oyelade O. J et al [8], in there paper they implemented the k-means clustering algorithm for analyzing students result data. The model was also combined with the deterministic model to analyze the students result. Database is taken from private institution in Nigeria. They also use Euclidian distance as a measure of similarity distance. They conclude that k means clustering algorithm is good to monitor the performance of students. It also enhances the decision making by academicians to monitor students' progress semester by semester andimprove future academic result.

E.Venkatesan, S.Selvaragini [16], in there paper they use expectation Maximization (EM) and k means algorithm, and sorting algorithms such as C4.5, k-Nearest neighbor and naïve Bayes for prediction of students performance, data is taken from four private Arts Science colleges in Chennai city of Tamilnadu, India. WEKA and Matlab is used to measure the operation of several data mining algorithms. They observe that best clustering algorithm is k means. Also accuracy is verified by classification algorithms J48, JRIP and CART by its various performance criteria. In this classification algorithms CART was found more serious than others.

Yann Ling Goh et al [27] in there paper they use k means clustering algorithm along with deterministic model is used to analyze the students' performance. Data set contains students score in A.Y. 2019 of a college. The number of students is 106 with 8 subjects.They conclude that this methodology will assist academic planners in measuring students' academic performance and assessing student's progression whether students are meeting course requirements.

Zhihui Wang [33] in there paper they use k means clustering algorithm on Iris data set.+e K-means algorithm and the improved K-means algorithm with student information are investigated. As a result, this paperproposes a mechanistic analysis of higher education management and student performance evaluation based on clustering algorithm to assess the quality of college classroom teaching from two perspectives: students' learning effects and teachers' teaching work, with the K-Means algorithm as the primary method. +e theory and application of clustering are highlighted based on a summary of data mining theory. This research presents a set of scientific and reasonable management capability evaluation index systems for universities, which serves as a strong foundation for relevant departments to conduct university administrationcapability evaluations in the future and, as a result, contributes significantly to raising the standard of university administration.

III. METHODOLOGY

K-means clustering algorithm is one of the most widely used algorithms in clustering techniques because of its simplicity and performance. Parallel k-Means algorithm, is used to solve the k-Means clustering problem. The first step in this algorithm is to decide the number of clusters. It is mandatory that the number of clusters decided should match the data. An incorrect choice of the number of clusters will invalidate the whole process. An empirical way to find the best number of clusters is to try parallel k-Means clustering with different number of clusters and measure the resulting sum of squares. Then the center of the clusters should be initialized. The closest cluster should be attributed to each data point and the position of each cluster is set to the mean of all data points belonging to that cluster. This process should be repeated until convergence. The performance of K-means clustering is highly affected when the dataset used is of high dimension. The accuracy and time complexity are highly dropped because of the high dimension data.

Figure 1: Generalised Pseudocode of Traditional k-means [5,8,9,18,25,27]

Srep 1: Step 1: Accept the number of clusters to group data into and the dataset to cluster as input values

Srep 2: Step 2: Initialize the first K clusters

- Take first k instances or
- Take Random sampling of k elements

Srep 3: Step 3: Calculate the arithmetic means of each cluster formed in the dataset.

Srep 4: Step 4: K-means assigns each record in the dataset to only one of the initial clusters - Each record is assigned to the nearest cluster using a measure of distance (e.g Euclidean distance).

Srep 5: Step 5: K-means re-assigns each record in the dataset to the most similar cluster and re-calculates the arithmetic mean of all the clusters in the dataset.

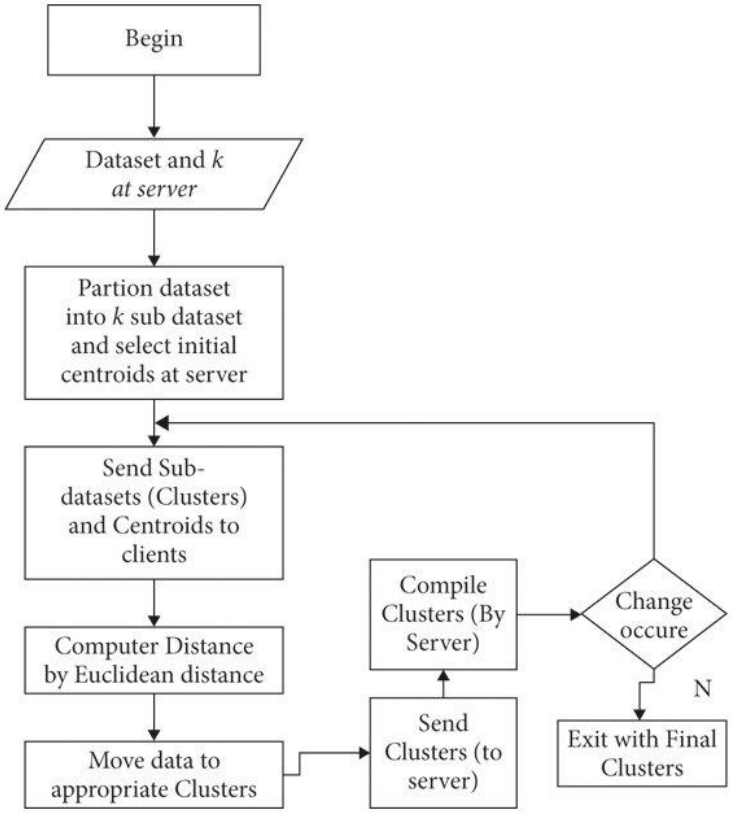


Fig. 2 Flow Chart of K Means Clustering Algorithm

Figure 3: Traditional k-means algorithm [5,8,25]

1. MSE = largenumber;
2. Select initial cluster centroids {mj}j K = 1;
3. Do
4. OldMSE = MSE ;
5. MSE1 = 0;
6. For j = 1 to k
7. mj = 0; nj = 0;
8. end for
9. For i = 1 to n
10. For j = 1 to k
11. Compute squared Euclidean distance d 2(xi, mj);
12. end for
13. Find the closest centroid mj to xi;
14. mj = mj + xi; nj = nj+1;
15. MSE1=MSE1+ d 2(xi, mj);
16. end for
17. For j = 1 to k
18. nj = max(nj, 1); mj = mj/nj;
19. end for
20. MSE=MSE1; while (MSE<OldMSE)

This produces a separation of the objects into groups from which the metric to be minimized can be calculated. The k-means is simple clustering algorithm that has been improved to several problem domains. After obtaining the k partitions, we will get the value of k which is used to predict the student performance by using different machine learning algorithms.

IV. EXPERIMENTAL ANALYSIS AND DISCUSSION

Student academic performance is predicted based on multiple input attributes. Algorithms such as, K-means are used on the input attributes to generate a classification model in-order to predict academic performance of students. In this research, all the pre-processing on the data is done by using different libraries from Python such as Pyspark etc. From the experimental point of view, the dataset is created by importing basic data sets of students from Kaggle.

All the above information will be consolidated as a whole form into complete dataset for the proposed methodology. After applying the K means algorithm, for Elbow method, we got the value of K and its corresponding cost. We have plotted Elbow graph which is used to predict the students' performance. Again, after applying the same K means algorithm, we got the value of K and its corresponding Silhouette score, with which we could plot Silhouette graph.

K means algorithm implementation

The dataset has the attributes "gender", "race/ethnicity", "parental level of education", "lunch", "test preparation course", "math score", "reading score", "writing score". In data processing for first 5 attributes are taken into account.

a. Then for all attributes in dataset following steps are implemented:

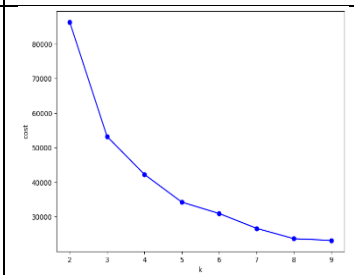
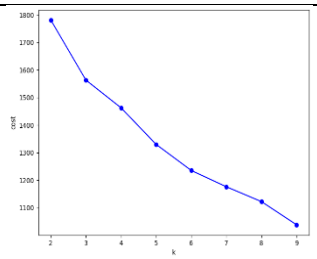
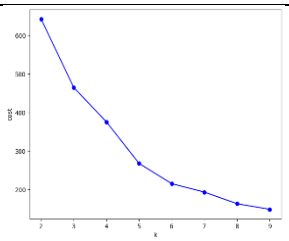
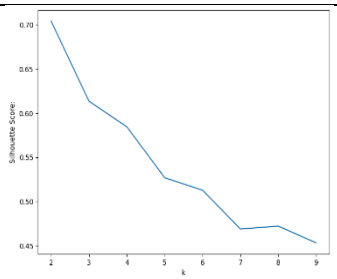
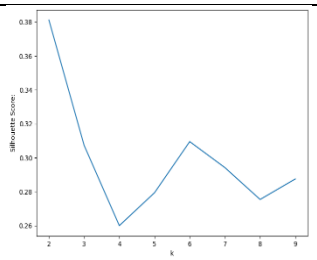
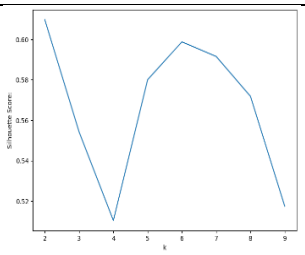
1. Vector assembler is used to assemble the data in to single columnvector which yielded thedf_features.
2. Then StandardScaler is used it creates another column i.e., df_standardized are generated. StandardScaler removes the mean and scales each feature/variable to unit variance.
3. PCA features are extracted using in new column i.e., pcaFeatures are generated. The Principal Component Analysis is a popular unsupervised learning technique for reducing the dimensionality of data. It increases interpretability yet, at the same time, it minimizes information loss. It helps to find the most significant features in a dataset and makes the data easy for plotting in 2D and 3D.

b. Then for last 4 columns in dataset following steps are implemented:

1. Vector assembler is used to assemble the data in to single columnvector which yielded thedf_features.
2. ThenStandardScaler is used it creates another column i.e., df_standardized are generated. StandardScaler removes the mean and scales each feature/variable to unit variance.
3. PCA features are extracted using in new column i.e., pcaFeatures are generated. The Principal Component Analysis is a popular unsupervised learning technique for reducing the dimensionality of data. It increases interpretability yet, at the same time, it minimizes information loss. It helps to find the most significant features in a dataset and makes the data easy for plotting in 2D and 3D.

Above steps a and b are used to create input and the elbow plot and silhouette score plot id plotted to identify proper clustering.

FIG2. ELBOW PLOT AND SILHOUETTE SCORE PLOT FOR ALL COLUMNS

for all columns with 300 entries	A) df_features	B) df_standardized	C) pcaFeatures
Elbow Plot			
Silhouette Score Plot			

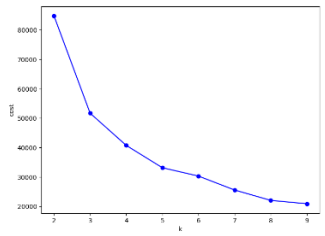
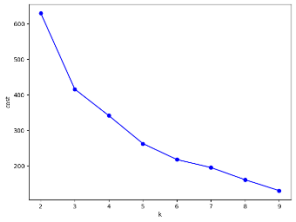
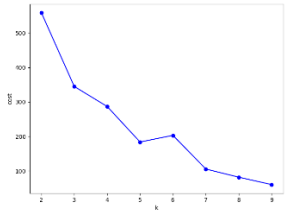
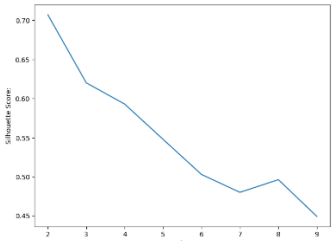
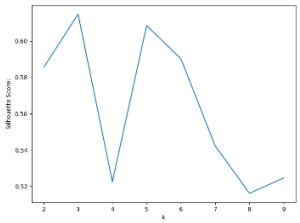
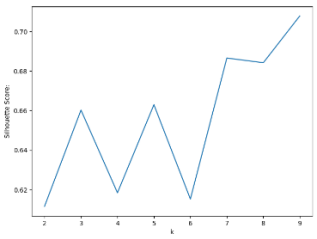
In the above plots Fig 2A i.e.Graph ofDf_features, we see elbow plot in which we have to see the first or most significant turning point of the curve which is visible as an elbow which suggest the right number of clusters [4]. In this case, it can be 4, 5, 6, 7 or any of these. The Silhouette score reaches its global maximum at the optimal k. This should ideally appear as a peak in the silhouette values versus-k plot. But there is no clarity with the Silhoutte plot. There is no a clear maximum or minima visible.

In the above plots Fig 2B i.e.Graph ofDf_Standardized, we see elbow plot in which we have to see the first or most significant turning point of the curve which is visible as an elbow which suggest the right number of clusters [4]., here we cannot see elbow like bend, so no clarity, same case with Silhoutte plot, no clear maxima, minima are visible.

In the above plots Fig 2C i.e. Graph of PCA features, we see an elbow plot in which we have to see the first or most significant turning point of the curve which is visible as an elbow which suggest the right number of clusters [4].In this case also, it can be 4, 5, 6, 7 or any of these.

To find exact answer, we can take help from Silhouette plot;theSilhouette score reaches its global maximum at the optimal k. This should ideally appear as a peak in the silhouette values versus-k plot. In this case, it is 5 and it is present in list of elbow point, so we can select 5 as a number of clusters for these.

FIG3. ELBOW PLOT AND SILHOUETTE SCORE PLOT FOR LAST FOUR COLUMNS

For Last four columns	A) df_features	B) df_standardized	C) pcaFeatures
Elbow Plot			
Silhouette Score Plot			

In the above plots Fig 3A i.e.Graph ofDf_features, we see Elbow plot in which we have to see the first or most significant turning point of the curve which is visible as an elbow which suggest the right number of clusters [4]. In this case, it can be 4, 5, 6, 7 or any of these. The Silhouette score reaches its global maximum at the optimal k. This should ideally appear as a peak in the silhouette values versus-k plot. But there is no clarity with the Silhoutte plot. There is no a clear maximum or minima visible.

In the above plots Fig 3B i.e.Graph ofDf_Standardized, In the above plots, we see the first or most significant turning point of the curve which is visible as an elbow which suggest the right number of clusters [4]. In this case also, it can be 4, 5, 6, 7 or any of these.

To find exact answer, we can take help from Silhouette plot,the Silhouette score reaches its global maximum at the optimal k. This should ideally appear as a peak in the silhouette values versus-k plot. In this case, it is 5 and it is present in list of elbow point, so we can select 5 as a number of clusters.

In the above plots Fig 3C i.e.Graph of PCA features, we see elbow plot in which we have to see the first or most significant turning point of the curve which is visible as an elbow which suggest the right number of clusters [4], so in this case there is no clarity, same case with Silhoutte plot, no clear maxima, minima are visible.

The Elbow Method: This is one of the most popular methods to determine the optimal number of clusters. It is a little bit simple approach. In this method, we calculate the cost which consists of sum of squared distances of points to their nearest centres.

The drawback of Elbow method is that sometimes we cannot get the optimal value of k. We can get the ambiguous value of k. In such ambiguous situation, we have to use the Silhoutte method.

The Silhoutte method: This method estimates a value which shows how a point is closer to its own cluster as compared to other clusters. The value of silhouette coefficient is between -1 to 1 [4].

One cannot bypass the Elbow method and consider only The Silhoutte one. The Elbow method is used to get a rough estimate of k whereas Silhoute value method is used to get the exact value of k. Both the methods conjunctively form a tool for us to take confident decision for the determination of value of k.

V. CONCLUSION

In this paper we have taken student dataset of 300 records from Kaggle, we have applied parallel K means algorithm on the dataset. Then vector assembler is used to assemble the data in to single (column) vector i.e., df_features are generated. Then StandardScaler is used it creates another column i.e., df_standardized are generated. PCA features are extracted using in new column i.e., pcaFeatures are generated. The df_features, df_standardized and pcaFeatures are used to create input and the elbow plot and silhouette score plot id are plotted to identify proper clustering. It is observed that for all attributes the PCA features results looks good. Also for last four columns the df_standardized results looks good. In both case we get five clusters i.e. K=5.

VI. ACKNOWLEDGEMENT

I would like to thank Dr.KailashPatidar, Research Guide, Dr. A.P.J. Abdul KalamUniversity,Indore, M.P. for giving me valuable guidelines and suggestions regarding this work. I would like to thank Dr.RanjitPatil, Principal, Dr. D. Y. Patil Arts, Science and Commerce College, Pimpri, Pune for writing research paper and useful discussions. I also like to express my sincere thanks to Dr. Bharat Shinde, Principal, VidyaPratishthan's Arts, Science & Commerce College, M.I.D.C., Baramati, Pune. (MS).

VII. REFERENCES

[1]. Data Mining Introductory and Advanced Topics, Margaret H. Dunhan, Pearson
[2]. Data Mining Practical Machine Learning Tools and Techniques, 3rd Edition, Ian H.witten, Eibe Frank, Mark A. Hall
[3]. Mining of Massive Datasets, 2nd Edition, Jure Leskovec, Anand Rajaraman, Jeffrey David Ullman
[4]. Data Mining, Concepts and Techniques, 3rd Edition, Jiawei Han, Micheline Kamber, Jian Pei
[5]. Prof. Prashant Sahai Saxena, Prof. M. C. Govil, “Prediction of Student’s Academic Performance using Clustering,” Special Conference Issue: National Conference on Cloud Computing & Big Data
[6]. Bindiya M Varghese, Jose Tomy J, Unnikrishnan A, Poulose Jacob K, “Clustering student data to characterize performance patterns,” (IJACSA) International Journal of Advanced Computer Science and Applications, Special Issue on Artificial Intelligence,
[7]. Md. Hedayetul Islam Shovon, Mahfuza Haque, “An Approach of Improving Student’s Academic Performance by using K-means clustering algorithm and Decision tree,” (IJACSA) International Journal of Advanced Computer Science and Applications, Vol.3, No. 8, 2012
[8]. Oyelade, O. J, Oladipupo, O. O., Obagbuwa, I. C., “Application of k-Means Clustering algorithm for prediction of Students’ Academic Performance,” (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, o. 1, 2010
[9]. Rakesh Kumar Arora, Dr. Dharmendra Badal, “Evaluating Student’s Performance Using k-Means Clustering,” International Journal of Computer Science And Technology, IJCST Vol. 4, Issue 2, April - June 2013, ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

[10]. Sharmila, R.C Mishra, "Performance Evaluation of Clustering Algorithms," International Journal of Engineering Trends and Technology (IJETT) - Volume4 Issue7- July 2013, ISSN: 2231-5381

[11]. Ramjeet Singh Yadav, P. Ahmed, A. K. Soni and Saurabh Pal, "Academic performance evaluation using soft computing techniques," CURRENT SCIENCE, VOL. 106, NO. 11, 10 JUNE 2014

[12]. Harwatia, Ardita Permata Alfiania, Febriana Ayu Wulandaria, "Mapping student's performance based on data mining approach (a case study)," The 2014 International Conference on Agro-industry (ICoA): Competitive and sustainable Agro industry for Human Welfare, Agriculture and Agricultural Science Procedia 3 (2015) 173 – 177

[13]. Patel, J. and Yadav, R.S. (2015) "Applications of Clustering Algorithms in Academic Performance Evaluation." Open Access Library Journal, 2: August 2015 | Volume 2 | e1623

[14]. Jyotirmay Patel, Ramjeet Singh Yadav, "Applications of clustering algorithms in academic performance evaluation"

[15]. Atul Prakash Prajapati, Sanjeev Kr. Sharma, Manish Kr. Sharma, "Student's performance analysis using machine learning tools," International Journal of Scientific & Engineering Research Volume 8, Issue 10, October-2017 ISSN 2229-5518

[16]. E.Venkatesan, S.Selvaragini, "Prediction of students academic performance using classification and clustering algorithms," International Journal of Pure and Applied Mathematics Volume 116 No. 16 2017, 327-333 ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)

[17]. Snehal Bhogan , Kedar Sawant , Purva Naik , Rubana Shaikh , Odelia Diukar , Saylee Dessai, "Predicting student performance based on clustering and classification," IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN:2278-8727, Volume 19, Issue 3, Ver. V (May-June 2017), PP 49-52

[18]. Mr. Shashikant Pradip Borgavakar, Mr. Amit Shrivastava, "Evaluating student's performance using k-means clustering," International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 6 Issue 05, May – 2017

[19]. Mrs .Mary vidya john, Akshata police patil, Anjali mishra, Bindhu reddy G, Jamuna N, "Clustering technique for student performance," International Research Journal of Computer Science (IRJCS), Issue 06, Volume 6 (June 2019), ISSN: 2393-9842

[20]. Noel Varela , Edgardo Sánchez Montero , Carmen Vásquez , Jesús García Guiliany , Carlos Vargas Mercado , Nataly Orellano Llinas , Karina Batista Zea , and Pablo Palencia, "Student performance assessment using clustering techniques," © Springer Nature Singapore Pte Ltd. 2019 Y. Tan and Y. Shi (Eds.): DMBD 2019, CCIS 1071, pp. 179–188, 2019. https://doi.org/10.1007/978-981-32-9563-6_19

[21]. N.Valarmathy, S.Krishnaveni, "Performance evaluation and comparison of clustering algorithms used in educational data mining," International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6S5, April 2019

[22]. Lubna Mahmoud Abu Zohair, "Prediction of Student's performance by modelling small dataset size," Abu Zohair International Journal of Educational Technology in Higher Education (2019) 16:27 https://doi.org/10.1186/s41239-019-0160-3

[23]. Mrs. Bhawna Janghel, Dr. Asha Ambhaikar, "Performance of student academics by k-mean clustering algorithm," International J. Technology. January – June, 2020; Vol. 10: Issue 1, ISSN 2231-3907 (Print), ISSN 2231-3915 (Online)

[24]. Marzieh Babaie, Mahdi Shevidi Noushabadi, “A review of the methods of predicting students' performance using machine learning algorithms,” Archives of Pharmacy Practice ¦ Volume 11 ¦ Issue S1 ¦ January-March 2020

[25]. Dr. G. Rajitha Devi, “Prediction of student academic performance using clustering,” International Journal of Current Research in Multidisciplinary (IJCRM) ISSN: 2456-0979 Vol. 5, No. 6, (June’20), pp. 01-05

[26]. Dewi Ayu Nur Wulandari; Riski Annisa; Lestari Yusuf, Titin Prihatin, “An educational data mining for student academic prediction using k-means clustering and naïve bayes classifier,” journal Pilar Nusa Mandiri Vol 16, No 2 September 2020

[27]. Yann Ling Goh, Yeh Huann Goh, Chun-Chieh Yip, Chen Hunt Ting, Raymond Ling Leh Bin, Kah Pin Chen, “Prediction of students' academic performance by k-means clustering,” Peer-review under responsibility of 4th Asia International Multidisciplinary Conference 2020 Scientific Committee

[28]. Revathi Vankayalapati, Kalyani Balaso Ghutugade, Rekha Vannapuram, Bejjanki Pooja Sree Prasanna, “K-means algorithm for clustering of learners performance levels using machine learning techniques,” Revue d'Intelligence Artificielle Vol. 35, No. 1, February, 2021, pp. 99-104

[29]. Rina Harimurti , Ekohariadi, Munoto , I. G. P Asto Buditjahjanto, “Integrating k-means clustering into automatic programming assessment tool for student performance analysis,” Indonesian Journal of Electrical Engineering and Computer Science Vol. 22, No. 3, June 2021, pp. 1389~1395 ISSN: 2502-4752, DOI: 10.11591/ijeecs.v22.i3.pp1389-1395

[30]. Rui Shang , Balqees Ara, Islam Zada, Shah Nazir , Zaid Ullah, and Shafi Ullah Khan, “Analysis of simple k-mean and parallel k-mean clustering for software products and organizational performance using education sector dataset,” Hindawi Scientific Programming Volume 2021, Article ID 9988318, 20 pages https://doi.org/10.1155/2021/9988318

[31]. Bao Chong, “K-means clustering algorithm: a brief review,” Academic Journal of Computing & Information Science ISSN 2616-5775 Vol. 4, Issue 5: 37-40, DOI: 10.25236/AJCIS.2021.040506

[32]. Said Abubakar Sheikh Ahmed, “Evaluating students’ performance of social work department using k-means and two-step cluster “a case study of mogadishu university”,” Mogadishu University Journal, Issue 7, 2021, ISSN 2519-9781

[33]. Zhihui Wang, “Higher education management and student achievement assessment method based on clustering algorithm,” Hindawi Computational Intelligence and Neuroscience Volume 2022, Article ID 4703975, 10 pages https://doi.org/10.1155/2022/4703975

[34]. Ahmad Fikri Mohamed Nafuri , Nor Samsiah Sani, Nur Fatin Aqilah Zainudin , Abdul Hadi Abd Rahman and Mohd Aliff, “Clustering analysis for classifying student academic performance in higher education,” Appl. Sci. 2022, 12, 9467. https://doi.org/10.3390/app12199467