Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:11(2), November, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed: Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available: www.ijmer.in

A STUDY ON CATEGORICAL DATA CLUSTERING

Dr.K. Asha Latha

1. Introduction

Over the past decade, the amount of information accumulated every second has become a treasure of inestimable value. Social media sites, sensors, transactions records and many other sources that come from everywhere are behind the Big Data phenomenon. Consequently, considerable efforts have been devoted to exploring such massive data in order to gain the maximum benefit from this treasure. As a solution to this problem, the research is intensified in the direction of parallel clustering methods (Lamari, 2017).

Clustering is a process, grouping a set of physical or abstract objects into classes of similar objects. In other words, the method of identifying similar groups of data in a dataset is called clustering (Koushik, 2016).

The classical definition of cluster was attributed by M. Porter (2008): "Educational cluster is a group of geographically neighbouring interconnected companies and organizations connected to them, working in a certain area and characterized by common activities and mutual reinforcement".

In clustering, the goal is to find data points that naturally group together, splitting the full data set into a set of clusters. Clustering is particularly useful in cases where the most common categories within the data set are not known in advance. If a set of clusters is optimal, within a category, each data point will be in general being more similar to the other data points in that cluster than data points in other clusters. Clusters can be created at several different possible grain-sizes: for example, schools could be clustered together (to investigate similarities and differences between schools), students could be clustered together (to investigate similarities and differences between students), or student actions could be clustered together (to investigate patterns of behavior) (cf. Amershi&Conati, 2006; Beal, Qu, & Lee, 2006).

A good clustering method will produce high quality clusters with high intra-class similarity - Similar to one another within the same cluster low inter-class similarity - Dissimilar to the objects in other clusters The quality of a clustering result depends on both the similarity measure used by the method and its implementation.

Type of data in cluster analysis include: Interval-scaled variables; Binary variables; Nominal, ordinal, and ratio variables; Variables of mixed types; Complex data types.

Clustering techniques can be broadly classified into many categories; partitioning, hierarchical, density-based, grid-based, model-based algorithms.

2. Review of Literature:

In a nutshell, He, Z., Xu, X., Deng, S., & Huang, J. Z. (2004) proposed an efficient clustering algorithm for analyzing categorical data streams; Kotsiantis et al. (2004) propounded five classification algorithms; Romero, C., & Ventura, S. (2007) unearthed application of data mining to traditional educational systems; IndrajitSaha and AnirbanMukhopadhyay (2008) demonstrated a variety of artificial and real life categorical data sets.; Do, H. J., & Kim, J. Y. (2008) indicated a new clustering algorithm for categorical data; Yu et al (2010) explored student retention by using classification trees; Ramaswami and Bhaskaran (2010) focused on developing predictive data mining model to identify the slow learners; Aranganayagi, S., &Thangavel, K. (2010) presented an incremental algorithm to cluster the categorical data; Sayal, R., & Kumar, V. V. (2011) overviewed of popular similarity measures of categorical attributes; Rezankova, H., Loster, T., &Husek, D. (2011) studied criteria based on variability measures; Baradwaj, B. K., & Pal, S. (2012) attempted data mining techniques in context of higher education; Kalaivani, K., &Raghavendra, A. P. V. (2012) presented categorical data set; Md.Hedayetul Islam Shovon (2012) presented a paper on prediction of student academic performance by applying K-means clustering; Sisodia, D., Singh, L., Sisodia, S., &Saxena, K. (2012) dealt with the study of various clustering algorithms of data mining; SwastiSinghal, Monika Jena (2013) introduced the WEKA tool; Chandrika, J., & Kumar, K. A.

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:11(2), November, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed: Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available: www.ijmer.in

(2013) delineated cluster the transactional data streams; Venkatesan, N. (2013) discussed the types of modeling technique; Kabakchieva, (2013) high potential of data mining applications for university management; Durairaj et al., (2014) proposed Educational Data mining; Natek, Srečko, and MotiZwilling., (2014) focused on the study of data mining techniques; Prashant et al. (2014) examined the clustering analysis in data mining; Veeramuthu et al (2014) analyzed how different factor affect a Students learning behavior; Shiwani and Roopali (2016) applied unsupervised learning algorithms; Gul'zamira D. Aitbayeva et al (2016) studied educational clusters; Sowmiya and Valarmathi. (2017) presented the literature review of the clustering algorithm for categorical and binary attributes; Abdul Rahmat (2017) studied transformational intellectual; Uddin J, Ghazali R, Deris MM (2017) proposed an alternative technique named Maximum Indiscernible Attribute (MIA); Sangam, R. S., & Om, H. (2017) proposed k-mode stream; Govindasamy, K., &Velmurugan, T. (2018) studied four clustering algorithms; Lakshmi Sreenivasa Reddy and Rajini (2018) strategic management tool; Qin, H., & Ma, X. (2018) propounded IG-ANMI; Amir Ahmad and SherozS.Khan (2019) presented taxonomy for the study of mixed data clustering algorithms.

Thus, the less trodden road is taken to find out clustering data with special reference to categorical clustering especially, in the State of Telangana.

3. Significance of the study

Clustering of categorical data is becoming increasingly important, since non-numerical data are ubiquitous and clustering can be used, for example, to optimize an anonymization process or to perform anomaly detection, or in any application where there is the need to automatically recognize the intrinsic structure of data.

An example of categorical attribute is shape whose values include circle, rectangle, ellipse, etc. Due to the special properties of categorical attributes; the clustering of categorical data seems more complicated than that of numerical data. Many algorithms focus on numerical data whose inherent geometric properties can be used naturally to define distance function between data points. However much of the data existed in the database is categorical where attributes values cannot be naturally ordered as numerical values. Categorical data has a different structure than the numerical data. The distance functions in the numerical data might not be applicable to the categorical data. Algorithms for clustering numerical data cannot be applied to categorical data (Suchitra, 2012).

An educational cluster is a group of educational institutions within certain territory, that as a final product form an educational service, competitive and interacting providers of necessary factors of industry, equipment, specialized services, utilities, research and development centres, which reinforce each other's advantages (figure 1). As well as the majority of competitive clusters, educational clusters occur naturally due to the existence and interaction of a significant amount of factors (Mukhametzyanova&Pugacheva, 2010).

The data containing categorical attributes pose a number of challenges on the existing clustering methods due to a) no natural order; b) high dimensionality; c) existence of sub-space clusters and d) conversion of categorical to numerical data.

4. Statement of the Problem

'A Study On Categorical Clustering Data'

5. Objectives of the study

1. To explore the *levels* of categorical data clustering among the ITI students in Bhadradri Kothagudem District of Telangana State.
2. To analyze the categorical data clustering among the ITI students in Bhadradri Kothagudem District of Telangana State with reference to the *gender.*
3. To study the categorical data clustering among the ITI students in Bhadradri Kothagudem District of Telangana State with special reference to *management.*

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:11(2), November, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed: Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available: www.ijmer.in

6. Research Hypotheses

Keeping in view the review literature and the objectives, the following hypotheses have been formulated. As pointed out earlier, this study is an investigation into the*categorical data clusteringin education*is rather a new area and number of questions and controversial issues awaiting answers and clarifications. Thus, this study throws light on the following negative research hypotheses.

1. There is no statistically significant difference on the levels of categorical data clustering among ITI students in BadradriKothagudem District of Telangana State.
2. There is no statistically significant difference on the categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to gender.
3. There is no statistically significant difference on the categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to management.

7. Methodology of the Study:

Sample

In order to select the representative sample for the study, simple random sampling technique was used. **One hundred and eighty students**, both boys and girls from Industrial Training Institutes of Government and Private from ***BadradriKothagudem*** district were selected for the present investigation.

Instrumentation

Table No.1 shows individual data for the total number of categories formed and the range of words in those categories

Recall Tests	No. of Categories formed and No. of words recalled	No. of clusters recalled and (No. of words recalled under each cluster)				Total clusters
		I	II	III	IV	
		Fruits	Flowers	Animals	Cities	
Recall Test -1	3(21)	0 (0)	1(6)	1(5)	2(8)	4
Recall Test -2	4(26)	1(3)	1(3)	4(9)	2(8)	8
Recall Test -3	4(33)	2(9)	2(4)	2(9)	1(8)	7

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:11(2), November, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed: Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available: www.ijmer.in

Administration of the instrument

The tool was administered to the selected sample. Every care has been taken to ensure their responses as objectively as possible. Suitable statistical techniques like t.-test and ANOVA along with SPSS (Statistical Package for Social Sciences) 16.0 was used for analysis.

8. ANALYSIS AND INTERPRETATION OF DATA

HO_1. There is no statistically significant difference on the levels of categorical data clustering among ITI students in BadradiKothagudem District of Telangana State.

Table 4.1 showing mean scores and ANOVA on the levels of Categorical data clustering.

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
GOVTIT Boys	45	24.4444	7.77688	1.15931	22.1080	26.7809	8.00	38.00
GOVTITI Girls	45	20.1333	7.47298	1.11401	17.8882	22.3785	8.00	39.00
PV TITI Boys	45	25.3556	6.97123	1.03921	23.2612	27.4499	11.00	38.00
PVT ITI Girls	45	20.5778	7.32830	1.09244	18.3761	22.7794	8.00	39.00
Total	**180**	**22.6278**	**7.68517**	**.57282**	**21.4974**	**23.7581**	**8.00**	**39.00**

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	952.461	3	317.487	5.809	.001
Within Groups	9619.600	176	54.657		
Total	**10572.061**	**179**			

It can be observed from the ANOVA table the calculated p-value is 0.001, which is highly significant. Moreover, the mean score of Private ITI Boys have is 25.3556 ± 6.97123 followed by Government ITI boys 24.4444 ±7.77688. Thus, it can be inferred that there is a statistically significant difference on the levels of categorical data clustering among ITI students in BadradriKothagudem District of Telangana State. Hence, the hypothesis formulated was **rejected**.

HO_2: There is no statistically significant difference on the categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to *gender.*

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:11(2), November, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed: Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available: www.ijmer.in

Table 4.2 showing mean scores and t-test on the Categorical Clustering.

Gender	N	Mean	Std. Deviation	Std. Error Mean
Boys Private ITI	45	25.3556	6.97123	1.03921
Girls Private ITI	45	20.5778	7.32830	1.09244
Boys Govt ITI	45	24.4444	7.77688	1.15931
Girls Govt ITI	45	20.1333	7.47298	1.11401

Levene's Test for Equality of Variances		**t-test for Equality of Means**						
F	Sig.	t	df	Sig. (2-tailed)	Mean Diff	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
.281	.598	2.681	88	.009	4.31111	1.60780	1.11596	7.50627
		2.681	87.861	.009	4.31111	1.60780	1.11589	7.50634
.008	.929	3.169	88	.002	4.77778	1.50777	1.78140	7.77416
		3.169	87.781	.002	4.77778	1.50777	1.78129	7.77426

The ANOVA table demonstrates the calculated p-value is 0.009, which is highly significant at 0.005 levels. Hence, it can be deduced that there is a statistically significant difference on the levels of categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to gender. Moreover, Mean ± Sd is found to be very high in boys when compared to girls. The descriptive statistics shows 24.4444 ± 7.77688 (Boys, Govt ITI) and 25.3556 ± 6.97123 (Boys Private, ITI). Thus, the hypothesis was **rejected.**

HO_3: There is no statistically significant difference on the categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to *management.*

Table 4.3 showing mean scores and t-test on the Categorical data Clustering

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:11(2), November, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed: Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available: www.ijmer.in

Gender	N	Mean	Std. Deviation	Std. Error Mean
Boys Private ITI	45	25.3556	6.97123	1.03921
Girls Private ITI	45	20.5778	7.32830	1.09244
Boys Govt ITI	45	24.4444	7.77688	1.15931
Girls Govt ITI	45	20.1333	7.47298	1.11401

Levene's Test for Equality of Variances		**t-test for Equality of Means**						
F	Sig.	t	df	Sig. (2-tailed)	Mean Diff	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
.335	.564	-.585	88	.560	-.91111	1.55690	-4.00513	2.18291
		-.585	86.968	.560	-.91111	1.55690	-4.00564	2.18342
.005	.944	-.285	88	.776	-.44444	1.56027	-3.54515	2.65626
		-.285	87.966	.776	-.44444	1.56027	-3.54516	2.65627

The ANOVA table demonstrates the calculated p-value is 0.556, which is highly significant at 0.005 levels. Hence, it can be deduced that there is a statistically significant difference on the levels of categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to management.

Further, Mean ± Sd is found to be very high in boys when compared to girls. The descriptive statistics shows it is 24.4444 ± 7.77688 (Boys, Govt ITI) and 25.3556 ± 6.97123 (Boys Private, ITI). Thus, the hypothesis was **rejected.**

9. Major Findings

1. The present study revealed that there is a statistically significant difference on the levels of categorical data clustering among ITI students in BadradriKothagudem District of Telangana State.

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:11(2), November, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed: Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available: www.ijmer.in

2. The findings of the study explicitly demonstrated there is a statistically significant difference on the levels of categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to gender.

3.The quantitative results obtained disclosed that there is a statistically significant difference on the levels of categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to management.

10. Discussion and Conclusions

HO_1: There is no statistically significant difference on the levels of categorical data clustering among ITI students in BadradriKothagudem District of Telangana State.

The mean score of Private ITI Boys have is 25.3556 ± 6.97123 followed by Government ITI boys 24.4444 ±7.77688. Thus, it can be inferred that there is a statistically significant difference on the levels of categorical data clustering among ITI students in BadradriKothagudem District of Telangana State. Hence, the hypothesis formulated was **rejected**.

HO_2. There is no statistically significant difference on the categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to gender.

The calculated p-value is 0.009, which is highly significant at 0.005 levels. Hence, it can be deduced that there is a statistically significant difference on the levels of categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to gender. Moreover, Mean ± Sd is found to be very high in boys when compared to girls. The descriptive statistics shows 24.4444 ± 7.77688 (Boys, Govt ITI) and 25.3556 ± 6.97123 (Boys Private, ITI). Thus, the hypothesis was **rejected.**

HO_3. There is no statistically significant difference on the categorical data clustering among ITI students in BadradriKothagudem District of Telangana State with reference to management.

The Mean ± Sd is found to be very high in boys when compared to girls. The descriptive statistics shows it is 24.4444 ± 7.77688 (Boys, Govt ITI) and 25.3556 ± 6.97123 (Boys Private, ITI). Thus, the hypothesis was **rejected.**

BIBILIOGRAPHY

Aranganayagi, S., &Thangavel, K. (2010). Incremental algorithm to cluster the categorical data with frequency based similarity measure.*World Academy of Science, Engineering and Technology*. Vol:4 2010-01-23

Baradwaj, B. K., & Pal, S. (2012). Mining Educational Data to analyze students' performance. *arXiv preprint arXiv:1201.3417*.

Chandrika, J., & Kumar, K. A. (2013).A Novel Approach for Clustering Categorical Data Streams. *International Journal of Innovation, Management and Technology*, *4*(5), 486.

Do, H. J., & Kim, J. Y. (2008), FAVC: Clustering Categorical Data Using the Frequency of Attribute Values Combinations. In *2008 3rd International Conference on Innovative Computing Information and Control* (pp. 304-304).IEEE.

Govindasamy, K., &Velmurugan, T. (2018).Analysis of Student Academic Performance Using Clustering Techniques. *International Journal of Pure and Applied Mathematics*, *119*(15), 309-323.

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:11(2), November, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed: Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available: www.ijmer.in

He, Z., Xu, X., Deng, S., & Huang, J. Z. (2004). Clustering Categorical data streams. *arXiv preprint cs/0412058*.

IndrajitSaha andAnirbanMukhopadhyay (2008) Improved Crisp and Fuzzy Clustering Techniques for Categorical Data. *International Journal of computer Sciences*. 35 (4).On line publication.

Kabakchieva,.D (2013) Analyzing University Data for Determining Student Profiles and Predicting Performance, Cybernetics and Information Technologies, Vol.1(3).

Kalaivani, K., &Raghavendra, A. P. V. (2012). Efficiency based categorical data clustering. In *2012 IEEE International Conference on Computational Intelligence and Computing Research* (pp. 1-4). IEEE.

Kotsiantis, S., C. Pierrakeas, P. Pintelas. Prediction of Student's Performance in Distance Learning Using Machine Learning Techniques. Applied Artificial Intelligence,Vol.18,2004,No5,411-426.

Lakshmi Sreenivasa Reddy and Rajini (2018)A Proposal to Predict Student's Performance using Data Mining Techniques.*International Journal of Computer & Mathematical Sciences.*IJCMS.Volume 7, Issue 2.pp.97-101.

Md. Hedayetul Islam Shovon,(2012) Prediction of Student Academic Performance by an Application of K-Means Clustering Algorithm, *International Journal of Advanced Research in Computer Science and Software Engineering,* Vol. 2(7).

Qin, H., & Ma, X. (2018).IG-ANMI: a novel initialization method for genetic clustering algorithm for categorical data. *International Journal of Science, Engineering and Technology (IJSET) UTY*, *1*(1), 53-66.

Rahamat (2017) Clustering in Education.*European Research Studies Journal* Volume XX, Issue 3A, 2017 pp. 311-324.

Ramaswami, M., R. Bhaskaran. A CHAID Based Performance Prediction Model in Educational Data Mining. – IJCSI *International Journal of Computer Science* Issues, Vol. 7, January 2010, Issue 1, No 1, 10-18.

Rezankova, H., Loster, T., &Husek, D. (2011).Evaluation of categorical data clustering.In *Advances in Intelligent Web Mastering–3* (pp. 173-182).Springer, Berlin, Heidelberg.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert systems with Applications*, *33*(1), 135-146.

Sangam, R. S., & Om, H. (2017). K-modestream algorithm for clustering categorical data streams. *CSI Transactions on ICT*, *5*(3), 295-303.

Sayal, R., & Kumar, V. V. (2011).A novel similarity measure for clustering categorical data sets. *International Journal of Computer Applications*, *17*(1), 25-30.

Shruti Sharma, Manoj Singh (2015) Clustering with Categorical Data-A Survey.International Journal of Engineering, Management & Sciences (IJEMS). Volume-2, Issue-12,pp.1-5.

Sisodia, D., Singh, L., Sisodia, S., &Saxena, K. (2012). Clustering techniques: a brief survey of different clustering algorithms. *International Journal of Latest Trends in Engineering and Technology (IJLTET)*, *1*(3), 82-87.

Sowmiya, N., &Valarmathi, B.(2007) A Review of categorical data clustering methodologies based on recent studies.*IIOABJ.* Vol. 8 (2.) pp.353-365.

Cover Page

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN:2277-7881(Print); IMPACT FACTOR :9.014(2025); IC VALUE:5.16; ISI VALUE:2.286
PEER REVIEWED AND REFEREED INTERNATIONAL JOURNAL
(Fulfilled Suggests Parameters of UGC by IJMER)
Volume:14, Issue:11(2), November, 2025
Scopus Review ID: A2B96D3ACF3FEA2A
Article Received: Reviewed: Accepted
Publisher: Sucharitha Publication, India
Online Copy of Article Publication Available: www.ijmer.in

Uddin J, Ghazali R, Deris MM (2017) An Empirical Analysis of Rough Set Categorical Clustering Techniques. *PLoS ONE* 12(1): e0164803.

Veeramuthu, P., Periyasamy, D. R., &Sugasini, V. (2014).Analysis of student result using clustering techniques. *International Journal of Computer Science and Information Technologies*, *5*(4), 5092-5094.

Venkatesan, N. (2013). Role of Data Mining Techniques in Educational and E-learning System. *Asia Pacific Journal of Research*, *2*.

Yu, C., S. DiGangi, A. JannaschPennell, C. Kaprolet.(2010) A Data Mining Approach for Identifying Predictors of Student Retention from Sophomore to Junior Year. *Journal of Data Science*, Vol. 8,pp. 307-325.