

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 9, September 2019

## A Survey on Mentality about the Student Curriculum Using Weka Tool in Data Mining Technique

P.Dhanalakshmi<sup>1</sup>, K.Radhakrishnan<sup>2</sup>

M.Phil Research Scholar, P.G & Research, Department of Computer Science, Dr.Ambedkar Govt. Arts College,  
Vyasarpadi, Chennai, India <sup>1</sup>

Assistant Professor, P.G & Research, Department of Computer Science, Dr.Ambedkar Govt. Arts College,  
Vyasarpadi, Chennai, India <sup>2</sup>

**ABSTRACT:** Nowadays, The purpose of this study was to mentality about the student and use for evaluating the curriculum quality and teaching. The quality of teaching necessary inside colleges. It is great theoretical and practical through the construction and implementation. The process working for Knowledge Discovery and Database. KDD is referred to as data mining. The data mining technique may improve the competence of educational institution. At the present a day the data in the educational institutions has augmented to large folds .These databases contains a lot of concealed information that can be use for get better student's performance. Should be make the idea through the **questionnaire were then spread among the students and finally collected from them.Each student over following the questionnaire with indolence and according to his our mind set real data** embryonic a growth mindset isn't easy, however it is achievable with time and effort. Students develop a growth mindset by recognising with the aim of the brain isn't fixed entity plus that improvements. It can be made through learning new information Growth way of thinking know how to be developed in the class room by providing students with a range of different strategy to help improve the way they study and their ability to cope with pressure.

**KEYWORDS:** Data Pre-Processing, Data quality, Clustering Technique

### I. INTRODUCTION

In this dissertation, introduces the method of key technology used in the Development rout. It was accumulates and relegates make-believe tale identifies consequential work and mediates it to computing educators and we identify research that gives well-fortified advice to a mind informative and rejuvenate the more impuissant segment students in the institution. Results of these studies give nearby into techniques for using the research process during that different process presaging student performance, compare the precision of data mining clustering algorithm, and demonstrate the implements. Clustering methods are divided into hierarchical and non-hierarchical methods according to the fragmentation technique of clusters. The success of both types of clustering methods varies according to the data set applied. In this study, both types of methods were tested on different type of data sets. Who take part in the educational process could gain by applying data mining on the data from the learning system. Since data mining represents the computational data process from contradictory perspectives,with the goal of excavating in- explicit information from the data, it can greatly avail every participant in the scholastic process in order to understanding the teaching process and it canters on discovering, detecting and expounding educational.The Simple K- means clustering, Hierarchical clustering, DBSCAN, clustering were grouped from related similarities, and analysis and Time Taken to build from

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 9, September 2019

different cluster mode was experimented performed by **WEKA TOOL**. Each data are significance to mention the particular information. to make the each segment to pre-processing on the prove to variety of results.

**What should think the student personality?** These are convention to the clustering concept with duration time to taken different type of moderate are concluded. Each category was building the conditioning format on the make the questionnaires conversions. It should also differentiate with the type of student performance and mindset monitoring level. Make the adaptation between the assortments of cluster are monitoring. Important for the emotional status of modern-day college students. College students are a special group, which come up to different regions, and mature in different environments with different cultures. But when they enter the College, facing the same living and working environment, it is very straightforward to make their environment change dramatically, and it is lying on your front to get negative psychological emotion, even nervousness and depression. If such negative emotions long for a time, it will concern the students' learning and life worth. Therefore, a study of the movement of college student's negative emotions and psychological problems of the mining model of the brain has become an important research topic of current.

## II. RELATED WORK

In this dissertation a survey on student mentality mention to recover the our mind set of thinking was easy to make the questionnaire process was each filled up the overview concept on the categorical level on the ration button, a large data set was build to make the pre-processing steps to condition are satisfied the answer on personality about the student think, the it should provide the each category are single group of portioning method was during to make the clustering method. A result give the problem on best solution for variety of category on the level by level technique are evaluated. Finally find the time taken on variety of clustering using to make the assign for data mining techniques as follows. Such as student motivation level, learning techniques used, strategy for preparing for student behaviour and habits, questions chosen for answering and way of expressing ideas may also be responsible for the final mindset on student.

- 1) The importance of the curriculum in improving and colleges form difficult courses of education and training has been increasing a top high level colleges approaches regarding the effectiveness of the components of the curriculum evaluation to make informed
- 2) Decisions regarding the outcomes and future actions relate to reserve,
- 3) Ask questions,
- 4) Welcome challenges and persist in meeting them,
- 5) Make mistakes,
- 6) Be an active and refresh,
- 7) Be inquisitive and seek new knowledge.

## III. DATA MINING TECHNIQUE USING WEKA TOOL

Most important to the quality of result the data Pre- processing, Clustering Technique was important. In best quality of data. By the processing the data, data quality can be measures in terms of accuracy, completeness, consistency, timeliness, believability, interpretability.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 9, September 2019

## WEKA TOOL

Now a days, weka is recognized as a landmark system in data mining and machine learning. it has achieved widespread acceptance with in academia and business circles and has become a widely used tool for data mining research. WEKA TOOL is a set of machine learning algorithm that can be applied to a Large Data set directly. The machine learning method is similar to data mining, its specified to find the patterns, adjust the program's action the different data mining system extract the data. A weka tool give the valuable information from large volumes of data.

## IV. PROPOSED WORK

This dissertation proposed a frame work to huge data sources computed to each based on the *student what should ever think?* The data was since collected from my college (**Dr.Ambedkar Government Arts College, vyasarpadi, Chennai-600 039**) and the information was gaining from **Different department and various courses**. There are multiple frameworks to handle the missing value of the noisy data should be measured make the pre-processing performs the variety of steps to improve the KDD. While the data are making the problems can deal with single data sources at instance level. A detecting level make the quality of data in applied the multiple data source of the different department and dissimilar to student performance.

### A. DATA PRE-PROCESSING

A full data set was unclear position. So filter to use the step on weka tool for pre processing. That was easy to build the training data in incomplete and noisy. Incomplete in lacking attributes values and lacking attribute of containing errors a incorrectly data was basically wrong but possible to data mining in weka tool for technique in pre-processing major issue are common consideration with different analysis

### B. CLUSTERING

Clustering is a process of dividing the data elements into groups which are similar to Each other. Each group is referred to as a cluster that consists of objects that are similar to one another and dissimilar to objects of another group. It is a technique that recognizes different patterns of data. Good clustering techniques will produce a good or a high quality cluster. It helps to analyse large volumes of data visually thus assists in making quick decisions. Clustering groups data instances into subsets in such a manner that similar instances are grouped together, while different instances belong to different groups and the groups are called as clusters.

#### 1) *Partition Based Clustering:*

All objects are considered initially single cluster. The objects are divided into n of partitions by iteratively locating the points between the partitions. The partitioning algorithms like K-means, K-medoids (PAM, CLARA, CLARANS, and FCM) and K- modes.Partition based algorithms can found clusters of Non-convexshapes.

#### 2) *Hierarchical-Based Clustering:*

There are two techniques of performing hierarchical based clustering Agglomerative (top-bottom) and Divisive (Bottom-up). In Agglomerative approach, initially one of object is selected and successively merges the neighbour object is selected and then to Successively merges the neighbour objects based on the distance as minimum, and the Maximum and average.

#### 3) *Density-Based Clustering:*

Data objects are categorized into core points, border points and noise points. All the core points are connected Together based on the densities to form cluster. Arbitrary shaped clusters are formed by various clustering Algorithms. Such as DBSCAN, OPTICS, DBCLASD, GDBSCAN, DENCLU and SUBCLU.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 9, September 2019

4) *Farthest First Clustering:*

Farthest places first is a variant of K means that places each cluster centre in turn At the point further from the existing cluster centres. This point must lie within the data area. This is greatly sped up the clustering in the most cases since less Reassignment and adjustment is needed.

5) *Simple EM –Expectation Maximisation (Class):*

EM assigns a probability distribution to each instance which indicates the probability of it belonging to each of the clusters. EM can decide how many clusters to create by crossvalidation, or you may specify A priori how many clusters to generate.

## V. EXPERIMENTAL RESULTS

484 students and 11 departments completed questionnaires both in weeks so our analysis of the effect on mindset attitude is based on this subset. Clustering model among identifier models provides to separate groups according to their calculated similarities by taking into consideration specific characteristics of Data First aim of usage of distance methods is to obtain similarity according to distance between data which is not grouped, which is present in this group.To evaluate the efficiency of the different clustering algorithms, some tests are realized.

Data set	Process Time (in sec.)				
	K mean Clustering	Farthest Cluster	DBSCAN Cluster	EM Cluster	Hierarchical Cluster
124	0.05 sec	0.01 sec	0.09 sec	0.07sec	0.011sec
139	0.00067 sec	0.00078 sec	0.000489sec	0.000567 sec	0.000789 sec
202	0.09872sec	0.00070 sec	0.000645 sec	0.000643 sec	0.000692 sec
239	0.00987 sec	0.00678 sec	0.000456 sec	0.000457 sec	0.000767 sec
332	0.0999 sec	0.00698 sec	0.000777 sec	0.000992 sec	0.006787 sec
407	0.0067sec	0.00567 sec	0.006782 sec	0.000567 sec	0.096787 sec

**Table 1: Time taken to form the respective number of clusters**

While performing the tests, six different algorithms were used. How to performing the cluster in time taken to form the respective number of cluster and processing on sum of the squared error mentioned the table 1, and Analysis of algorithms was made on nine different data sets and mentioned the table 2.

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 9, September 2019

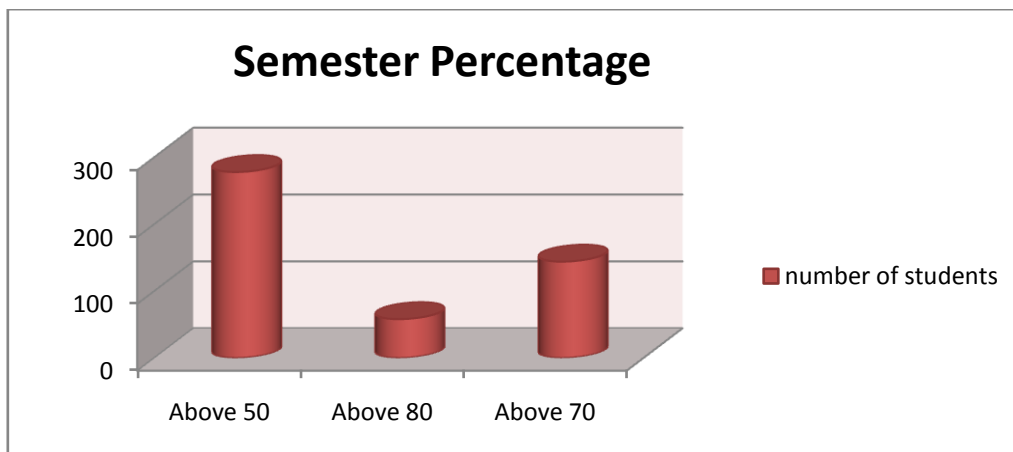
Data set	# Cluster Instance	K mean Clustering	Farthest Cluster	DBSCAN Cluster	EM Cluster
You study at College for the sake	10	0.1651	0	2189.9	1647
Attendance in class room	7	0.2861	2241	1874	2854
Semester percentage	4	4556700	789690	876500	876
Parent visit to college	7	87765	0.876	9543578	907
Ridiculed By Fear	2	9876	675438	5467	987654
College for important	6	78976	436700	98765	657489
Unsure about the future dream	12	0.9876	0.333	987	0
Which term of study is right for you	6	0.6798	0.218	97	7654
Distance between college and home	8	0.7765	765490	984	97654

*Table 2: Sum of Squared Errors of Clustering Process*

### EXAMPLE FOR CLUSTERING GRAPHICAL VIEW

*Student's performance in Semester Percentage:*

Figures show the results of, A Student's Semester Percentage Figs.(a) Data was Pre-Processing view and using Data mining technique (b) Graphical clustering point in weka tool.



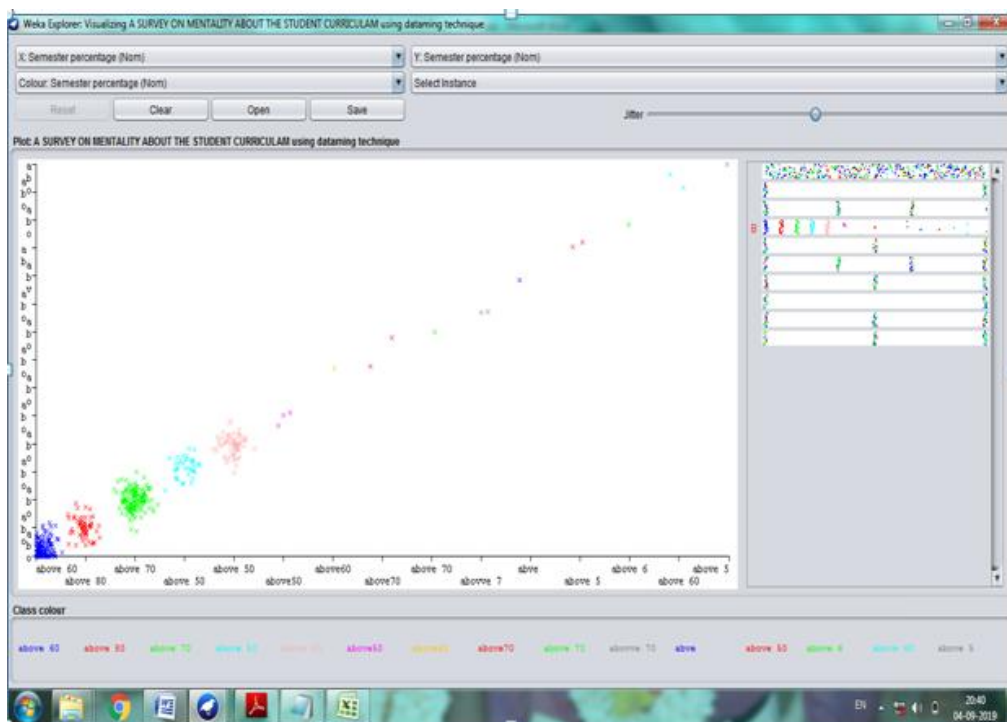
(a)

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 9, September 2019



(b)

## VI. CONCLUSION

In this study, different clustering algorithms were compared in terms of processing on Time, sum of squared errors of Different datasets. As a result of this study, It is who Under stood that which category of Student Affect in study. And then another part to improve your Knowledge Study. Algorithm produces more accurate results rather than Differentiate between Two type ways of clustering performed through the result from (**Hierarchical or Non – Hierarchical**). A performance was Assign the group for a Reasoning mainly Study importance. One Group of students in one way part, different Way. (**Interested to Study or not interest to study**).

## REFERENCES

- [1] Danger of a fixed mindset : implications of self-theories research for computer science education Laurie Murphy, Lynda Thomas ACM SIGCSE Bulletin 40 (30), 271-275, 2008.
- [2] Anoopkumar M, Educational Data Mining to Predict Student concept A Review on Data Mining Techniques and Factors Used .
- [3] FidanKaya Gülagiz1 and SuhapSahin Comparison of Hierarchical and Non-Hierarchical Clustering Algorithms.
- [4] Developing and assessing students' entrepreneurial skills and mind-set Sven G Bilan, Elliabeth C Kisenwether , Sarath E Rasa, John C Wise, Journal of Engineering Education 94(2), 233-243,2005.
- [5] Survey of clustering algorithm RuiXu, Donald C Wunch, Intitue of Electrical and Electronic Engineers (IEEE), 2005.
- [6] Developing a positive mind-set toward the use of technology for classroom instruction, Mabel CPO Okojie, Antony olinock, International Journal of instruction Media 33(1), 33-42, 2006.
- [7] Azwa Abdul Aziz, Nor Hafieza Ismail and Fadhilah Ahmad, "First Semester Computer Science Students' Academic Performances Analysis by Using Data Mining Classification Algorithms", International Conference on Artificial Intelligence and Computer Science (AICS 2014), September 2014

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 9, September 2019

- [8] Hierarchical clustering schemes , Stephen C Johnson, Psychometrika 32(3), 241-254, 1967.
- [9] Life satisfaction, happiness, and the growth mindset of healthy perfectionists among Hong Kong Chinese gifted students, David W Chan , Roeper Review 34(4), 224-233, 2012
- [10] Student ‘s Examination Result Mining: A predictive Approach, Vaneet Kumar, Vinod Sharma, International Journal of advanced computer Science and Applications(IJSER) 3 (11), 2012.
- [11] Datamining in educational system using weka, Sunitha B Aher, LMRJ Lobo, International Conference on Emerging Technology Trends (ICETT) 3, 20-225, 2011.
- [12] Cristobel Romero, Sebastian Ventura "educational data mining: a review of the state of the art" iee transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 40, no. 6, november2010 .
- [13] Effectiveness of school based universal social, emotional, and behavioural programs: Do they enhance student’s development in the area of skill behaviour, and adjustment? MarchinSkład, Rene Diekstra, Monique De Ritte, Jehonathan Ben, CarolienGravesteijn, Psychology in the schools 49(9), 892-909, 2012.
- [14] The state of educational data mining in 2009: A review and future vision, Ryan SJD Baker, KlinaYacef, JEDMI Journal of Educational Data Mining1(1), 3-17, 2009.
- [15] A Data mining and survey study on diseases associated with paraesophageal hernia, Jianji Yang, Judith Logan, AMIA Annual Symposium Proceedings 2006, 826, 2006.
- [16] AshishDutt, SaeedAghabozrgi, MaizatulAkmalBinti I smail, and HamidrezaMahroeian, "Clustering Algorithms Applied in Educational Data Mining", International Journal of Information and Electronics Engineering, Vol. 5, No. 2, March 2015
- M.Jayakameswaraiah, S.Ramakrishna, "A Study on Prediction Performance of Some Data Mining Algorithms", International Journal of Advance Research in Computer Science and Management Studies, ISSN: 2321-7782 (Online) Volume 2, Issue 10, October 2014
- [17] Azwaabdulaziz, Nurhafiezaismail, Fadhilahahmad, "mining students’ academic performance", Journal of Theoretical and Applied Information Technology ISSN: 1992-8645,E-ISSN: 1817-3195, Vol. 53 No.3, July 2013
- [18] A.DineshKumar ,Dr.V.Radhika, "A Survey on Predicting Student Performance", International Journal of Computer Science and Canada, 2008, pp. 182–185. [www.educationaldatamining.org](http://www.educationaldatamining.org).
- [19] Romero, C. & Ventura, S, Educational Data Mining: a Survey from 1995 to 2005, Expert Systems with Applications, Elsevier, pp. 135-146. (2007).
- [20] SrečkoNatek, MotiZwilling. Student data mining solution–knowledge management system related to higher education institutions. Expert Systems with Applications, 2014, 41:6400–6407