International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 14
November 2017

# A Bigdata Approach For Classification And Prediction Of Student Result Using R

Akhila & Veer Ankulu
[1]M. Tech Student, Department of CSE, MallaReddy Engineering College For Women's, Maisammaguda, Dhulapally, RangaReddy, Telangana, India.

[2]Assistant Professor, Department of CSE, MallaReddy Engineering College For Women's, Maisammaguda, Dhulapally, RangaReddy, Telangana, India.

**Abstract—** *As of late the measure of information put away in instructive database is developing quickly. The put away database contains concealed data which if utilized guides change of understudy's execution and conduct. In this paper insightful exhibiting approach is used for removing this covered information. Information is gathered, a prescient model is figured, expectations are made, and the model is approved as extra information winds up noticeably accessible. The prescient models will help the teacher to see how well or how ineffectively the understudies in his/her class will perform, and consequently the educator can pick genuine instructive and instructional interventions to redesign understudy learning comes about. The usage is done in Hadoop system with MapReduce and Revolutionary R Enterprise RRE*

**Keywords**—big data analytics; educational data mining; predictive modelling; learning analytics; Hadoop; Horton works sandbox; revolution r enterprise; clustering; regression

## I. INTRODUCTION

Instruction part has a considerable measure of information as understudy data. Utilization of PCs on this information can extricate important data to give quality training. Because of this blend of instruction and PC (information mining), another exploration group is developing i.e. Instructive Data Mining (EDM). Instructive information mining is worried about mounting strategies to decide realities from information, particularly obscure learning driven example from instructive store with a specific end goal to feature the quality and shortcomings of the understudy

An instructive foundation needs a rough earlier information of enlisted. This causes them to distinguish promising understudies and furthermore gives them a chance to focus on and enhance the individuals who might presumably get bring down evaluations. There is no supreme scale for measuring information yet examination score is one scale which demonstrates the execution pointer of understudies. Quality instruction is a standout amongst the most encouraging obligations of any nation to his comrades. Quality training does not mean abnormal state of learning created. In any case, it implies that instruction is created to understudies in effective way so they learn with no issue. For this reason quality training incorporates highlights like: technique of instructing, ceaseless assessment, classification of understudy into comparable sort, with the goal that understudies have comparable destinations, instructive foundation and so on.

In this paper class X understudy's information of Central Board of Secondary Education in all finished India is considered for input dataset. Since it contains petabytes of understudy information, these datasets are considered as large information. In huge information idea the conventional information mining calculations are meant Map Reduce calculations for running them on Hadoop bunches by deciphering their information examination rationale to the Map Reduce work which is to be keep running over Hadoop groups. Hadoop bunches are intended for putting away and handling gigantic measure of information in an appropriated registering condition.. MapReduce programming model comprises of two independent and unmistakable undertakings that Hadoop programs perform. The first is the guide work, which takes an arrangement of information and believers it into another arrangement of information, where singular components are separated into tuples (key/esteem sets). The lessen work takes the yield from a guide as information and joins those information tuples into a littler arrangement of tuples.

The bunching technique is utilized to recognize scholastically atrisk understudies and sort the understudies appropriately. Since there are numerous calculations for information grouping, the K-Means technique is utilized here. The various relapse calculation is utilized for anticipating understudy comes about. The two calculations are made an interpretation of to MapReduce calculations to keep running on hadoop groups. In Kmeans bunching it comprises of 2 sections Map and Reduce. The guide work plays out the methodology of doling out each example to the nearest focus while the decrease work plays out the technique of refreshing the new focuses. Keeping in mind the end goal to diminish the cost of system correspondence, a combiner work is created to manage halfway blend of the middle of the road esteems with a similar key inside a similar guide errand. The numerous straight relapse mapreduce calculation is connected on understudy past outcome information to construct a prescient model. This model would then be able to foresee the understudy last outcome as review or checks. This will be useful for educator, understudy and their folks to know ahead of time about understudy last anticipated outcome and will empower them to take preventive measure.

## II. RELATED WORKS

There are numerous organizations that have made utilization of learning investigation to enhance understudy achievement and maintenance. Ourania Petropoulou, Katerina Kasimatis, Ioannis Dimopoulos, and Symeon Retalis, [6] composed LAe-R: another learning examination apparatus in Moodle for evaluating understudies' execution. A testing and requesting errand for the instructors in elearning situations is the appraisal of understudies' execution. A few learning administration frameworks (LMS) like Moodle offer a few evaluation instruments, for example, tests, scales, "exemplary" rubrics, and so forth. Deshpande[8] propose the utilization of Hadoop Framework and the ET-L process for Hadoop for performing forecasts in view of the datasets.

Hurn [5] characterize learning examination, how it has been utilized as a part of instructive establishments, what learning investigation devices are accessible, and how staff can make utilization of information in their courses to screen and anticipate understudy execution. They likewise give points of interest of a few issues and worries with the utilization of learning investigation in advanced education. Weizhong Zhao[7] composed Parallel KMeans Clustering Based on MapReduce for bunching , Data grouping has been gotten impressive consideration in numerous applications, for example, information mining, record recovery, picture division and example characterization. Learner-Centered Approach to Learning Analytics , thought is to take the demonstrated innovation based answer for tending to the maintenance and accomplishment of in danger understudies (the Online Student Profile framework created in CPCC's 2003-08 ) and work with accomplice schools to convey both the OSP and the related workforce and staff advancement exercises so as to enhance maintenance.

Learning investigation (LA) is a multi-disciplinary field including machine learning, counterfeit consciousness, data recovery, measurements, and perception. LA is additionally a field in which a few related zones of research in TEL join. These incorporate scholarly investigation, activity look into, instructive information mining, recommender frameworks, and customized versatile learning. M.A. Thüs audit late productions on LA and its related fields and guide them to the four measurements of the reference demonstrate. Besides, we distinguish different difficulties and research openings in the zone of LA in connection to each measurement. Kenneth Wottrich[7] propose an examination in 2010 to portray and show the execution of MapReduce applications on ordinary, adaptable groups in view of central application information and preparing measurements. He recognized five basic qualities which characterize the execution of MapReduce applications. At that point he made five separate seat check tests, each intended to seclude and test a solitary trademark. The after effects of these benchmarks are useful in developing a model for MapReduce applications.

There are three basic arranging issues in MapReduce, for instance, zone, synchronization and tolerability. The most widely recognized goal of planning calculations is to limit the consummation time of a parallel application and furthermore accomplish to these issues. There are numerous calculations to comprehend this issue with various procedures and methodologies. Some of them get center to change information region and some of them executes to give Synchronization handling.

## III. PROPOSED SYSTEM

The primary objective of this paper is to recognize scholastically atrisk understudies and to build up a prescient model to anticipate understudy scholarly execution in instructive organizations, which predicts their future outcomes. Understudy scholastic execution is influenced by various elements. The extent of this examination is restricted to the examination of learning movement on their scholastic execution. The proposed framework comprises of two functionalities:

a)Identifying scholastically in danger understudies

b)Prediction of understudy result

A. Distinguishing scholastically in danger understudies

The information gathered from various applications require legitimate technique for separating learning from extensive storehouses for better basic leadership. This makes an extraordinary test for establishments utilizing conventional information administration component to store and process colossal datasets. So it is required to characterize another worldview called "Huge Data Analytics" to re-assess current framework and to oversee and process enormous information. We actualize a part of Big Data Analytics known as "Learning Analytics". Learning examination ( LA) alludes to the elucidation of an extensive variety of information delivered by and assembled for understudies keeping in mind the end goal to evaluate scholarly advance, foresee future execution, and spot potential issues. Fig.1 demonstrates the means for recognizing scholastically in danger understudies utilizing LA.

The initial phase in any LA exertion is to gather information from different instructive situations and frameworks. This progression is basic to the fruitful disclosure of helpful examples from the information. The dataset is gathered from different CBSE schools in all finished India and is accessible in SQL design in MySQL Server. Since it contains petabytes of understudy information, these datasets are considered as Big Data. Hadoop structure can be utilized for savvy and speedier huge information preparing, which would improve the examining process.Hadoop is an open-source programming system for putting away information and running applications on groups of product hardware,it gives monstrous capacity to information.

Guide Reduce is a system for composing applications that procedure a lot of organized and unstructured information in parallel over a group of thousands of machines, in a dependable, blame tolerant way. Correspondingly HDFS is a document framework that gives dependable information stockpiling and access over every one of the hubs in a Hadoop group. It interfaces together the document frameworks on numerous nearby hubs to make a solitary record system.The datasets should be arranged and transferred to Hadoop Distributed File System (HDFS) and utilized further by different hubs with Mappers and Reducers in Hadoop bunches. The datasets are transferred to Hortonworks Data Platform (HDP) for investigation, utilizing an instrument SQOOP (hadoop reverberate framework extend). It is utilized to import information from social databases, for example, MySQL, Oracle to Hadoop HDFS, and fare from Hadoop document framework to social databases

Information pre-handling additionally permits changing the information into a reasonable arrangement that can be utilized as contribution for a specific LA technique. A few information pre-handling assignments, obtained from the information mining field, can be utilized as a part of this progression. These incorporate information cleaning, information combination, information change, information diminishment, information displaying, client and session distinguishing proof, and way consummation. The datasets are overseen and pre-prepared by Apache Hive. Hive gives a distribution center structure and SQL like access for information in HDFS and other hadoop input sources. The information in required configuration is accessible in HDFS by the utilization of Sqoop. These information is cleaned and incorporated by HiveQl dialect gave by Hive. Pre-handling is utilized to perform information operation to make an interpretation of information into a settled information design before giving information to calculations or instruments. The information investigation process will then be started with this organized information as the info.

After information is accessible in the required organization for information examination calculations, information investigation operations will be performed. The information examination operations are performed

for finding important data from information to take better choices towards execution with information mining ideas. It might either utilize expressive or prescient examination for understudy's execution assessments. Investigation can be performed with different machine learning and also custom algorithmic ideas, for example, relapse, order, grouping, and model-based proposal. For Big Data, similar calculations can be meant MapReduce calculations for running them on Hadoop groups by deciphering their information investigation rationale to the MapReduce work which is to be keep running over Hadoop bunches. MapReduce is the handling system for Apache Hadoop. MapReduce enables programming to take care of information parallel issues for which informational collection can be sub-isolated into little parts and handled freely. The framework parts the info information into different lumps, each of which is alloted a guide undertaking that can procedure the information in parallel. Each guide assignment peruses the contribution as an arrangement of (key, esteem) matches and creates a changed arrangement of (key, esteem) combines as the yield. The system rearranges and sorts yields of the guide undertakings, sending the moderate (key, esteem) sets to diminish errand, which bunches them into definite outcomes.

**Parallel K-Means Algorithm Based on MapReduce**

The info dataset is put away on HDFS as a grouping document of <key, value> sets, each of which speaks to a record in the dataset. The key is the balanced in bytes of this record to the begin purpose of the information document, and the esteem is a string of the substance of this record. The dataset is part and all around communicate to all mappers. Thusly, the separation calculations are parallel executed. For each guide undertaking, PK-Means develop a worldwide variation focuses which is an exhibit containing the data about focuses of the groups. Given the data, a mapper can figure the nearest focus point for each specimen.Algorithm MAP (key,value)

The last phase of the procedure comprises of perception of the aftereffects of information examination. Perception is an intelligent approach to speak to the information bits of knowledge. This should be possible with different information representation programming resembles Gephi and so forth. Gephi is an open-source arrange examination and perception programming bundle written in Java on the NetBeans stage, at first created by understudies of the University of Technology of Compiègne (UTC) in France.

The came about yield of parallel K-Means grouping calculation in hadoop is in twofold configuration. To comprehend the outcome in comprehensible configuration we have to change over the twofold arrangement into .txt or GraphML, for this we utilize Clusterdump instrument in Mahout. Mahout is the datamining library of Apache Hadoop. The came about yield in GraphML can specifically open in Gephi and it will come about the factual examination. By utilizing calculations, for example, measured quality, Fruchterman Reingold we can examine the outcome and parcel them into bunches in light of the aftereffect of grouping calculation

B. Expectation of understudy result

The forecast of understudies' outcome is essential for instructive organizations, in light of the fact that the nature of showing process is the capacity to address understudies' issues. Breaking down the past execution of these understudies would give a superior point of view of the plausible scholastic execution of understudies later on. This can possibly be accomplished utilizing the idea of Predictive Analytics. Prescient investigation incorporates an assortment of measurable procedures from displaying, machine learning, and information mining that examine present and chronicled realities to make forecasts about future

The above figure demonstrates the means for anticipating understudy comes about. The initial step is to gather the dataset for forecast. The dataset is partitioned into two sets – one for preparing information (preparing set) and other for test information (test set). The insights is done on the preparation dataset and a prescient model is produced utilizing the preparation dataset. The preparation and test datasets are transferred to Hortonworks Data Platform (HDP) for investigation, utilizing an apparatus called SQOOP. It is utilized to import information from social databases, for example, MySQL, Oracle to Hadoop HDFS, and fare from Hadoop document framework to social databases

Information Analysis stage is done in Hortonworks Sandbox. This stage incorporates information cleaning, information designing, information subsetting and so

forth. The Hortonworks Sandbox is a solitary hub execution of the Hortonworks Data Platform (HDP). In this stage just the suitable factors for understudy's execution assessment are separated from the info dataset utilizing the Map Reduce worldview. The hortonworks sandbox gives a segment called Hive, which is utilized for extricating important information from dataset. The Hive bolsters an inquiry organize, HiveQL which is same as that of SQL, yet preparing is done in view of guide lessen programming model.

The last three stages are actualized in RRE. Upheaval R Enterprise for Windows is an improved, upheld adaptation of the open source R dialect. It incorporates RevoScaleR, Revolution's bundle for factual investigation of extensive informational collections. RevoScaleR gives capacities to performing adaptable and to a great degree elite information administration, investigation, and representation. The information control and investigation works in RevoScaleR are proper for little and expansive datasets, yet are especially helpful in three normal circumstances: 1) to dissect informational indexes that are too huge to fit in memory and, 2) to perform calculations appropriated more than a few centers, processors, or hubs in a bunch, or 3) to make versatile information examination schedules that can be created locally with littler informational collections. RevoScaleR gives another information record sort with expansion .xdf that has been streamlined for "information lumping", getting to parts of a Xdf petition for free handling. Xdf records store information in a paired arrangement. Techniques for getting to these documents may utilize either level (lines) or vertical (sections) square dividing. The record organize gives quick access to a predefined set of lines for a predetermined arrangement of segments.

Different Linear Regression calculation is utilized to recognize connection amongst needy and free factors in preparing dataset. Once the connections amongst reliant and autonomous factors are discovered, at that point a direct relapse display is made. The model is of the frame

$$Y = \beta0 + \beta1X1 + \beta2X2 + \beta3X3 + .....$$

Where Y is the predicted value, β0, β1, β2, β3 are regression coefficients and X1, X2, X3 are dependent variables.

Once the model is made, test dataset is just expected to apply in this direct relapse model to produce anticipated esteem. In this venture we have taken four test marks (GP1, GP2, GP3 and GP4) of every understudy as preparing dataset, they are utilized to produce direct relapse display condition. GP4 is taken as autonomous variable and GP1, GP2 and GP3 as needy variable. Utilizing various straight relapse calculation discovered connection between GP4 against GP1, GP2 and GP3. Utilizing this relationship another prescient model is made. The test dataset contains just three imprints GP1, GP2 and GP3. Applying these needy factors to the relapse condition to discover the anticipated esteem GP4.

**NETWORKD3**

NetworkD3 works exceptionally well with the latest rendition of RStudio (>=v0.99, download). When you utilize this form of RStudio, diagrams will show up in the Viewer Pane. Not exclusively does this give you a convenient method for seeing and tweaking your charts, yet you can likewise send out the diagrams to the clipboard or a PNG/JPEG/TIFF/and so on document. The bundle can be downloaded from CRAN.

## IV. EXPERIMENTAL RESULTS

The clustered output of the training dataset is given in the fig. 3. It consists of 6 clusters which contain all the dataset and partitioned them into different clusters.

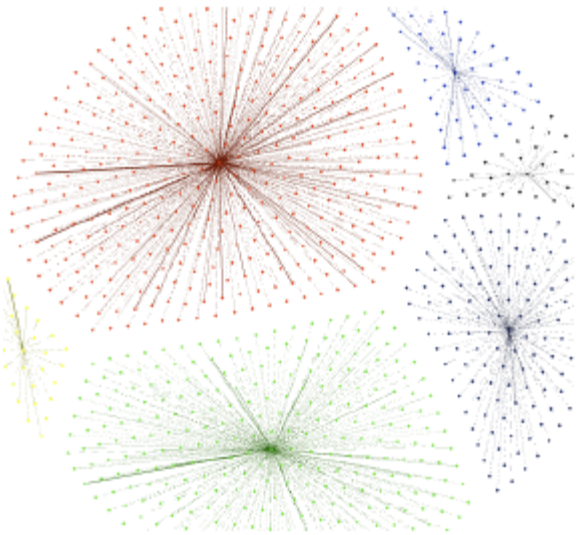

Fig 3. Clustered Output

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue 14
November 2017

| | SCH | RROLL | CNAME | SEX | GP1 | GP2 | GP3 | GP4 |
|---|---|---|---|---|---|---|---|---|
| 1 | 03001 | 1100878 | AAHASH | M | 10 | 10 | 10 | 10 |
| 2 | 03001 | 1100879 | PATEL AARSH PANKAJ KUMAR | F | 06 | 07 | 05 | 06 |
| 3 | 03001 | 1100880 | ABHISHEK PANDEY | M | 06 | 07 | 06 | 06 |
| 4 | 03001 | 1100881 | SHAIKH ATIYA NAIMUDDIN | F | 08 | 08 | 07 | 07 |
| 5 | 03001 | 1100882 | ADRESHA DAVE | F | 09 | 09 | 09 | 08 |
| 6 | 03001 | 1100883 | ALEESHA BABU | F | 08 | 06 | 06 | 05 |
| 7 | 03001 | 1100884 | AMAN AHANKAL | M | 10 | 10 | 10 | 09 |
| 8 | 03001 | 1100885 | AMINOOH BABU | M | 05 | 05 | 05 | 05 |
| 9 | 03001 | 1100886 | ANISH VERMA | M | 07 | 06 | 06 | 06 |
| 10 | 03001 | 1100887 | ANSHIKA BADGIE | F | 09 | 09 | 08 | 08 |
| 11 | 03001 | 1100888 | ARNAV PUROHIT | M | 08 | 09 | 10 | 08 |
| 12 | 03001 | 1100889 | SHAIKH BUSHRA KHATUN | F | 07 | 08 | 06 | 07 |
| 13 | 03001 | 1100890 | DEEP NARENDRA LUNKER | M | 07 | 06 | 06 | 06 |
| 14 | 03001 | 1100891 | DEEPIKA ROAT | F | 06 | 05 | 04 | 05 |
| 15 | 03001 | 1100892 | DHARMIK | M | 07 | 07 | 08 | 07 |
| 16 | 03001 | 1100893 | MADHIEN MANOJ DESAI | M | 06 | 06 | 08 | 05 |
| 17 | 03001 | 1100894 | HARSH MANHARIA | M | 05 | 05 | 08 | 07 |
| 18 | 03001 | 1100895 | HARSH PRASAD | M | 07 | 08 | 08 | 07 |
| 19 | 03001 | 1100896 | HARSHIL JAIN | M | 08 | 08 | 09 | 09 |
| 20 | 03001 | 1100897 | HARSHITA J MEENA | F | 06 | 07 | 05 | 06 |
| 21 | 03001 | 1100898 | HARSHITA SHARDA | F | 09 | 09 | 09 | 09 |
| 22 | 03001 | 1100899 | JUI ATUL KUMAR CHAWDA | F | 06 | 06 | 05 | 06 |
| 23 | 03001 | 1100900 | KALPAV RAJESH KUMAR AMIN | M | 07 | 06 | 06 | 06 |
| 24 | 03001 | 1100901 | GALI KALYAN SEMMAR | M | 09 | 09 | 10 | 09 |
| 25 | 03001 | 1100902 | KAPIL SEMRA | M | 05 | 05 | 05 | 05 |
| 26 | 03001 | 1100903 | MADHAV OJHA | M | 08 | 09 | 07 | 07 |
| 27 | 03001 | 1100904 | SHAIKH MISBAH I | F | 07 | 09 | 08 | 08 |
| 28 | 03001 | 1100905 | MAUMEETA JAYAPRASAD | F | 09 | 10 | 10 | 10 |
| 29 | 03001 | 1100906 | NITISH ROAT | M | 09 | 09 | 10 | 08 |
| 30 | 03001 | 1100908 | PANCHAL PARTH NAVIN | M | 09 | 09 | 10 | 10 |
| 31 | 03001 | 1100909 | PRACHI DHAVA | F | 09 | 09 | 09 | 09 |
| 32 | 03001 | 1100910 | PRASATI MISHAL | F | 07 | 05 | 06 | 06 |

Fig 4. Training dataset

The figure 4 shows input training dataset, using which the predictive model is built. The training dataset contains information like school code (sch), roll number (rroll) of each student, name (cname, mname, fname), sex and marks for 4 tests (gp1, gp2, gp3, gp4). In this training dataset the variable gp4 is taken as dependent variable and gp1, gp2 and gp3 are taken as independent variables. Around 150000 student records are taken in training dataset

| | SCH | RROLL | CNAME | SEX | GP1 | GP2 | GP3 |
|---|---|---|---|---|---|---|---|
| 1 | 03096 | 1105769 | JAYDEEP RAVAL | M | 07 | 09 | 06 |
| 2 | 03096 | 1105770 | KRUNALBHAI VASAVA | M | 10 | 10 | 10 |
| 3 | 03096 | 1105771 | KRUSHABEN CHHASATIA | F | 08 | 10 | 08 |
| 4 | 03096 | 1105772 | MAYURIBEN DARJI | F | 06 | 08 | 06 |
| 5 | 03096 | 1105773 | NIMUBHAI ROHIT | M | 10 | 10 | 09 |
| 6 | 03096 | 1105774 | PIYUSHBHAI CHAUDHARI | M | 07 | 09 | 07 |
| 7 | 03096 | 1105775 | PRATIK BARIYA | M | 06 | 09 | 06 |
| 8 | 03096 | 1105776 | RAHUL RATHAVA | M | 07 | 10 | 06 |
| 9 | 03096 | 1105777 | RAKESH RATHVA | M | 08 | 09 | 06 |
| 10 | 03096 | 1105778 | RUTANGI VASAVA | F | 08 | 08 | 08 |
| 11 | 03096 | 1105779 | SANDIPKUMAR RATHWA | M | 07 | 09 | 07 |
| 12 | 03096 | 1105780 | SHREYABEN PATEL | F | 08 | 10 | 08 |
| 13 | 03096 | 1105781 | SHREYABEN PATEL | F | 07 | 08 | 08 |
| 14 | 03096 | 1105782 | SHREYASKUMAR CHAUDHARI | M | 08 | 10 | 08 |
| 15 | 03096 | 1105783 | SUNILKUMAR VANKAR | M | 09 | 10 | 09 |
| 16 | 03096 | 1105784 | SUNILKUMAR RATHVA | M | 08 | 10 | 09 |
| 17 | 03096 | 1105785 | SURBHI DAGAR | F | 10 | 10 | 10 |
| 18 | 03096 | 1105786 | VANDANABEN RATHWA | F | 07 | 09 | 08 |
| 19 | 03096 | 1105787 | VEENABAHEN PATEL | F | 08 | 10 | 09 |
| 20 | 03096 | 1105788 | MUKULKUMAR CHANDAR | M | 10 | 10 | 09 |
| 21 | 03096 | 1105789 | UTKARSHKUMAR PATEL | M | 07 | 10 | 07 |
| 22 | 03096 | 1105790 | KRISHNPAL GAJJAR | M | 10 | 10 | 09 |
| 23 | 03096 | 1105791 | UMESHKUMAR PARMAR | M | 08 | 10 | 07 |
| 24 | 03096 | 1105792 | SAMILKUMAR VANKAR | M | 09 | 10 | 09 |
| 25 | 03096 | 1105793 | ARJUNBHAI RATHWA | M | 08 | 10 | 08 |
| 26 | 03096 | 1105794 | VANSHIT ROHIT | M | 10 | 10 | 09 |
| 27 | 03096 | 1105795 | RIDDHIBEN AMIN | F | 07 | 10 | 07 |
| 28 | 03096 | 1105796 | JAGRUTIBEN PARMAR | F | 07 | 10 | 07 |
| 29 | 03096 | 1105797 | MAHESHBHAI AHIR | M | 08 | 10 | 08 |
| 30 | 03096 | 1105798 | SWETABEN RATHWA | F | 08 | 10 | 08 |
| 31 | 03096 | 1105799 | JANVI PARMAR | F | 09 | 10 | 09 |
| 32 | 03097 | 1106018 | AADIL SHAUKATALI ANKUSI | M | 06 | 07 | 06 |

Fig 5. Test dataset

Figure 5 shows the test dataset, a new dataset contains all the columns in training dataset except one column, i.e.; mark for one test (gp4) in training dataset. The gp4 for test dataset is predicted from the statistical relationship found in training dataset.

| | SCH | RROLL | CNAME | SEX | GP1 | GP2 | GP3 | GP4 |
|---|---|---|---|---|---|---|---|---|
| 1 | 03096 | 1105769 | JAYDEEP RAVAL | M | 07 | 09 | 06 | 6.63 |
| 2 | 03096 | 1105770 | KRUNALBHAI VASAVA | M | 10 | 10 | 10 | 9.98 |
| 3 | 03096 | 1105771 | KRUSHABEN CHHASATIA | F | 08 | 10 | 08 | 8.15 |
| 4 | 03096 | 1105772 | MAYURIBEN DARJI | F | 06 | 08 | 06 | 6.04 |
| 5 | 03096 | 1105773 | NIMUBHAI ROHIT | M | 10 | 10 | 09 | 9.51 |
| 6 | 03096 | 1105774 | PIYUSHBHAI CHAUDHARI | M | 07 | 09 | 07 | 7.59 |
| 7 | 03096 | 1105775 | PRATIK BARIYA | M | 06 | 09 | 06 | 6.15 |
| 8 | 03096 | 1105776 | RAHUL RATHAVA | M | 07 | 10 | 06 | 6.77 |
| 9 | 03096 | 1105777 | RAKESH RATHVA | M | 08 | 09 | 06 | 6.15 |
| 10 | 03096 | 1105778 | RUTANGI VASAVA | F | 08 | 08 | 08 | 8.01 |
| 11 | 03096 | 1105779 | SANDIPKUMAR RATHWA | M | 07 | 09 | 07 | 7.09 |
| 12 | 03096 | 1105780 | SHREYABEN PATEL | F | 08 | 10 | 08 | 8.06 |
| 13 | 03096 | 1105781 | SHREYABEN PATEL | F | 07 | 08 | 08 | 7.56 |
| 14 | 03096 | 1105782 | SHREYASKUMAR CHAUDHARI | M | 08 | 10 | 08 | 8.15 |
| 15 | 03096 | 1105783 | SUNILKUMAR VANKAR | M | 09 | 10 | 09 | 9.06 |
| 16 | 03096 | 1105784 | SUNILKUMAR RATHVA | M | 08 | 10 | 09 | 9.06 |
| 17 | 03096 | 1105785 | SURBHI DAGAR | F | 10 | 10 | 10 | 9.98 |
| 18 | 03096 | 1105786 | VANDANABEN RATHWA | F | 07 | 09 | 08 | 6.63 |
| 19 | 03096 | 1105787 | VEENABAHEN PATEL | F | 08 | 10 | 09 | 9.06 |
| 20 | 03096 | 1105788 | MUKULKUMAR CHANDAR | M | 10 | 10 | 09 | 9.51 |
| 21 | 03096 | 1105789 | UTKARSHKUMAR PATEL | M | 07 | 10 | 07 | 7.23 |
| 22 | 03096 | 1105790 | KRISHNPAL GAJJAR | M | 10 | 10 | 09 | 9.51 |
| 23 | 03096 | 1105791 | UMESHKUMAR PARMAR | M | 08 | 10 | 07 | 7.65 |
| 24 | 03096 | 1105792 | SAMILKUMAR VANKAR | M | 09 | 10 | 09 | 9.06 |
| 25 | 03096 | 1105793 | ARJUNBHAI RATHWA | M | 08 | 10 | 08 | 8.15 |
| 26 | 03096 | 1105794 | VANSHIT ROHIT | M | 10 | 10 | 09 | 9.51 |
| 27 | 03096 | 1105795 | RIDDHIBEN AMIN | F | 07 | 10 | 07 | 7.23 |
| 28 | 03096 | 1105796 | JAGRUTIBEN PARMAR | F | 07 | 10 | 07 | 7.23 |
| 29 | 03096 | 1105797 | MAHESHBHAI AHIR | M | 08 | 10 | 08 | 8.15 |
| 30 | 03096 | 1105798 | SWETABEN RATHWA | F | 08 | 10 | 08 | 8.15 |
| 31 | 03096 | 1105799 | JANVI PARMAR | F | 09 | 10 | 09 | 9.06 |
| 32 | 03097 | 1106018 | AADIL SHAUKATALI ANKUSI | M | 06 | 07 | 06 | 6.43 |

Fig 6. Predicted Result

Figure 6 shows the final predicted result after applying test dataset in predictive model developed using training dataset.

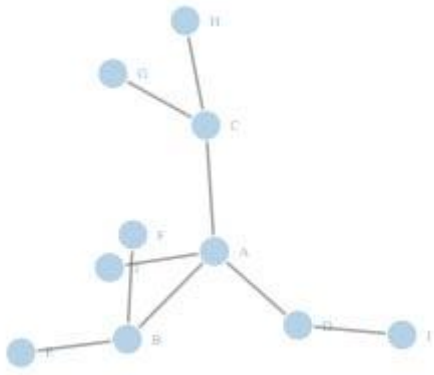

Fig 7. Nearest Values from the data

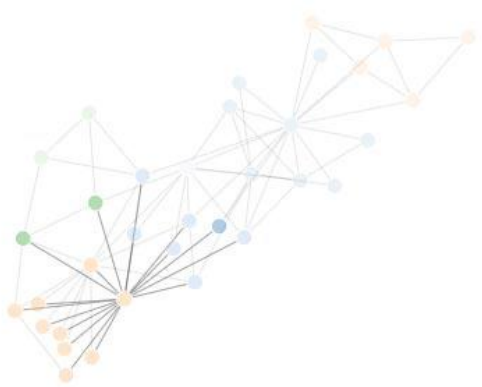

Fig 8. Nearest Values from the data

Figure 8 shows the final predicted result after applying test dataset in predictive model developed using network d3.

## V. CONCLUSION

In this paper we have displayed another approach called Learning Analytics and Predictive examination to distinguish scholastically in danger understudies and to anticipate understudies learning results in instructive foundations. The prescient models will help the teacher to see how well or how inadequately the understudies in his/her class will perform, and henceforth the educator can pick. It likewise encourages educators to foresee about understudies achievement and disappointment in examination and furthermore they can give appropriate advices to avoid disappointment in the examinations..

## VI. REFERENCES

*[1] Clint McElroy,The Online Student Profile Learning System: a Data Mining problems. Soft comput, 10.1007/s00500-008-0323*

*[2] Alcala, J., Sanchaz, L., Garcia, S., Del Jesus, M. ct. (2007). KEEL :A software tool to assess Evolutionary Algorithms to Learner-Centered Approach to Learning Analytics, April 2011.*

*[3] George Lepouras, Akrivi Katifori, Costas Vassilakis, Angeliki Antoniou, Nikos Platis, Towards a Learning Analytics Platform for Toolkit for Teachers. Educational Technology & Society 15, 3, 58– 76.*

*[4] Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., and Schroeder, U. 2012. Design and Implementation of a Learning Analytics Supporting the Educational Process.*

*[5] Bulletin of the IEEE Technical Committee on Learning Technology – Special Issue on "State-of-the-Art in TEL"*

*[6] M.A. Chatti, A.L. Dyckhoff, U. Schroeder, and H. Thüs, A Reference Model for Learning Analytics, International Journal of Technology Enhanced Learning (IJTEL), Volume 16, Number 1, January 2014*

*[7] Shreyas Kudale1, Advait Kulkarni2, Asst. Prof. Leena A. Deshpande3, Predictive Analysis Using Hadoop:, IEEE REVISTA IBEROAMERICANA DE TECNOLOGIAS DEL APRENDIZAJE, VOL. 9, NO. 3, AUGUST 2014*

*[8] Sérgio André Ferreira and António Andrade, Academic Analytics: Mapping the Genome of the University A Survey, International Journal of Innovative Research in Computer and Communication Engineering Vol. 1, Issue 8, October 2013.*