PREDICTING ACADEMIC COURSE PREFERENCE USING HADOOP INSPIRED MAPREDUCE

J.HimabindhuPriyanka, C. Akhila, N.Pranay Reddy, V. Sravan

[1]Assistant professor, Department of Computer Science and Engineering
Anurag Group of Institutions, Hyderabad, Telangana, India

2, 3, 4 Department of Computer Science and Engineering
Anurag Group of Institutions, Hyderabad, Telangana, India

ABSTRACT

With the rise of new innovations, new scholastic patterns brought into Instructive framework which brings about huge information which is difficult for understudies to really like to those scholastic courses which are useful in their modern preparing and expands their vocation possibilities. Guide Reduce occupations run over Hadoop Clusters by parting the enormous information into little lumps and cycle the information by running it equal on conveyed groups. Information separated utilizing Map Reduce will be useful in dynamic for understudies to decide courses picked for modern trainings. In this venture, we are inferring ideal courses for seeking after preparing for understudies dependent on courses which are primarily being preferred by the majority of the understudies. Here, Map Reduce yield is gotten after total of results.

Keywords: Mapper, Reducer, Shuffle, Aggregation

INTRODUCTION

Data mining is one of themaximumoutstandingregions in currenttechnology for retrieving significantrecords from massivequantity of statisticsthe usage of parallel processing of statistics. There is massivegain to Educational area of following Data Mining Techniques to examinestatisticsenter from students, feedbacks, trendyeducationaldevelopmentsetc which facilitates in impartingbesttraining and decision-making technique for students to growth their professionpotentialities and properchoice of guides for business trainings to fulfil the abilityhole pertains amongnumber onetraining and enterprise hiring students. Data Mining has outstandingeffect in educationalstructureswhereintraining is weighed as number oneenter for

societal progress. Big statistics is the risingarea of statistics mining. It is a time period for datasets which might be so big or complicated that conventionalstatistics processing application software program is incompetent to cope with them. Big statisticsconsists ofaccumulating of statistics for garage and evaluationmotive which advantagemanipulate over operations like searching, sharing, visualization of statistics, question processing, updating and preserveprivateness of records. In Big statistics, right here is extraordinarilybig dataset this is analyzed computationally to reveal patterns, developments and associations. Hadoop is one approach of massivestatistics and solution to issuesassociated withdealing with of hugestatistics. Hadoop is an open-supply programming paradigm which plays parallel processing of programs on clusters. Big Data technique can assist colleges, institutions, universities to get a complete aspect approximately the students. It facilitates in answering questions associated with the studying behaviors, higherexpertise and curriculum developments, and destinyroutechoice for students which facilitates to create charmingstudyingstories for students. The hassle of noticeablybiglength of dataset may be solved the usage of Map Reduce Techniques.

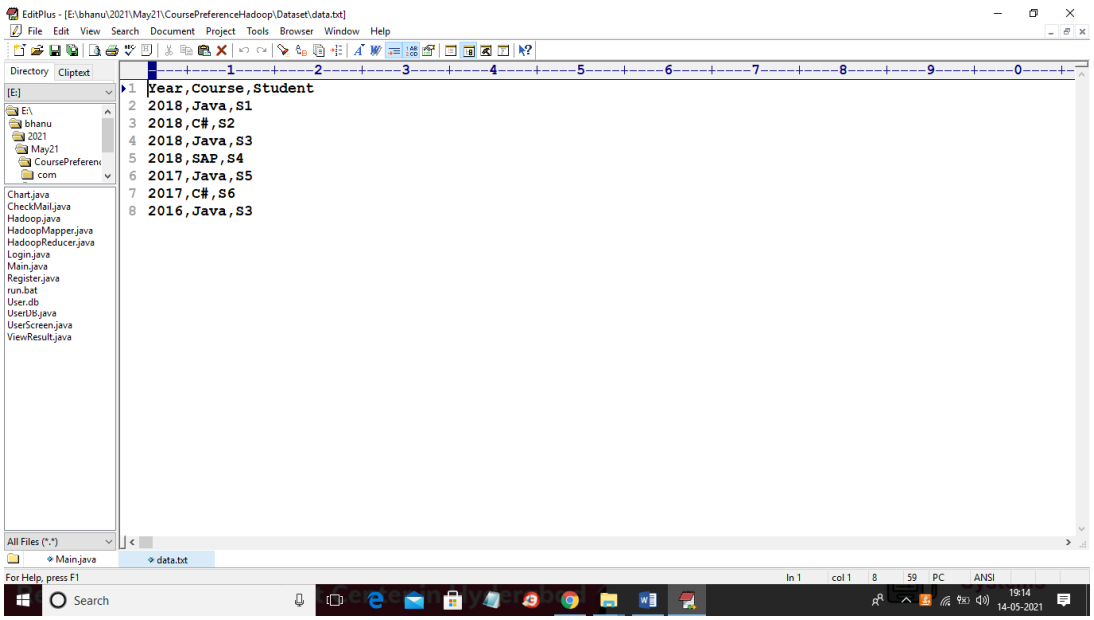
PROPOSED SYSTEM

In our proposed system we are analyzing the courses that are being preferred year wise using mapreduce and providing the results. This provides the user information regarding the course preference since various years. This would reveal whether there is progress or regress in the number of students preferring the course and the student could decide whether to opt a course or not.

EXPERIMENT AND RESULT

Here we are finding course preferences using HadoopMapReduce Framework. We will give student course details as input and MapReduce and then MapReduce will read each course details and then find out course preference percentage.

To implement this project we have used below dataset

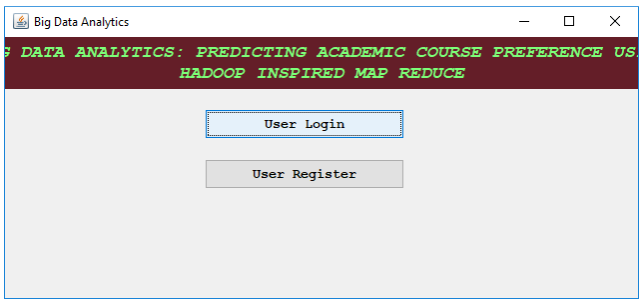


Above dataset you can change but all values should be comma separated and each row must have 3 values separated by comma. You can increase dataset size also by adding new rows. To implement this project we have designed following modules

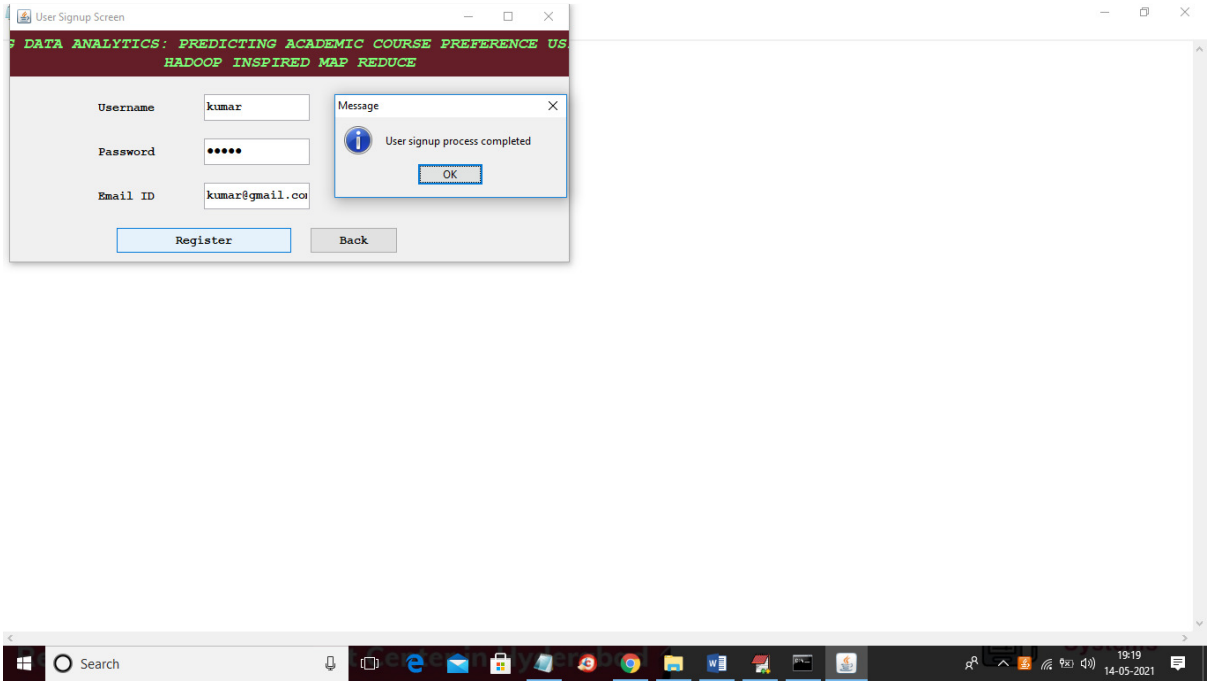
1) Register: Using this module we will signup with the application
2) Login: Using this module user can login to application
3) Upload Dataset: using this module user can upload course dataset to application
4) Apply MapReduce technique using this module we will apply MapReduce algorithm on uploaded dataset and selected course and then MapReduce will find out preference %.
5) Course Preference Percentage Graph: using this module we will plot course preference output in the form of graph

SCREEN SHOTS

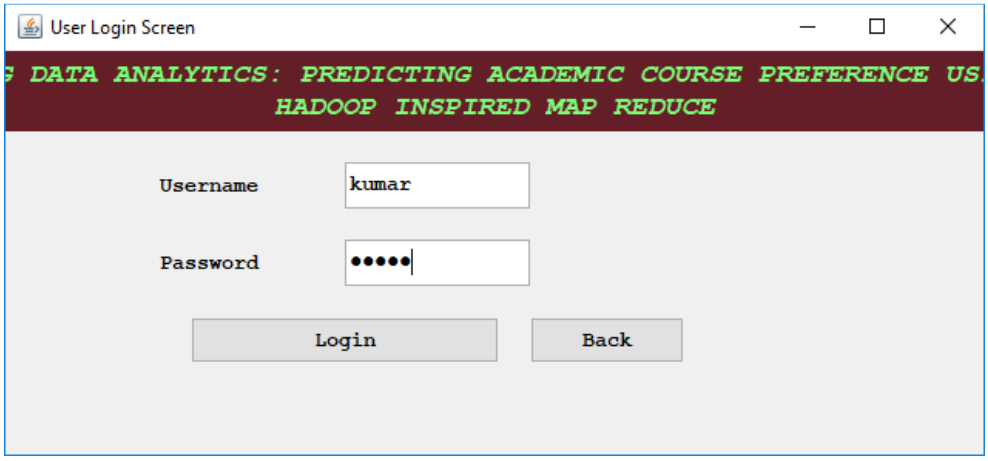
To run project double click on 'run.bat' file to get below screen



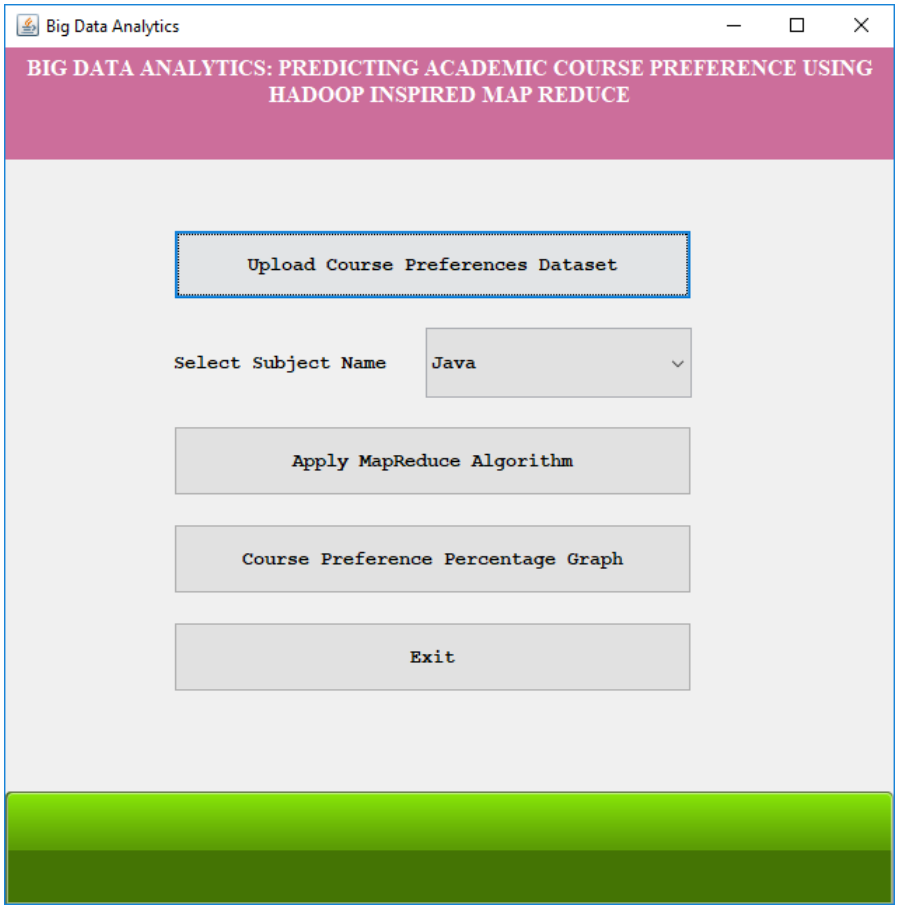
In above screen click on ‘User Register’ button to signup new user



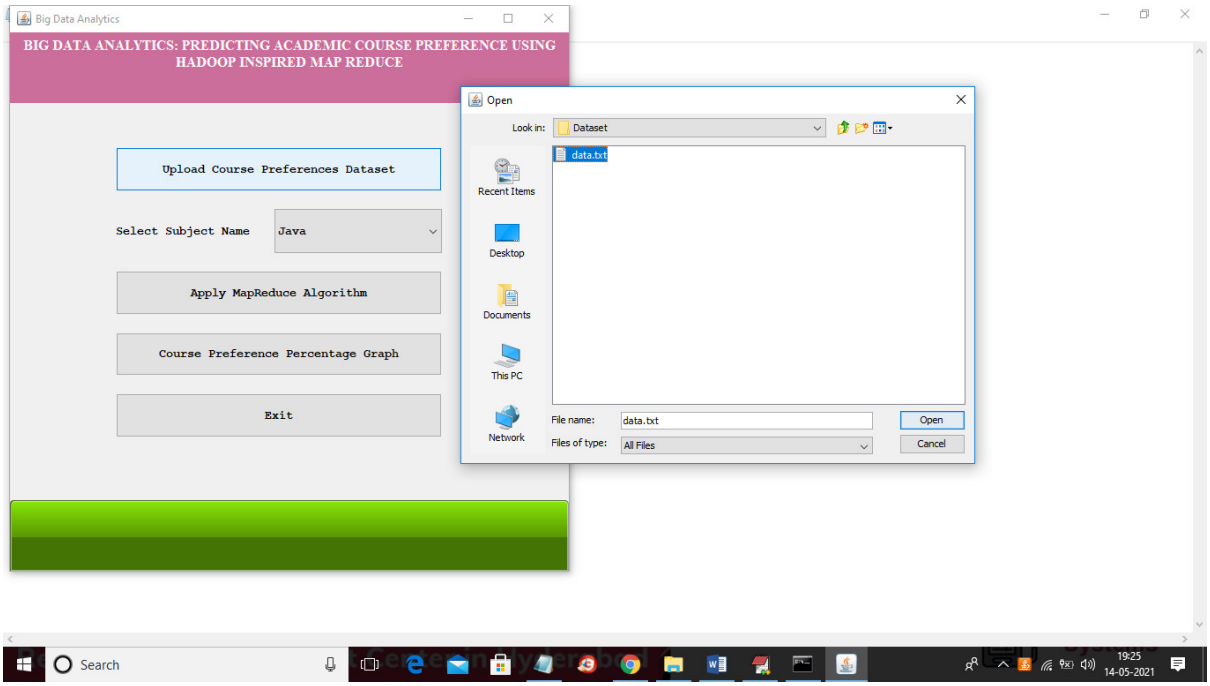
In above screen one user is signup and then click on ‘OK’ button to get below screen and then click on ‘User Login’ button to get login as user



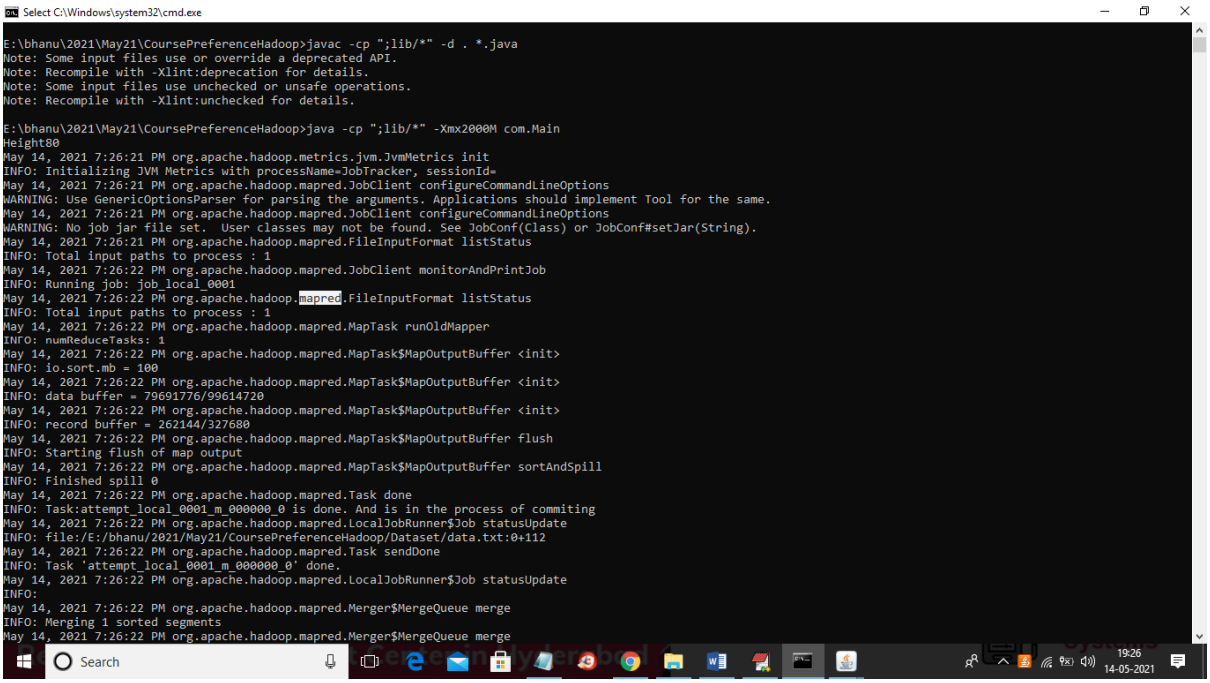
In above screen user is login and then click on ‘Login’ button to get below screen



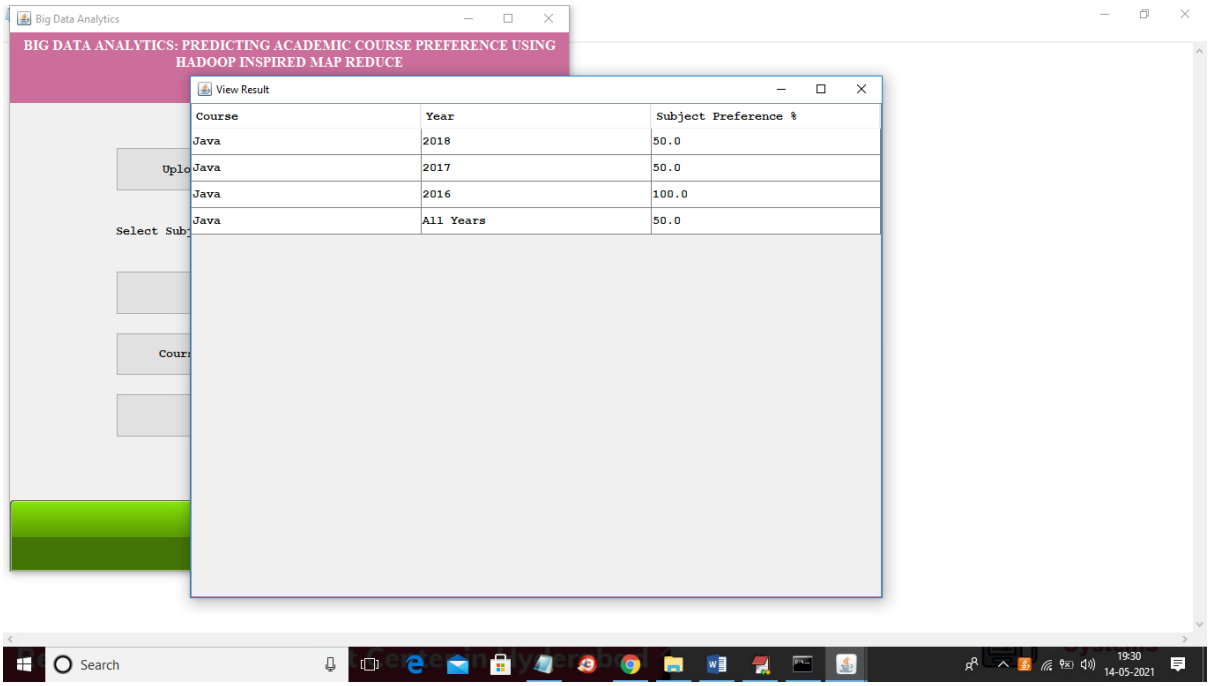
In above screen click on ‘Upload Course Preferences Dataset’ button to upload dataset and to get below screen



In above screen selecting and uploading 'data.txt' file and then click on 'Open' button to load dataset and then select desired subject from drop down box and then click on 'Apply MapReduce Algorithm' button to apply MapReduce and to get course preference %

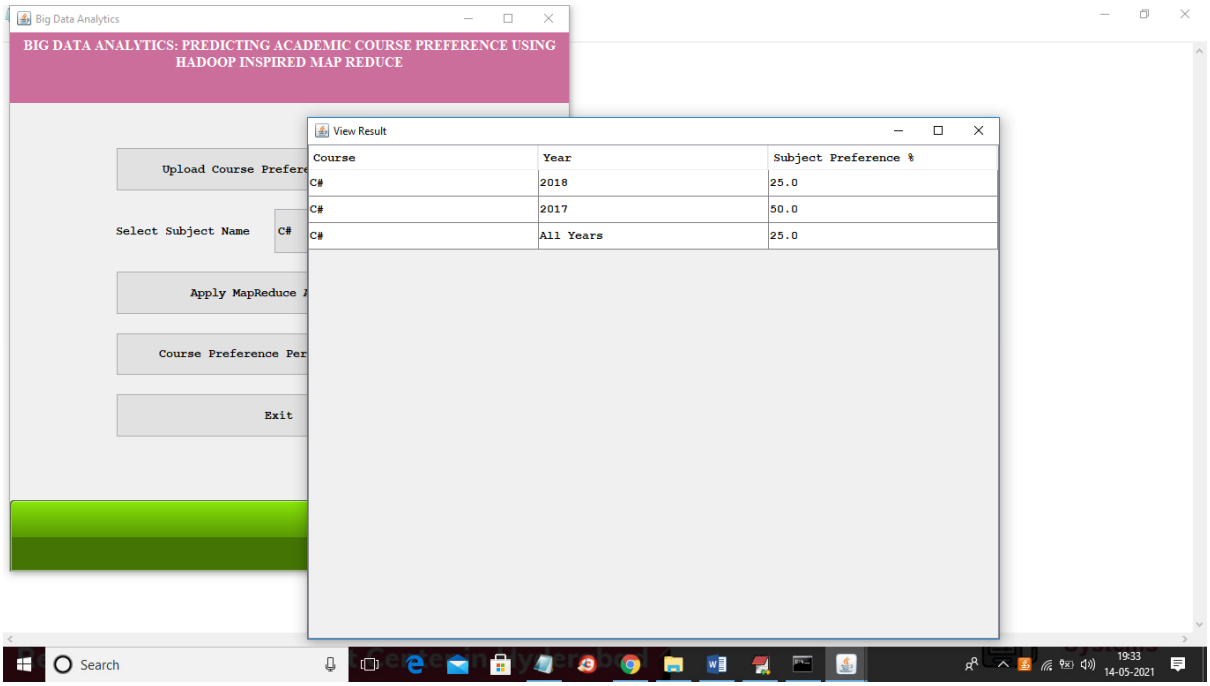


In above screen MapReduce starts processing and then will get below output

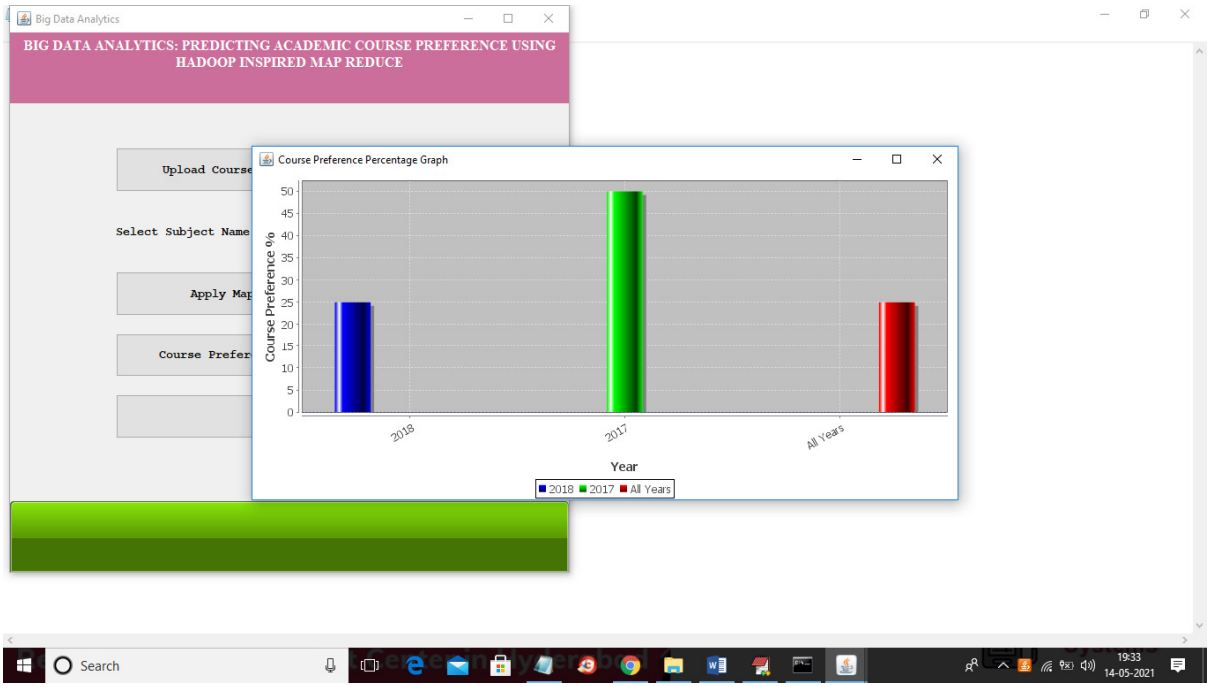


In above screen for selected course 'Java' we got above preference % and similarly you can select other courses and test and now click on 'Course Preference Percentage Graph' button to get below graph

In above graph x-axis represents YEARS and Y-axis represents course preference percentage of selected course 'Java'. Now below is the output of C#



Below is C# course Graph



CONCLUSION

Since the information generated from numerous instructional establishments is massive in size, it might be hard for a scholar to research the dataset and locate the publications which might be being broadly speaking favored and that are trending. So right here comes the answer of MapReduce that is used to method big information and with the aid of using this the scholar can examine the information and locate the publications which might be in call for and people that are trending. The Map Reduce Framework includes Map and Reduce Functions with unmarried Resource Manager which acts as a grasp and one Node supervisor which acts as slave in keeping with cluster node. The enter dataset is fed into the mapper and after passing through shuffle phase, reducer presentations the output after aggregating the tuples received from mapper and are in the shape of pair. The consumer is even furnished with graphical displays for every course which makes the scholar to research the information effectively and accurately.

REFERENCES

1. SonaliAgarwal, G. N. Pandey, M. D. Tiwari, “Data Mining inEducation: Data Classification and Decision Tree Approach”, International Journal of e-Education, e-Business, e-Management and e-Learning, Vol. 2, No. 2, April 2012.
2. Jongwook Woo, "Apriori -Map/Reduce Algorithm." Proceedings ofthe International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (World Comp), 2012.
3. N.Tajunisha,M.Anjali,"PredictingStudentPerformanceUsingMapReduce”,IJECS,Vol.4,Issue 1, Jan 2015, p. 9971-9976.
4. Mallikarjuna Reddy, V. Venkata Krishna, L. Sumalatha," Face Recognition based on Cross Diagonal Complete Motif Matrix", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.10, No.3, pp. 59-66, 2018.DOI: 10.5815/ijigsp.2018.03.07.
5. Ayaluri MR, K. SR, Konda SR, Chidirala SR. 2021. Efficient steganalysis using convolutional auto encoder network to ensure original image quality. PeerJ Computer Science 7:e356 https://doi.org/10.7717/peerj-cs.356
6. Shankar M.Patil, Praveen Kumar, “Data Mining Model for Effective Data Analysis of Higher Education Students UsingMapReduce”,IJERMT,ISSN:2278-9359,Vol.6,Issue 4,April 2017.

7. A.M. Reddy, V. V. Krishna, L. Sumalatha and S. K. Niranjan, "Facial recognition based on straight angle fuzzy texture unit matrix," 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), Chirala, 2017, pp. 366-372, doi: 10.1109/ICBDACI.2017.8070865. (
8. A.M. Reddy, K. SubbaReddy and V. V. Krishna, "Classification of child and adulthood using GLCM based on diagonal LBP," 2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), Davangere, 2015, pp. 857-861, doi: 10.1109/ICATCCT.2015.7457003.
9. Swarajya Lakshmi V Papineni, SnigdhaYarlagadda, HaritaAkkineni, A. Mallikarjuna Reddy. Big Data Analytics Applying the Fusion Approach of Multicriteria Decision Making with Deep Learning Algorithms *International Journal of Engineering Trends and Technology*, 69(1), 24-28, doi: 10.14445/22315381/IJETT-V69I1P204.
10. Srinivasa Reddy, K., Suneela, B., Inthiyaz, S.,Kumar, G.N.S., Mallikarjuna Reddy, A." Texture filtration module under stabilization via random forest optimization methodology "International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, No.3, May - June 2019
11. MadhaviVaidya,"ParallelProcessingofclusterby MapReduce",IJDPS,Vol.3,No.1,2012.
12. Jeffrey Dean, Sanjay Ghemawat, " MapReduce: Simplified Data Processing on Large Clusters", Google, Inc, OSDI 2010.