Random Forest Algorithms and Prediction of Student Satisfaction in Higher Education Organization

- **Poonam Singh**   • **Bhavana Narain**

Abstract: Higher education is the basic requirement of today's youth. Datamining is a domain that works for large datasets. It provides various standard algorithms to get knowledge from a large dataset. It works on structured and unstructured datasets. Prediction of student satisfaction in any educational organization is the first and foremost priority. In past years manual methods were used for surveys of student satisfaction. The arrival of technology has changed the pattern of the survey. Technology has increased the reach of the organization. In our work, we have used the random forest, the technique of data mining for survey and analysis of student satisfaction in educational organizations. Data collection, preprocessing of data set and feature extraction are done. Dataset is generated by the questioner. We have designed a google form for collecting data.

Keywords: *Student, Datamining, Education, Organization, Algorithm.*

Poonam Singh
MSIT, MATS University, Raipur, C.G. India
Email-id: poonam20phd@gmail.com

Bhavana Narain
MSIT, MATS University, Raipur, C.G. India
Email-id: narainbhawna@gmail.com

Introduction:

Satisfaction is the feeling of happiness that is found when a person fulfills his needs and desires. Student satisfaction is a short-term approach resulting from the evaluation of students' educational experiences. It is a positive antecedent of student loyalty and is the result and outcome of an educational system [1][2][3][4]. Student satisfaction can be defined as a function of the relative level of experiences and perceived performance about educational service during the study period. Satisfaction refers to the evaluation of service and comprises cognitive, affective, and attitudinal components. Higher satisfaction is associated with consumer loyalty and positive effect, in research that has mostly tested cross-sectional associations. Therefore, institutional outcomes such as word-of-mouth recommendations, enrolment, and retention relate to student satisfaction; these associations are generally interpreted as indicating consequences of

satisfaction but reciprocal relations are possible [5][6][7][8]. Student satisfaction can also have important implications for teaching due to its bidirectional associations with student engagement and achievement [9][10][11][12]. The national Higher Education Regulatory Authority(Nhera), will be set up to regulate in a 'light but the tight' and facilitative manner, meaning that a few important matters –particularly financial probity, good governance, and fully online and offline public disclosure of all finance, procedures, faculty/staff, course, and educational outcomes will be very effectively regulated while leaving the rest to the judgment of the HELs, which is essential to institutional autonomy, innovation, and pursuit of excellence.

Contribution and Organization of Paper

Hoang Thai Son et al. (2018) This paper reviews the role of student satisfaction and factors that may influence student satisfaction and presents the empirical research with the case with universities. assessing the satisfaction of male and female students.

Diah Dharmayanti et al, (2018) Discussed in the paper Education sector all over the world is now facing a demanding challenge which is about new innovation in providing education and quailed environment. research analyzes the students' satisfaction, students' loyalty, competitive advantage, and financial sustainability. The results of this study show that: Student value creation has a significant effect on students' satisfaction. Students' satisfaction has significant effects on students' loyalty.

Van Deuren, Rita et. al. (2017). Student Satisfaction in Higher Education: A comparative study of a Public and a Private college. 1.

Materials and Methods:

In reference to our previous research on student satisfaction, this study used a survey approach. To excel in our research, we have developed a questionnaire based on the results of the literature review measuring students' satisfaction. In total 30 items were created in the categories of teachers' roles (such as fairness and s, mentoring)course curriculums (such as relevant assignments and project works are given to help us learn the subject), college facilities administrative staff(such as the administrative staff responds to my needs promptly), future prospects (such as the college organization guest speakers to help us be informed about job market), college location (e.g. the college is accessible to transportation)and overall satisfaction (such as I have gained positive environment from the institute. Later on, demographic characteristics such as gender, age, course, and year of students were measured.

Data mining is a technique was the analysis of a big amount of data. This technique is helpful in our domain as our domain details millions of data in sets. As we know that classification is one of the methods to analyze a big amount of dataset. With labels. In our work quality ratio is the mail label and so we have selected a classification method for the analysis of our dataset. We have developed a program in Python. This program is for Random Forest, J48, and C4.5 classifiers. The flow of our work starts from the collection of raw data, Preprocessing, and development of the prepared dataset. In the next step, we go for the training, testing, and validation part. Training of data is done in five rounds. This training is for the given three algorithms. Trained data is then tested with the given dataset.

Preparation of Dataset

For our work, we have designed a google form and circulated it in various colleges and universities of Chhattisgarh. Raw data were collected and preprocessing was applied to remove noisy data.

All questions were related to students of higher education. These questions helped us to find students' view about higher education. We have also gone through National Education Policy 2020. We have finalized six features on which work has been done.

Feature Extraction and Feature Selection

In our work we have selected six features which are given as follows

Table 1. Details of Feature Selection

Student Satisfaction	Academic Support	Admin Support	Infrastructure	University Reputation	Placement

Proposed Methodology

We have proposed two hypotheses for our work

H1 To find out student satisfaction in higher education based on five parameters

H2 To find the accuracy of student satisfaction.

Machine Learning techniques are widely preferred over statistical analysis. Because of that in this research proposal, a data mining technique will be used. Data mining supervised and unsupervised data and techniques include clustering, classification, association, etc. tools to study the data. In this research work-appropriate algorithm preferred to fulfill the objective of the research objective [13][14] [15][16] [17][18] [19][20][21]. In our work, we have collected datasets and received various patterns of student interest in various higher education organizations.
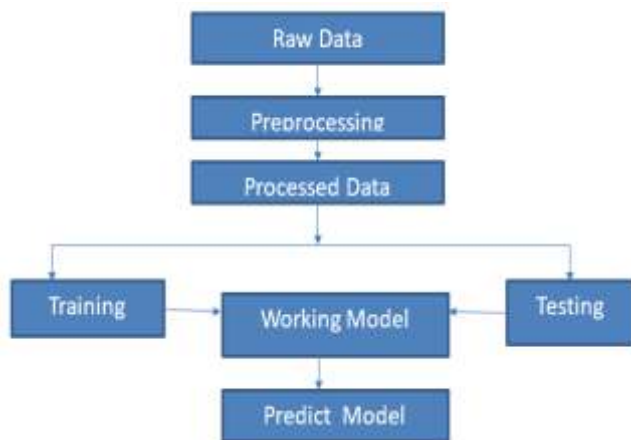

Fig. 1. Machine Learning Process Cycle

Random Forest Algorithm

Forest Algorithm Work as a large Collection of the Correlated decision tree well, the name forest is because we use a lot of decision tree But this algorithm of random forest creates a lot of decision trees and use them to make a classification (Amrehn, (2018). That is why it is a technique based on the random technique, that was presented before so, how do the random forest algorithm works let's see

$$m \ = \ \begin{matrix} fa1 & fb1 & fc1 \\ fa2 & fb2 & fc2 \\ fan & fbn & fcn \end{matrix}$$

we show here in this example a matrix m supports this matrix s is a matrix of training sample that we will submit to this algorithm to create a classification model.

Expected Outcome of The Proposed Work:

Dataset is of supervised nature. We have selected random forest techniques of data mining. The program is simulated in Weka.

- Factors that influenced student enrollment.
- The value of college tuition.
- Student satisfaction with financial aid.
- Faculty feedback, fairness, and availability.
- The class registration process.
- Campus safety.
- Staff helpfulness.
- Whether students feel welcome on campus.

In our algorithm satisfaction level is the dependent variable and the rest of the parameters are the independent variable. The random forest algorithm shows the best result in our dataset in terms of satisfaction level.

Conclusion and Future Work

- Our study will help education organizations to reduce the dropout ratio of the students.

- This ratio can be reduced to a significant level and improve the performance level of the education organization.

Our technique is Random Forest Classification. Best action must be taken for student satisfaction by comparing the result of all classification techniques.

References :

1. Mudasar Shahzad, "key factors for determining student satisfaction in distance learning courses: a study of Allama Iqbal open university", Turkish online journal of distance education,2011, vol. 12

2. Amran Rasli, Ahmad Reza Shekarchizadeh, Muhammad Jawad Iqbal, "Perception of Service Quality in Higher Education: Perspective of Iranian Students in Malaysian Universities", International Journal of Academic Research in Management (IJARM),2012, Vol. 1, No.1, ISSN: 2296-1747

3. AzleenIlias, Hishamuddin Fitri Abu Hasan, "Student Satisfaction and Service Quality", ccse journals,2008, vol.1.

4. Diah Dharmayanti, Hatane Semuel, Devie, The Students Satisfaction, Student Loyalty, Competitive Advantage and Financial Sustainability on Private Universities in Surabaya", Advances in Social Sciences Research Journal, Vol.5, No.10

5. Farida Haji Hassan, Nooraini Mohmd Sherif, "students' need recognition for higher education at private colleges in Malaysia", Sunway academic journal vol.3, 61–71

6. Heather J. Green, Michelle Hood, David L. Neumann, "Predictors of Student Satisfaction with University", Psychology Learning & Teaching,2015, Vol. 14(2) 131–146

7. HishamuddinFitri Abu Hasan, AzleenIlias," Service Quality and Student Satisfaction: A Case Study at Private Higher Education Institutions",ccse journals,2008, vol. 1, No.3

8. Jinqing Zhang, Pengchao Zhang and Bin Xu "Analysis of College Students' Public Opinion Based on Machine Learning and Evolutionary Algorithm", Hindawi, Vol. 2019

9. M. Samaha and N. S. Hawi, "Relationships among smartphone addiction, stress, academic performance, and satisfaction with life," Computers in Human Behavior, vol. 57, pp. 321–325, 2016.

10. Z. Demirtas, "Relationship between school culture and student achievement," Egitim VeBilim-Education and Science, vol. 35, no. 158, pp. 3–13, 2010.

11. 11.L. Yu and S. Ding, "Broad learning-based optimization and prediction of questionnaire survey: application to mind status of college students," Complexity, vol. 2018, Article ID5736030, 9 pages, 2018.

12. Gaganjot Kaur, Amit Chhabra, Improved J48 Classification Algorithm for the Prediction of Diabetes, International Journal of Computer Applications (0975 – 8887) Volume 98 –No.22, July 2014

13. Edy Budiman, Haviluddin Nataniel Dengan, Awang Harsa Kridalaksana, Masna Watiand Purnawansyahn Performance of Decision Tree C4.5 Algorithm in Student Academic Evaluation, Springer Nature Singapore Pte Ltd. 2018 R. Alfred et al. (Eds.): ICCST 2017, LNEE 488, pp. 380–389.

14. Hoang Thai Son and Ngo Thuy Ha, Pham Thi Minh Khuyen, "Measuring Students' satisfaction with higher education service – An experimental study at Thai Nguyen University" International Journal of Business Marketing and Management,2018, vol.3 No 4 P.P. 21-34,2018

15. IM Salinda Weerasinghe, R. Lalitha, S. Fernando, "Students' Satisfaction in Higher Education Literature Review", American Journal of Educational Research, 2017, vol. 5, No. 5, 533-539,2017

16. 16 Jalal R. M. Hanaysha, Haim Hilman Abdullah, AriWarokka, "Service Quality and Students' Satisfaction at Higher Learning Institutions: The Competing Dimensions of Malaysian Universities' Competitiveness", Journal of Southeast Asian Research,2011, vol.4.

17. Jane Whitney Gibson "A Comparison of Student Outcomes and Student Satisfaction in Three MBA Human Resource Management Classes Based on Traditional Vs. Online Learning", Journal of College Teaching & Learning,2008, vol.5, No.8

18. Nara M. Martirosyan,D. Patrick Saxon, Reubenson "WanjohiStudent Satisfaction and Academic Performance in Armenian Higher Education", American International Journal of Contemporary Research,2014,vol.4, No.2

19. Qidong Cao,T homas E. Griffin, XueBai "The Importance of Synchronous Interaction for Student Satisfaction with Course Web Sites", Journal of Information Systems Education, 2014, vol.20 (3)

20. Sein min, Cheychorkhoon "Demographic Factors in the Evaluation of Service Quality in Higher Education: International Students' Perspective", International Review of Management and Business Research journal,2013, vol.2, issue 4.

21. Sepideh Farahmandian, HadiMinav and, Mona Afshardost, "Perceived service quality and student satisfaction in higher education", IOSR Journal of Business and Management, 2013, vol.12, Issue 4.

22. Sik Sumaedi, I Gede, Mahatma Yuda Bakti, Nur Metasari, "The Effect of Students' Perceived Service Quality and Perceived Price on Student Satisfaction", Management Science and Engineering, 2011, vol.5, no. 1