Enhanced Elective Subject Selection for ICSE School Students using Machine Learning Algorithms

Poulami Dash* and V. Vaidhehi

Department of Computer Science, Christ University, Hosur Road, Bengaluru – 560029, Karnataka, India;
Poulami.dash@cs.christuniversity.in, vaidhehi.v@christuniversity.in

Abstract

Objective: Academic advising requires a lot of expertise, time and responsibility. To assist the human advisors in an efficient way, upcoming of computerized advising system is a necessity **Methods/Statistical Analysis:** Course Advisory System has been implemented using WEKA tool to recommend subjects for 8th class students of ICSE board. Machine learning algorithms – Naïve Bayes, J48, PART, Random Forest and KNN have been modeled and tested on the data set. The performance of each classifier has been compared and analyzed **Findings:** It is inferred that no advising system has been developed to assist school students in subject selection. Research work based on Indian students' requirements is minimal. Research work based on students' data caters more on binary class problems whereas the addressing of multi class problems is minimal. This work proposes an advising system for the school students of 8th standard of ICSE board to choose their electives. **Application/Improvements:** This work focuses on Indian educational system of school students. The approach takes care of the school students which will add its advantage to the existing systems. As school students are more vulnerable by taking wrong decisions, the course Advisory system will assist them in analyzing their academic history and help them choose their electives wisely. The classification algorithms might give a better accuracy with increasing instances. The Course advisory system can be enhanced using ensemble approach.

Keywords: Course Advisory System, Feature Selection Algorithms, ICSE School Students, Machine Learning Algorithms, Subject Selection

1. Introduction

In the education system, students have to select the subjects that they to learn from a pool of various options. It is a crucial decision making process as the future opportunities largely depend on the subjects that they had studied at the school level. This research work focuses on assisting human advisors to advice 8th class students of ICSE board to choose their electives.

Indian Certificate of Secondary Education (ICSE) is a board of education in India. The subjects offered in 8th class are divided into three groups. First group consists of the compulsory subjects- English, History & Civics, Geography and Indian language. Second group consists of nine options- Mathematics, Science, Environmental Science, Mathematics, Commercial studies, Technical Drawing, A modern foreign language, A classical language, Computer science, Environmental science and Agricultural Science out of which two subjects has

to be chosen. The third group consists of three options – Economic Application, Computer Application and Commercial Application out of which one subject has to be chosen. If the choice of subjects is inappropriate, then the student might get barred throughout life from reading that particular subject. For example, if a student doesn't choose math in 9th and 10th then the student might find it difficult to study in the later years, so advising the students must be done carefully.

The advisor generally takes decisions based on their previous experiences and the current rules of the school. This suggests for a model of computerized advisory system which uses the concept of training the system with the past data and testing the system with the new data. To make advising more profound the current research study proposes the use of machine learning algorithms to train the system. WEKA supports machine learning algorithms in an efficient way. As academic problems are liable to repeat themselves so the solutions given to past

problems can be used to give solution to similar current problems.

The need of computer based academic advising systems is elaborated[1,2]. The decisions taken by human advisors can sometimes be biased. To overcome this issue, computerized system can assist human advisors in taking academic decisions. Generally human advising is based on past experiences and the current rules of the institution. Course advisory expert system[1] is designed as case based and rule based reasoning system to recommend courses to a student. The recommendation is done on the basis of the academic information, academic details and family background details of students. Generally students choose courses depending on the risk level[2]. According to the risk theory, there are three dimensions of course selection- lecture's style of teaching, difficulty of the course and the learning value. There are various other reasons which can decide the choice like- graded performance, classroom environment and motivation. Students' data is collected using questionnaire and risk perception is studied. In[3] employed the Intelligent Academic Advising using Genetic algorithms and decision tree to track the progress of the students and help the human advisors and students to achieve their academic goal. In[4] suggests an online course recommender which suggests courses to the students depending on their learning styles. The information about the learning styles and learning pattern of students is collected through the 44 item questionnaire. In[5], software tools can help faculty and staff who advise undergraduate students and is used in course planning.

In[6] online-based academic advising system is designed based on educational curriculum of each institution, grading system and university policies using Decision tree. In[7] a distance education advisor is designed based on interest and academic performance of students using user-based collaborative filtering. In[8] machine learning techniques like IB1, Naïve Bayes, C4.5, and RBF are used to improve learning experience for online students. In[9] decision tree algorithm is used to analyze the data and improve the student's leaning process. In[10] student's performance is analyzed and the grade is predicted using classification algorithms- Naïve Bayes, K-means, Expectation Maximization, K- nearest neighbors and Decision tree.

In[11] Artificial Neural Network and Web mining is used to identify the learning pattern of individual student. In[12] classification model is designed to predict student academic performance. In[13] classification algorithms are used to identify the college students who are at the risk of failing in their first year. In[14] early warning system is designed to estimate the student's success in exams using naïve Bayes algorithm. A web based advising system is designed for undergraduate students[15].

From the literature, it is inferred that the availability of advising system to assist school students in subject selection is very minimal. Also the research work based on Indian student's requirements for academic advising is minimal. It is also evident that the research work based on student's data caters more on binary class problems whereas the addressing of multi class problems is minimal. This current paper proposes an advising system which advises the school students of 8th standard of ICSE board to choose their electives.

2. Methodology

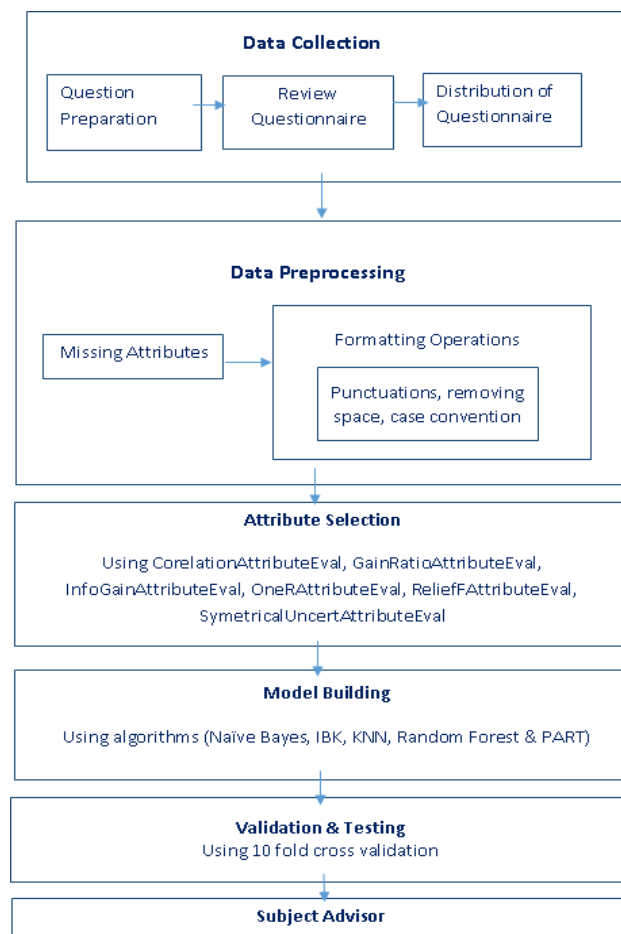The steps involved in implementing the Course Advisory System are given in Figure 1.

Figure 1. Advisory system.

Figure 1: steps involved in the process

- Data collection- a questionnaire was prepared to collect data of students within the age group of 14-18 belonging to ICSE board who are studying / completed 8th standard. The questionnaire consists of 34 questions. Pilot study was conducted to review the questionnaire. The questionnaire targets at collecting detailed information about 6th and 7th class academic record, focusing particularly on the subjects- Mathematics, Science and English. The questionnaire also aims at capturing detailed information about the family and financial condition of the student. The questionnaire was modified to make sure that people of every age group can understand and fill the questionnaire. After drafting the final questionnaire it was put into social media sites for collection of real time data. 324 student instances are considered for this research work.

- Data preprocessing- some data were not filled properly with more than 5 missing attributes, so those data were not included in the final data set. Punctuations and space was removed. Case was converted to lowercase.

- Attribute selection- using attribute selection algorithms like- CfsSubsetEval, 25 attributes were found to be important.

- Other attribute selection algorithms like- CorelationAttributeEval, GainRatioAttributeEval, InfoGainAttributeEval, OneRAttributeEval, RelieFAttributeEval, SymetricalUncertAttributeEval the ranking of all the attributes were found. Using majority voting technique, the final ranking of all the attributes was listed.

- Model building- with the selected attributes classification algorithms were applied to get the outcome. The algorithms used were- Naïve Bayes, J48, IBK, Random Forest and PART. Naïve Bayes algorithm assumes that the value of each attribute contributes equally to the outcome. KNN assigns weight to the neighbors. The nearest neighbors contribute more to the outcome than the more distant ones. PART generates decision list; outcome is based on the best attribute. IBK helps to select the value of K based on cross validation and also do distance weighting. Random Forest is used to create forest of random trees.

- Validation and testing- 10- fold cross validation was used to measure the accuracy of the dataset.

Description about each attribute:

Awareness- this attribute asks whether the student is aware of web based advisory System. From the collected data we can infer that 116 students are aware of the advisory systems and 208 students are not aware of the advisory systems. *Feel to opt any other course-* this attribute checks the number of students happy with their choice and whether they feel to opt any other course after joining. From the collected data we can infer that 140 students wanted to change the subjects and 184 students do not want to change the subjects. *Gone through-* this attribute checks whether the students have consulted any advisory system before choosing the subjects. From the collected data we can infer that 84 students took help in choosing the subjects and 240 students did not take any help in choosing the subjects. *Ready-* this attribute enquires whether the students are ready to accept a new web based course advisory system that is proposed in the present work. *After joining-* this attribute enquires whether the students have consulted any web based advisory system after joining the course. *Course Advisor to be-* this attribute identifies the type of course Advisory system like a machine interface, human interface, face to face counseling or telephonic.

If machine- if the Course Advisory System is a machine then the students may use it to as a web based system or automated voice based system. *Meeting expectation-* This attribute asks whether the course taken meets their expectation. *Source of information-* this attribute enquires the source from which they got inspired to choose a particular subject. Out of all the options, internet seemed to have been the major source of information. *Parent's qualification-* this attribute collects the highest academic degree of the parent of student. This attribute helps us to analyze the parental qualification. *Family status-* this attributes helps to understand the family status of the student. *Physical disability-* this attribute takes into account whether the student has some physical disability.

Level of attendance- this attribute accounts whether the student is punctual in class. *Class participation-* this attribute determines the attentiveness of the student on a scale of 5. *Medium of education before 6th-* this attribute helps to know the medium of education before 6th. This attributes helps the advisor in deciding whether the student can perform well in a particular recommendation of a subject. If the medium of education before 6th was some other language other than English then the student might find it hard to perform well with difficult subjects.

Previous board- ICSE board being the toughest board of education, anybody with any other board before 6th standard might find it difficult to adjust with the difficult subjects if recommended. *Favorite subjects-* this attribute helps in knowing the interest of the student in a particular course and recommend accordingly. *Least favorite subject-* this attribute helps the advisor to know the subjects that the student is not interested in, so that the advisor may never recommend those subjects to the student. *Subject best performed in –* this attribute helps the advisor in knowing the subject that the student is best in. In turn it helps to recommend a combination having that subject. *Subject struggled in-* this attribute helps the advisor in knowing a particular subject that the student is poor in. *Overall 6th percentage-* this attribute records the mark of the student in percentage. This attribute helps the advisor in knowing whether the student is below or above average. *Overall 7th percentage-* this attribute collects the mark of the student in percentage. This attribute along with the 6th percentage helps in deciding the overall performance of the student. *Area of interest-* this attribute focuses on the field in which the student wants to build the career so that while recommending a particular subject, the field of the student can be kept in mind to get better outcome. *Factors-* this attribute lists out the reasons behind choosing a particular subject. This attribute helps in understanding the mindset of the student while choosing a particular subject and leaves the responsibility on the hands of the advisor to direct the students on the path of their choice. *Role-* this attribute helps the advisor in understanding the role played by parents and teachers in helping the student choose the subject.

Long term interest – this attribute helps the advisor in understanding where the student wants to see himself in the upcoming years. Guided recommendation can help them reach their goal. *Extracurricular activities-* this attribute accounts for the extracurricular activities that the student is interested in apart from academics. As that sometimes might help the advisor in knowing the hidden talent of the student and recommend accordingly. *Connection-* this attribute tries to connect the extracurricular activities with academics. *Help-* this attribute helps the advisor to know the extent to which the student need help from the advisor. *Group 1-* lists out the electives liable to be chosen by the student. There are a total of nine subjects available out of which two subjects has to be chosen. *Group 2-* There are a total of three subjects available under this attribute out of which only one subject has to be chosen. *Class-* this attribute lists out all the combinations possible including two subjects from group 1 and two subjects from group 2.

3. Implementation

The procedure for implementing the course advisory system includes the following activities:

- The student should be of 8th class.
- The students should be of ICSE board.
- The students should fill the questionnaire.
- The questionnaire should take detailed information about the previous class academic record (Math, Science and English)
- The questionnaire should capture detailed information about the subjects best performed and the subjects worst performed.
- The questionnaire should clearly mention the combination of subjects to choose.
- The questionnaire should check the family background of the student.
- The students choosing Computer Science from group I should not choose Computer Applications from group II.
- After submitting the questionnaire the system should be able to recommend subjects to the student based on the trained records.

Initially 400 instances were recorded out of which only 324 instances were considered for mining. 76 instances had incomplete information. Initially 34 attributes pertaining to the students were recorded out of which only 25 attributes were finalized based on attribute selection algorithms.

4. Results

Accuracy is a performance measure considered for various multiclass classifiers. Figure 2 is the graphical comparison of the Classification algorithms with and without Feature Selection. It can be inferred that with feature selection the efficiency of each classification algorithm has increased. The accuracy of classification algorithm Random Forest has increased from 99.07 % to 99.69 %. In case of PART the accuracy has increased from 96.29 % to 96.91 %. IBK also shows an increase in efficiency from 99.07 % to 99.69 %. Overall, it can be inferred that the classification

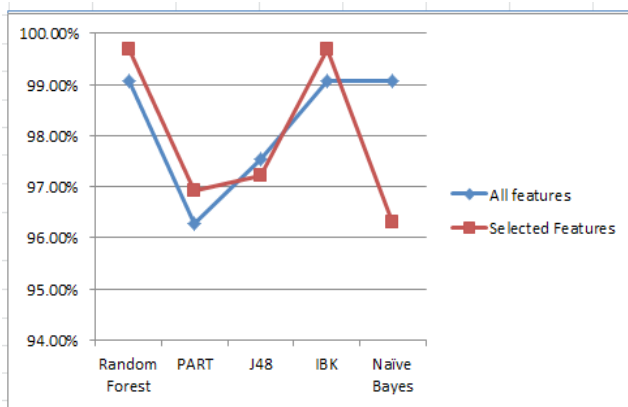algorithms- Random Forest and IBK are equally suitable for current data set.

Figure 2. Comparison of performance of all classifiers using feature extraction methods.

5. Conclusions

8th class students generally choose the elective due to parental pressure, peer pressure, class environment, grade and motivation which might not end up giving them the outcome that they expect in the future. Course Advisory System helps the students to know their strengths and weakness academically thus helping the students to choose an elective which will be helpful in realizing their future expectations. The approach taken in the current paper takes care of the school students which will add its advantage to the existing systems. As school students are more vulnerable to taking wrong decisions, the course Advisory system will assist them in analyzing their academic history and help them choose their electives wisely. The classification algorithms might give a better accuracy with increasing instances. The Course advisory system can be enhanced using ensemble approach.

6. References

1. Daramola O, Emebo O, Afolabi I, Ayo C. Implementation of an intelligent course advisory expert system: cased-based course advisory expert system. International Journal of Advanced Research in Artificial Intelligence (IJARAI). 2014; 3(5):6–12.
2. Zocco D. Risk theory and student course selection. Research in Higher Education Journal. 2009 Jan; 3:1–29.
3. Abdulwahhab RS, Makhmari HSA, Battashi SNA. An educational web application for academic advising. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) 8th GCC Conference and Exhibition; 2015 Feb 1–4. p. 1–6.
4. El–Bishouty MM, Chang T–W, Graf S, Kinshuk, Chen N–S. Smart e-course recommender based on learning styles. Journal of Computers in Education. 2014 Mar; 1(1):99–111. Crossref
5. Matte N, Dodson T, Guerin JT, Goldsmith J, Mazur M. Lessons learned from development of a software tool to support academic advising. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) Zone 1 Conference of the American Society for Engineering Education; 2014 Apr 3–5. p. 1–8. Crossref
6. Roushan T, Chaki D, Hasdak O, Chowdhury MS, Rasel AA, Rahman MA, Arif H. University course advising: overcoming the challenges using decision support system. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) 16th International Conference on Computer and Information Technology (ICCIT), Khulna, Bangladesh; 2014 Mar 8–10. p. 13–8.
7. Taha K. Automatic academic advisor. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) 8th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom); 2012 Oct 14–17. p. 262–8.
8. Singh S, Lal SP. Educational courseware evaluation using machine learning techniques. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) Conference on e-Learning, e-Management and e-Services, Kucing, Malaysia; 2013 Dec 2–4. p. 73–8.
9. Pathan AA, Hasan M, Ahmed MF, Farid DM. Educational data mining, A mining model for developing students' programming skills. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), Dhaka, Bangladesh; 2014 Dec 18–20. p. 1–5. Crossref
10. Manne S. Moodle system using datamining techniques. 2015 Mar; 5:1– 5.
11. Kolekar SV, Sanjeevi SG, Bormane DS. Learning style recognition using artificial neural network for adaptive user interface in e-learning. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) International Conference on Computational Intelligence and Computing Research, Coimbatore, India; 2010 Dec 28–29. p. 1–5. Crossref
12. Guo B, Zhang B, Xu G, Shi C, Yang L. Predicting students performance in educational data mining. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) International Symposium on Educational Technology (ISET), Wuhan, China; 2015 Jul 27–29. p. 125–8.
13. Gray G, McGuinness C, Owende P. An application of classification models to predict learner progression in tertiary education. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) International Advance Computing Conference (IACC), Gurgaon, India; 2014 Feb 21–22. p. 549–54. Crossref

14. Goker H, Bulbul HL. Improving an early warning system to prediction of student examination achievement. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) 13th International Conference on Machine Learning and Applications, USA; 2014 Dec 3–6. p. 568–73. Crossref

15. Gamulin J, Gamulin O, Kermek D. Comparing classification models in the final exam performance prediction. In the Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia; 2014 May 26–30. p. 663–8. Crossref