



An Intuitive Implementation of Course Recommendation System Based on Learner's Personality

Asad Jatoi, Memoona Sami*, Muzamil Nawaz and Junaid Baloch

Department of Software Engineering, Mehran University of Engineering and Technology, Jamshoro, Pakistan

*Corresponding Author: Memoona Sami, Department of Software Engineering, Mehran University of Engineering and Technology, Jamshoro, Pakistan.

Received: March 16, 2022

Published: April 20, 2022

© All rights are reserved by Memoona Sami, et al.

Abstract

The work covered in this study concerns learner classification; a learner can be divided into numerous groups based on how comfortable he or she is with the course. In a broader sense, a learner can be theoretical, practical, or hybrid: a mix of the two. The research will use the Naive Bayes method to determine the degree of excellence of learners in their recognized learning style in a particular competency once the learner's type has been determined. The level of quality is divided into four categories: exceptional, good, medium, and fair. Recommendation will be given to the student for further improvements within the skillset or for newer skills once the classification of the learner in a category has been determined, as well as his or her level of proficiency in the area. The study places a greater emphasis on determining a learner's learning style because it is so important to the study. The research's main building block is the identification of learning styles, as the second portion of the research's recommendation is entirely based on superior categorization results.

Keywords: Classification; Algorithm; Recommendation System; Naïve Bayes; Learning Style

Introduction

Smart studying involves ordered studying, it is believed that when the things are in proper order, the flow is well-known then doing it will be easier. In-order to recommend Student a course first it's necessary to identify the root factors, such as the type of learner and their learning style [1]. For 30 years it's been said that most of the people are better at learning by either auditory sensors, visual sensors or both [2]. But in recent research the concept of smart studying is hot. Visual Learning has more types introduced, with charts and graphs, with patterns and figures, with facts and raw data. Every student has a different mechanism of learning. Our study is focused on identifying types of learners and then recommending them the best-fit course for them.

Knowing the type of learner someone is, it becomes easier for them to learn perfectly and retain. Learning techniques differ from person to person. For some the best approaches are Taking Notes,

drawing mind-maps, studying in group, rewriting, lecturing etc. [4]. The mentioned techniques differ from individual to individual. The instructors believe the same. A Teacher gives all their best to create a competent class environment but most of students may differ in learning styles and a single teacher may not be able to convey the same piece of information to every student, therefore; we always have some average students in a class, and that's not rare.

Background

Literature reviews

We studied many research papers and applications to conduct this work. Different scholars' contribution was studied in detail. Few papers included datasets from different sources singleton or multiple in many cases showing that the study has to be versatile in nature. The table 1 shows the use of different algorithms in [4,7,8,10,11,14,15] studies.

Paper	Collaborative	Content- Based	KNN	Regression Analysis	Naïve Bayes	Decision Tree
[4]	X	X	X	✓	X	X
[7]	✓	X	X	X	X	X
[8]	X	X	X	✓	X	X
[10]	✓	✓	X	X	X	X
[11]	X	X	X	X	✓	X
[14]	X	X	X	X	X	✓
[15]	✓	X	✓	X	X	X

Table 1: Use of various Algorithms by different scholars.

As per results achieved from different research, figure 1 shows the score of algorithms.

Similarly, the scholars have collected the dataset for their model training process from different resources. Table 2 shows the resources of data collection used in different research papers.

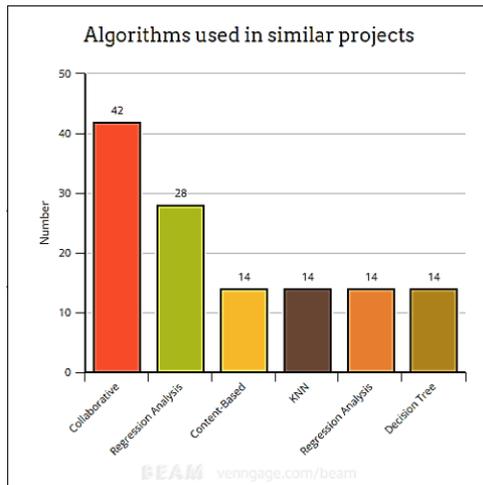


Figure 1: Chart of frequently used Algorithms used in various research.

The pi-chart in figure 2 shows the percentage of dataset sources. According to this figure institutional data and historic data LMS are used in more in projects than other dataset sources.

Machine learning

Machine Learning is the process of making a machine learn without providing the additional human logic to resolve the path for machine. At some extent the machine learning is automated, yet to improve the accuracy we still need the well purified dataset which in turn require human logic expertise.

Paper	Historic Data (LMS)	Institute's Data	Survey	Questionnaire Quiz	Static Data	Grades
[1]	X	✓	X	X	X	✓
[7]	X	✓	X	X	✓	X
[8]	✓	X	X	X	X	X
[9]	X	✓	X	X	✓	X
[11]	✓	X	X	✓	X	✓
[12]	✓	✓	X	X	X	X
[13]	X	X	✓	X	X	X
[15]	X	X	X	✓	X	X

Table 2: Resources from where data is obtained by different scholars.

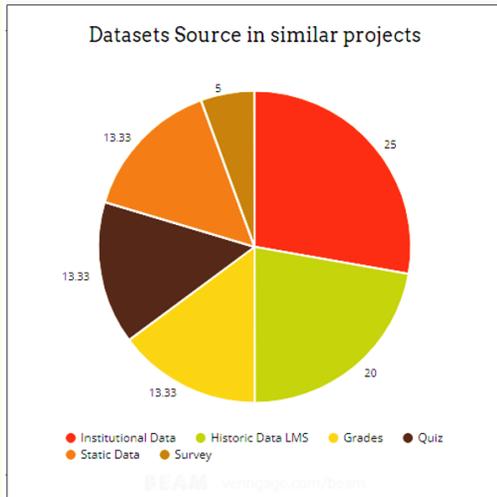


Figure 2: Percentage of data collection from various methodologies.

For this study, a model trained using Machine Learning library was required so the use of PHP Machine Learning was done. Machine Learning Algorithm Naïve Bayes was best fit for the implementation.

Dataset collection and purification

For any machine learning algorithm, the dataset is the basic building block. A good dataset defines the accuracy of an algorithm. The dataset for the study was gathered from different student and the dataset of around 500 students was collected. The students were asked to answer 30 different questions in related to java purely at first. The responses received from the Google Forms was limited so the alternative of re-generating the dataset was used with which a dataset of worth 5lac entries was generated using Java. Figure 3 shows dataset obtained from Google Forms, the dataset is all mumbled at first.

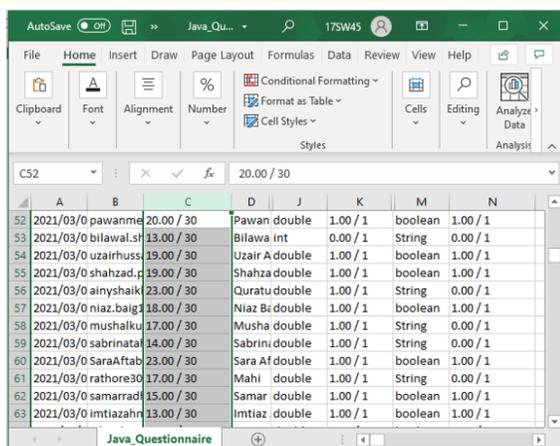


Figure 3: Dataset collected from Google Forms, Before Purification (CSV file).

The dataset is auto-generated from Google Forms (the third party service that gives the platform to conduct survey online and take the results as a Comma-Separated Values (CSV)) the results obtained contains absurd information that may not be required by algorithm so a filtered csv is created by removing all the absurd data from the file. The filtered CSV file has undergone the phase of pre-processing that is termed as purification process. Data purification is as vital as data gathering. The data that we have, always contains extra unnecessary information, one which was necessary at the beginning but later was absurd. Cutting off such information is important before processing the dataset for algorithm training process. For a classification-based algorithm training the dataset must have two important columns: one for the field upon which learning will be based. And other the result or classification label generated after experimenting with the first field. The figure 4 shows dataset after performing purification in Google Forms' generated CSV file.

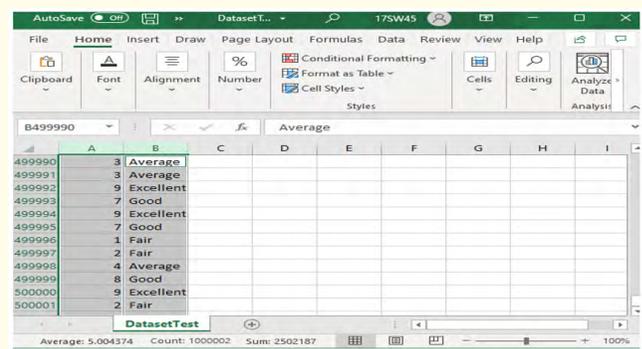


Figure 4: Updated Dataset after altering-out absurd data fields from Google's form generated CSV file.

The file is then passed in code for model training part.

Naïve Bayes Algorithm

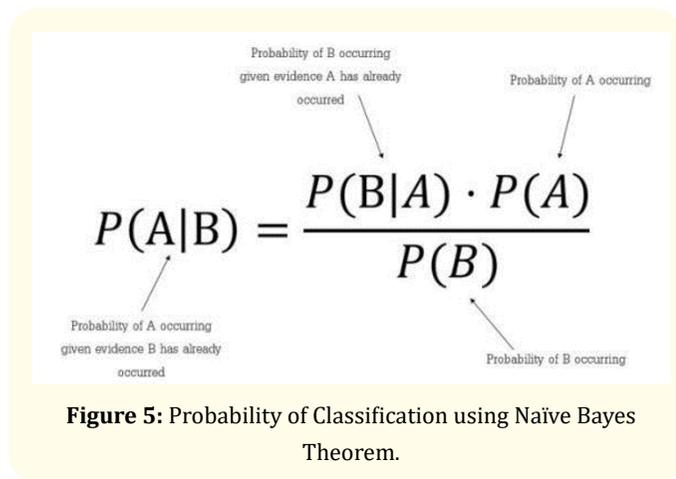
Having a dataset which comprises of scores from a questionnaire and all those have to be processed in order to classify the expertise-type a candidate has in the subject is reason why Naive Bayes algorithm is used.

Naive Bayes algorithm works on 'dependency of a variable over another variable' factor. In the dataset obtained we have 'Score' field upon which the classification lies. Then there is a derived dataset which has pre-determined classes for particular scores. Half of the entities having score are labeled with the proper classes i-e Average, Excellent, Fair, Good. Now the derived dataset is passed with Naive Bayes algorithm so that the algorithm learns the pattern how the data is being labeled.

Moreover, in future the research will have multiple fields such as: head_noding: Boolean, questioning: Boolean, discipline, marks: integer etc. upon which the classification will be made, keeping present and contemporary times hardships in mind Naive Bayes algorithm is perceived to be the best choice. Naïve Bayes algorithm, which uses concept of mathematical probability for filtration of data. It is named after the mathematician Bayes, as it follows the mathematical formula of Bayes probability theorem [3]. It's more accurate because it uses probabilities of categories and adds newer type into the category that is nearly like the probabilistic values. Naïve Bayes algorithm uses the concept of dependency events.

In PHP ML library Naïve Bayes is already available and is easily accessible because of its open-source contributors. In-order to train the model the dataset is required. The larger the dataset is; the more accuracy is generated [16-18].

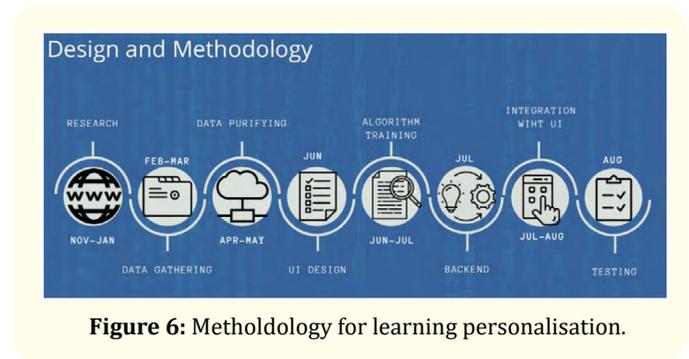
The pictorial form of the Naïve Bayes theorem is shown in figure 5.



Proposed methodology

The algorithm training process required a dataset when data was completely purified the training process of algorithm started. As stated above the study shows three types of super learners, Theoretical, Practical and Hybrid. But in each super-type we have sub-types of super-types. If a learner is classified to be a Theoretical learner, then how much good he or she really is justified afterwards. Sub-types involve 5 types: Excellent Learner, Good Learner, Average Learner, Fair Learner. An excellent learner is

the learner who possess the ability to score 9-10 in whole section, a good learner can score 6-8 questions right, average learner can answer 3-5 questions correctly and fair learner can get 0-2. The ranges are specified within sub-types. Based upon these ranges one can identify the goodness or level of excellence of a student in a subject. Figure 6 shows the complete methodology.



If a student is an Excellent Learner in a category that means he has achieved the perfection in the category move to next section. If a learner is Excellent in all three sections that he is good to move to the next skillset unlocking level, such a learner will be recommended the next programming technology to focus on which shows he or she is already proficient in the one they have opted. The algorithms that we used were SVM (Support Vector Machine: Clustering or Regression), Naïve Bayes and Linear Regression [5]. Out of these only two algorithms were fit for the nature of our study which were Naïve Bayes and SVM's Linear SVC. We tried the SVC algorithm on Google Collab and calculated the accuracy of algorithm and at the same time we used PHP ML algorithm Naive Bayes theorem and conclusion was that we got better accuracy and precision results from Naïve Bayes theorem. So, the algorithm that was finalized was done based on the accuracy results. With Naive Bayes algorithm 100% accuracy was achieved because the dataset used was big enough to learn all the sub-types of the learners.

In-order to test the algorithm, using Model Manager, we have to restore the saved algorithm. After restoration, predict method of Model Manager object will predict the classification label of the given field.

In-order to test the accuracy, Recall, F1 score etc. Metric library of PHP ml is used. Accuracy is measured using the given formula:

$$Accuracy = \frac{Corrected\ Predictions}{Total\ Predictions} \times 100$$

The algorithm can have variations. Instead of using Naïve Bayes algorithm other algorithms such as: Linear Regression, SVM (Support Vector Machine) etc. could have been used too, but better accuracy was obtained with Naïve Bayes algorithm, proving it to be best of its competent.

Results and Discussion

The model accuracy is tripled in our case because we are finding the learner type in every section and based on all results combined a proper type of learner is identified and based on identification of learner's type the best-fit course is being recommended to the learner. The accuracy is checked by Metrics package of PHP ML Library the accuracy checked and was displayed on a PHP page. The figure 12 shows the accuracy, precision, F1 score and Recall values for our study. The probability of accuracy for the study is 1 which in percentage means 100% since the dataset was large enough therefore more accurate results are found which are shown in figure 7.

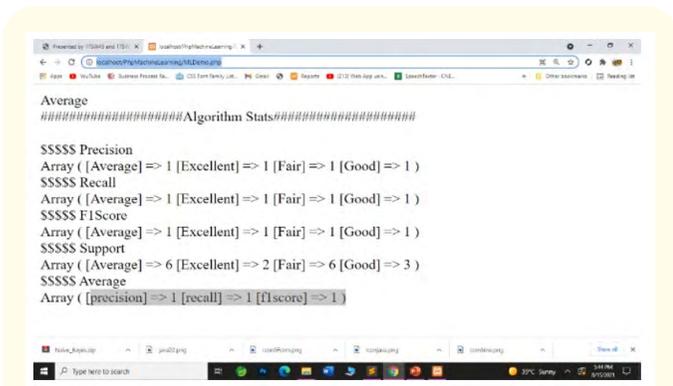


Figure 7: Statistics of Naïve Bayes Algorithm generated by performing algorithm on a test dataset.

Figure 8 shows different accuracy results achieved by using different algorithms on same dataset.

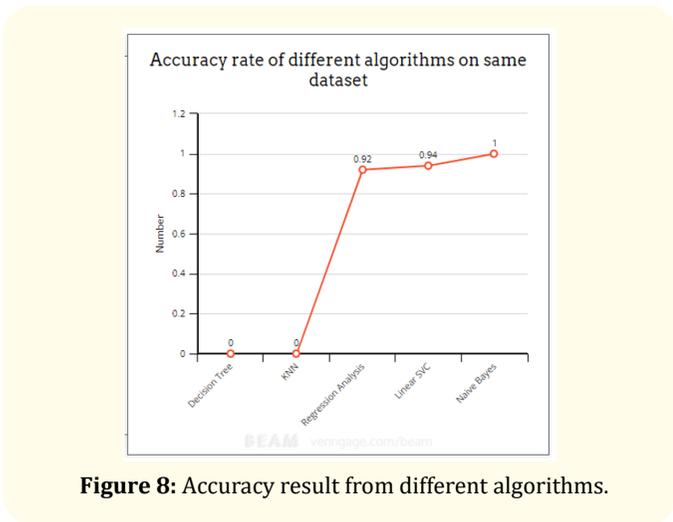


Figure 8: Accuracy result from different algorithms.

Currently this study is identifying the type of learner based on singleton value: score, the questionnaire [6]. If a smart class was established here, we would have many other resources to compare a student. On the website without interacting with the teacher of student we really can't add other important factors in the application such as: Nodding in Class, Ability of asking questions in class, ability to take notes and answering the teacher asked questions. Such factors can't be directly automated in the web. If these factors would have been added in the website, then even foremost better results would have been achieved and we could have cut off the questionnaire part of the study since everything is being calculated automatically there is no such need as questionnaire.

Conclusion and Future Work

This paper presents the complete solution for smart learning system. Likewise, other LMS already have such stronger recommendation systems. This study can also be used to better guide candidates regarding their knowledge base. It doesn't only compare based on the single attribute, but this is used to classify students into different categories.

In future, we predict that the classrooms will be smart enough to have surveillance cameras in it for auto attendance system. If the surveillance cameras are set in the classroom, then using Python's Open CV library we can easily detect Nodding, Notes taking, Answering, Questioning properties of every student. And we already have smart system which means that the results of students would be available online for a particular subject. We can obtain the marks of a student in a subject and we get the results from Open CV and easily we can identify the learning style of a student. And the smart educational system must already be having several courses to opt. Based on the automated data we received from Students through several automated ways, it's lot easier to identify the type of learner. In future we are intending to use this above- mentioned approach and see the results for it. If there are no smart classrooms, we can get a surveillance camera and start researching on it furthermore.

Bibliography

1. Yu Haixia., et al. "Artificial Intelligence-Based Quality Management and Detection System for Personalized Learning". *Journal of Interconnection Networks* (2021): 2143004.
2. "Types of Learners in a college" (2021).
3. "Working of Naive Bayes Algorithm?" (2021).
4. Sáiz-Manzanares María Consuelo., et al. "Teaching and learning styles on moodle: An analysis of the effectiveness of using stem and non-stem qualifications from a gender perspective". *Sustainability* 13.3 (2021): 1166.

5. "Top 6 Regression Algorithms" (2021).
6. Monsalve-Pulido Julián., et al. "Autonomous recommender system architecture for virtual learning environments". *Applied Computing and Informatics* (2020).
7. Bourkoku Outmane., et al. "A personalized e-learning based on recommender system". *International Journal of Learning and Teaching* 2.2 (2016): 99-103.
8. AK Milićević. "E-Learning personalization based on hybrid recommendation strategy and learning style identification". *Computers and Education* 56.3 (2011): 885-899.
9. Dascalu Maria-Iuliana., et al. "A recommender agent based on learning styles for better virtual collaborative learning experiences". *Computers in Human Behavior* 45 (2015): 243-253.
10. Monsalve-Pulido Julián., et al. "Autonomous recommender system architecture for virtual learning environments". *Applied Computing and Informatics* (2020).
11. Dorça Fabiano A., et al. "An Automatic and Dynamic Approach for Personalized Recommendation of Learning Objects Considering Students Learning Styles: An Experimental Analysis". *Informatics in Education* 15.1 (2016): 45-62.
12. Noor Rizwana and Farman Ali Khan. "Personalized recommendation strategies in mobile educational systems". 2016 Sixth International Conference on Innovative Computing Technology (INTECH). IEEE, (2016).
13. Laksitowening K A., et al. "E-Learning Personalization Using Triple-Factor Approach in Standard-Based Education". *Journal of Physics: Conference Series* 801.1 (2017).
14. TA Syed and V Palade. "A Personalized Learning Recommendation System Architecture for Learning Management System" (2017).
15. Ramya Sree, P., et al. "Personalized e-learning system based on user's performance and knowledge: an adaptive technique". *International Journal of Recent Technology and Engineering* 8.4 (2019): 8695.
16. Sami Memoona., et al. "LERNEN, Connecting Professionals and Students". In the 14th Asia Pacific International Conference on Information Science and Technology (APIC-IST 2019), Beijing, China, June 2019.
17. Wei Xin., et al. "Personalized Online Learning Resource Recommendation Based on Artificial Intelligence and Educational Psychology". *Frontiers in Psychology* 12 (2021): 767837-767837.
18. Bulger Monica. "Personalized learning: The conversations we're not having". *Data and Society* 22.1 (2016): 1-29.
19. Zhou Yuwen., et al. "Personalized learning full-path recommendation model based on LSTM neural networks". *Information Sciences* 444 (2018): 135-152.
20. SAddAR SALAHuddin., et al. "Evaluating Performance of Hibernate ORM based Applications using HQL Query Optimization". *Oriental Journal of Computer Science and Technology* 11.2 (2018): 115-125.
21. Qureshi Umair Mujtaba., et al. "Indoor localization using wireless fidelity (WiFi) and bluetooth low energy (BLE) signals". 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE). IEEE, (2019).

Assets from publication with us

- Prompt Acknowledgement after receiving the article
- Thorough Double blinded peer review
- Rapid Publication
- Issue of Publication Certificate
- High visibility of your Published work

Website: www.actascientific.com/

Submit Article: www.actascientific.com/submission.php

Email us: editor@actascientific.com

Contact us: +91 9182824667