



E Nurse – Student Health Risk Analysis and Personalized Evidence-based Diet Recommendation System with Machine Learning Techniques

Balaji V¹, P Sundresh², Pushpak L³, Saketh Guptha S B⁴, Shrivatsa Chitgopkar⁵

Assistant Professor, Computer Science and Engineering, The National Institute of Engineering, Mysore, India¹

Student, Computer Science and Engineering, The National Institute of Engineering, Mysore, India²⁻⁵

Abstract: In the past and present generation, the change in lifestyle of young people, less number of physical activities and eating fast food led to an alarming increase of diseases especially with college students. The unhealthy diet of students is responsible for unbalanced weight. The body mass index is used to classify a person's mass or weight as underweight, normal weight, overweight and obese based on tissue mass (bone, fat, and muscle) and height. Underweight leads to malnutrition, vitamin deficiencies, decrease in immune system function, growth issues and development issues. Major risk of being overweight leads to type-2 diabetes. Machine learning algorithms can solve many health problems. The project suggests the creation of an electronic college nurse that can perform the same functions as a human college nurse. Efficient Machine learning algorithms such as Naïve Bayes, Random Forest or Decision Tree used to build ML models to train the educational datasets and predicts student health problems. All these algorithms' results are analysed and compared to find the best algorithm to predict student health risk. Many educational factors such as age, gender, food habits, family issues, pressure, academic results, height, weight etc.... are considered to predict student mental health problems. In our project work we build an application software with ML model that can predict the student health problems based on health factors and also recommends suitable diet plan. System aims at providing personalized, evidence-based diet recommendations for the students. Project also analyses trends in health metrics and enabling dynamic, context-aware advice for the students. Proposed system is a real time medical system useful for educational sector and students and built using Microsoft tools such as Visual Studio tool and SQL Server tool.

Keywords: Data science, Naïve Bayes, SVM, Random Forest, GUI, Student, Mental Issues, Machine learning, fertilizer.

I. INTRODUCTION

In India and other Asian countries, we Every Educational institute wants 100% academic performance. Basically, student academic performance depends on so many factors like gender, communication skills, grasping capability, health issues, family issues, extra circular activities etc.... one should understand the student behaviours or health issues and can come up with precautionary measures to avoid less academic marks and over all performances. Results of every student will vary based on their skills and based on their health conditions. It is very important to identify student health problems or issues at early stages to take some measures to overcome the health issues and to improvise the academic performances. This system applies supervised machine learning (ML) techniques to find the student health risks/problems.

Basically, Students often experienced difficulties in learning an introduction to any academic course. This proposed work attempts to investigate how successful students regulate their learning in this course. The answer to these questions will provide teachers with useful information to better comprehend how students learn, and which strategies are effective in learning. Efficient Machine learning algorithms such as *Naïve Bayes*, *Random Forest* or *Decision Tree* used to build ML models to train the educational datasets and predicts student health problems. All these *algorithms' results are analysed and compared to find the best algorithm* to predict student health risk. Many educational factors such as *age, gender, food habits, family issues, pressure, academic results, bmi, height, weight* etc... are considered to predict student mental health problems. In our project work we build an application software with ML model that can predict the student health problems based on health factors and also recommends suitable diet plan. System aims at providing personalized, evidence-based diet recommendations for the students. Project also analyses trends in health metrics and enabling dynamic, context-aware advice for the students.

II. METHODOLOGY

Analysis of Student Health Factors and Risk Prediction Process

- Training the Datasets
- Building Model (ML)
- Model Evaluation (Confusion Matrix)
- Selecting best Model
- Result Analysis
- Data Visualization
- Line Graph, Pie Chart (outputs)

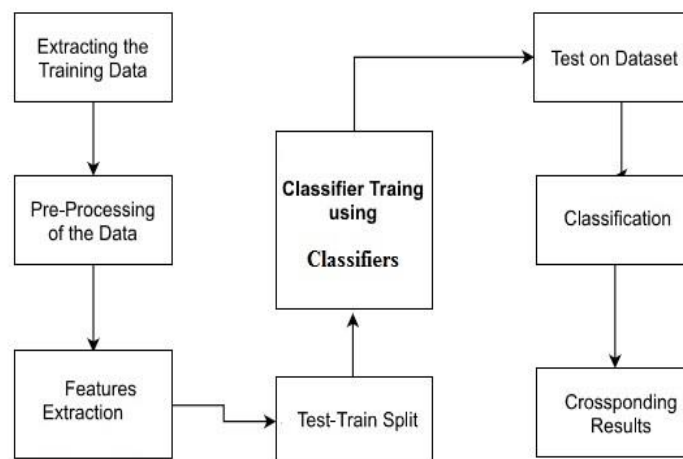


Fig: 1 Student Health Risk Prediction Process Flow Diagram

Step 1. Required data extracted from the server. In our project we extract student health datasets (training datasets) from the server.

Step 2. data pre-processing is done, where we remove the irrelevant data and extract the required data for processing. In our project irrelevant data means student id, name, mobile etc... all these are irrelevant data.

- Step 3.** once data pre-processing is done, desired data is inputted to the efficient supervised learning algorithms such as Random Forest and Naïve Bayes/Decision Tree algorithms for processing.
- Step 4.** Algorithms process the educational data and finds the factors associated with student health and predicts student's health risk/problem.
- Step 5.** Algorithms are tested and results are compared to find the best algorithm.
- Step 6.** Efficiency of both algorithms compared, and best algorithm is chosen.
- Step 7.** Using that best algorithm, student health outputs on graphical user interface.

III. EXPERIMENT RESULTS

3.1 RF ALGORITHM RESULTS

DISCUSSION

Here we build a real time application useful for the educational department. This project build using Microsoft technologies. Educational Training datasets trained using RF algorithm and we got very good results. RF algorithm is programmed in such a way that, it works for dynamic datasets. RF algorithm logic is written and it's our own library. We are getting around 88% of accurate results and it takes around 1000 milli seconds for prediction.

Constraint	NB Algorithm
Accuracy	88 %
Time (milli secs)	1006
Correctly Classified (precision)	88 %
Incorrectly Classified (Recall)	12 %

IV. CONCLUSION

Lack of health consciousness, food diet and physical activities in college students may lead to health issues/problems. Analysis of student health risk/problems such as stress, depression, anxiety, type 2 diabetes etc... Plays a challenging role in the current education sector. It is important to rectify the health factors associated with the student health risk. Identification of student health risks at early stages and providing suitable personalized evidence based diet recommendations for the students is one of the critical factor in the current college sector. Proposed system is an automation for student health risk prediction using ML algorithms. We implemented 2 ML models and got very good results. In real time our system can be applied for colleges.

V. FUTURE ENHANCEMENTS

- We can add more ML algorithms and predict student health risk and compare algorithms to find best algorithm.
- We can add more factors for prediction.
- We can increase the size of the datasets.

**REFERENCES**

- [1]. Research on the College Students' Psychological Health Management based on Data Mining and Cloud Platform, Fang Li, 2016.
- [2]. PREDICTIVE ANALYSIS OF STUDENT STRESS LEVEL USING NAÏVE BAYESIAN CLASSIFICATION ALGORITHM. Monisha S, Meera R. Vijay, Swaminath.R, Dr. Arun Raj L, 2020.
- [3]. Machine Learning Techniques for Stress Prediction in Working Employees, U SRINIVASULU REDDY, ADITYA VIVEK THOTA, A DHARUN, 2020. Y. Yuan and F.-Y. Wang, "Blockchain: The State of the Art and Future Trends", *Acta Automat. Sin.*, vol. 42, no. 4, pp. 481-94, 2016.
- [4]. Title: Classification Algorithms based on Mental Health Prediction using Data Mining, Vidit Laijawala, Aadesh Aachaliya, Hardik Jatta, Vijaya Pinjarkar, 2020.
- [5]. Agrawal R, Swami A, "Mining association rules between sets of items in large databases," *Acm Sigmod International Conference on Management of Data*, ACM, pp. 207-216, 1993.
- [6]. Koh, H. C., and Tan, G., "Data mining applications in healthcare," *Journal of Healthcare Information Management Jhim*, vol. 19, pp. 64-72, 2005.
- [7]. Yan, R W. Cong, L I. Lei and Y Y, "Improvement of Apriori Algorithm Based on Association Rules and the Application in the Insurance CRM of China," *Science Technology and Engineering*, vol. 21, pp. 6469-6473, 2009.
- [8]. Pradhan, G. N, and Prabhakaran B, "Association rule mining in multiple, multidimensional time series medical data," *Journal of Healthcare Informatics Research*, vol. 1, pp. 92-118, 2017.
- [9]. Jiang, H. B, and Yang, D L, "Application Research on Fast Discovery of Association Rules Based on Air Transportation," *International Conference on Service Systems and Service Management*, pp. 1-6, 2007. 723
- [10]. Liu G, Jiang H, Geng R, and Li H, "Application of multidimensional association rules in personal financial services," *International Conference on Computer Design and Applications*, vol. 26, pp. 3877-3879, 2010.
- [11]. Qi W, Yan J, Huang S, Guo L, and Lu R, "The application of association rule mining in college students' mental health assessment system. *Journal of Hunan University of Technology*," vol. 6, pp. 94-99, June 2013.
- [12]. Meng Q, and Sha J, "Tree-based frequent item sets mining for analysis of life-satisfaction and loneliness of retired athletes," *Cluster Computing*, vol. 2, pp. 1-9, May 2017.