

“A Real time Application for Student Stress Level Prediction in education sector using Bayesian-KNN (NB with KNN) Algorithm”

^[1]Manoj Kumar V, ^[2]Mayur V, ^[3]Bharath M, ^[4]Khajabandur N M
Information and Science and Engineering, Malnad College of Engineering
Hassan-573202, India

Email id:manojkumarv724@gmail.com, abj@mcehassan.ac.in, mayurkainij@gmail.com, mbharath0326@gmail.com
mittalakodkhajabandu@gmail.com

Abstract— College understudies are enduring from numerous mental wellbeing issues counting mental push, somatization, fanatical, interpersonal affectability, discouragement, uneasiness, threatening vibe, fear, distrustfulness and mental disturbance, which can bring a parcel of negative impacts to them. Clearly, the mental prosperity issues of college understudies not because it were particularly impact their have advancement, but as well impact the consistent quality of the campus. We require a system to handle the understudy mental wellbeing issues. Here we are mainly targeting on predicting levels of stress . We can observe several variables related to mental issues such as age, sex, work culture, assignments, conflicts within family, companions issues, participation, education etc. Utilizing machine learning strategies we anticipate understudy stress and levels. In the project we used two administered learning algorithms such as knn and bayes algorithm and applied combination of these for prediction and got better results with an accuracy 94.5% which is more compared to 80.5% and 85.7% for KNN and bayes respectively, we named that combination as hybrid algorithm. Moreover our system is going to provide some useful suggestions based on the stress levels of understudy. In this proposed system we develop automation for education sector for student stress and levels prediction.

Keywords— *Stress, Educational Data, Bayes, KNN, Data Science, Machine Learning, Administered Learning, Student performance, Academic.*

I. INTRODUCTION

Stress is regarded as a problem because it creates an imbalance in the life of every character and it is additionally regarded as a serious issue for psychological adjustments and trauma reduction. All Colleges basically analyze whether understudies have mental wellbeing issues or what kind of issues they have within the understudies. Most colleges moreover pay increasingly consideration to students' mental emergency checking and anticipation. Some Colleges usually analyze whether students have mental issues or disorders or what kind of trouble they are going through. Some colleges conducts manual counseling where lecturers sit with students and try to identify mental health disorders for low academic performances . These methods cannot discover covered up connections within the mental information. Since it is a psychological data it cannot find hidden relationships [3]. We require a framework to handle the understudy mental wellbeing issues, here we are primarily concentrating on understudy push forecast. There are so numerous components related to push such as work stack, assignments, family issues, companions issues, participation, instructing etc. ML may be a domain to detect future based on

what happened earlier. Utilizing ML techniques we anticipate understudy levels of stress conjointly proposed framework will grant recommendations depending on the stress levels of understudy [6].

Predicting the student stress at early stages is a tough task in the current medical sector. The prediction of student prone to stress can help medical sector to take proper decision regarding their treatments. Stress prediction has become a global health issue and is an area of concern. Student Stress needs to be recognized at early stages and current works takes much time for diagnosis and treatments. Our system is automated for stress levels prediction using supervised learning techniques. As we use ML algorithms for stress prediction, we will get more accurate and efficient results. System useful for educational sector and doctors to take faster and better decisions. All the input parameters we are taking are more relevant for stress prediction to get better results and we are using around 1600 records which is obtained from Kaggle however specific source link is currently unavailable.

II. RELATED WORKS

This section summarizes and compares existing studies on stress levels prediction using AI/ML models. The studies reviewed focus on the algorithms, datasets, and methodologies used to classify and predict student stress.

A. *Paper Title: Classification Algorithms based Mental Health Prediction using Data Mining Authors: Vedit Laijawala, Aadesh Aachaliya, Hardik Jatta, Vijaya Pinjarkar Year: 2020 [1]*

Description: The emotional, psychological and social welfare of an individual is revealed by their mental health. It influences how an individual will think, feel and face a situation. Different components contribute to mental prosperity issues which lead to mental ailment like thrust, social uneasiness, debilitation, over the top compulsive clutter, steady subjugation, work environment issues and character disorders.

Limitations: The above mentioned work has used very small datasets which leads to less accurate results. For mental health prediction data mining technologies used which do not predicts properly. Huge data is required.

B. Paper title: Predictive Analysis of Student Stress level using naïve bayesian classification algorithm Authors: Monisha S, Meera R, Vijay Swaminath R, Dr Arun Raj L Year: 2020 [3]

Description: Overall academic performance and social compulsion have created a pressurized psychological state for students. Frequently stated stress factors should be reduced to help the students to excel in their academic performance and activate the social activities, thereby decreasing individual health issues like migraine headaches, spectacle wearing and so on.

Limitations: This concept predicts stress of only college students. Only ML model built, not suitable for real time. Algorithms used here takes more time processing. Less efficient results.

C. Paper Title: Machine Learning Techniques for Stress Prediction in Working Employees Authors: U srinivasulu reddy, aditya vivek thota, a dharun Year: 2020 [2]

Description: Now a days Mental disorders are usual problem amongst operating IT experts within the enterprise those days. With converting manner of lifestyle and paintings work cultures, there`s an growth within the threat of stress among the employees. Here, they used ML strategies to analyze extend plans in working grown-ups and to contract down the factors that unequivocally choose the thrust levels.

Limitations: Less parameters used for predicting levels of stress. Boosting technique is not suitable for GUI based and real time application. Ready Tools used for stress prediction.

D. Research on the College Students' Psychological Health Management based on Data Mining and Cloud Platform Authors: Fang li Year: 2016

Description: In later a long time, college students' mental issues are getting to be increasingly noticeable, and due to the need of high-quality assets in mental wellbeing instruction, it is troublesome. In this paper, the author analyses the college students' psychological and mental health management based on data mining and cloud platform.

Limitations: Applicable for education sector not for real world. Less parameter used in the project. Less efficient results.

Constraints	[1]	[2]	[3]	[4]	Our System
Less Parameters used	✓	✓	✗	✓	More factors will be used
Less Datasets used	✗	✓	✓	✓	Huge datasets will be used
Real Time Implementations	✗	✗	✗	✗	✓
Dynamic Data	✗	✗	✗	✗	✓
Prediction of levels of stress	✗	✗	✗	✗	✓
Solution Recommendations	✗	✗	✗	✗	✓

TABLE I. COMPARATIVE TABLE

III. PROPOSED WORK

Since lifestyle and work cultures changes everyday, there is an increase in risks related to mental health and stress among the college students. System finds parameters that accurately determine the levels of stress. Levels were identified based on sex, age, family issues and availability of any benefits in education sector. By identifying the accurate levels of stress we can come up with approaches to reduce levels of stress and create a good and peaceful learning environment for students. System uses ML algorithms or AI algorithms to find stress of a student. System can be developed as a real time application which is useful for colleges. As Visual Studio and SQL Server is more supportive with real time application, we use these technologies for application development. Finding the risk factors which affects the student's mental status is the major motive of the system.

3.1 Advantages of Proposed system:

Our proposed System analyses and predicts the student stress with better results. Faster Decision Making is possible. Creates Neat working atmosphere in the organization. We can predict the students who are finding difficulty with stress in less time and can come up with some solutions to overcome the stress. Useful in real time (any type of organization).

3.2 Scope of the Project:

Project is an browser or web based app which can be used through extensions like google chrome, opera, mozilla firefox etc after deployment. This system is an UI based application which requires internet to access in real time. Project can be used in real time which will be useful for educational sector. System accessed by lecturers and students in real time. System works based on the educational related parameters to predict stress and its levels. System helps education sector to get better academic results and student performances.

IV. PROPOSED METHODOLOGY

Step 1: Information Collection (Stress Information)

Typically the primary step within the methodology where we gather stress information. Information collected from numerous sources which includes parameters like Sex, Age, Monetary Issues, Family Issues, Wellbeing Issues, Favoritism Settle, Weight, Customary, Interaction etc.

Step 2: Information Arrangement

Here stress information analyzed and as it were important information extricated. Since all features were required feature extrication not done as it was relevant to our work. All features mentioned used for prediction of levels of stress. Binning method is used for prior processing where we removed irrelevant data and fixed missing values.

Step 3: Finalize Constraints

Inputs utilized for stress level expectation are gotten. Parameters such as Sexual orientation, Age, Monetary Issues, Family Issues, Wellbeing Issues, Prejudice Settle, Weight, Normal, Interaction etc.

Step 4: ML Calculations- Administered Learning

Administered learning is an approach to ML that's based on preparing information that incorporates anticipated answers.

Bayes based learning Algorithm and KNN algorithm

Algorithm We considered above algorithms because of the following reasons

1. Efficient classifier
2. Works fine for restricted no of inputs as well as huge number of parameters.
3. Works fine for little data-set as well as huge data-set.
4. More exact outcomes

Step 5: Stress level forecasting

Framework predicts the levels of stress based on the inputs utilizing ML calculation. We utilize 2 distinctive calculations for stress expectation “KNN calculation and Bayes Rule Algorithm”.

Step 6: Results

Proposed system predicts stress or stress free with levels, we have class labels of results as 0 which indicates stress free, 1 indicates 25% stress, 2 indicates 50% stress and 3 indicates 100% stress level. K fold used, tested for all possible ratios 50:50, 60:40, 70:30 and for 90:10 we got best results.

Step 7: Visual Representation

All the outputs and predictions will be displayed for users in GUI

V. ALGORITHMS USED

5.1 KNN Algorithm Steps:

1. Determine K (no of nearest neighbors)
2. Calculate distance (Euclidean, Manhattan)
3. Determine K-minimum distance neighbors
4. Gather category Y values of nearest neighbors
5. Use simple majority of nearest neighbors to predict value of query instance

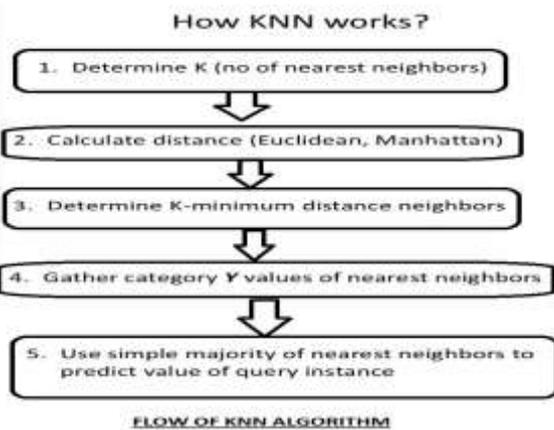


Fig. 1. KNN Algorithm Flow

Pseudo-code for KNN

- Training algorithm
 - For each training example $\langle x, class(x) \rangle$, add the example to the list *Training*
- Classification algorithm ($R^n \rightarrow V$)
 - Let $V = \{v_1, \dots, v_l\}$ be a set of classes
 - Given a query instance x_q to be classified
 - Let $X = \{x_1, \dots, x_k\}$ denote the k instances from *Training* that are nearest to x_q
 - $\forall i: 1 \dots l \quad vote_i = \{x \in X \mid class(x) = v_i\}$
 - Return v_j such that $|vote_j|$ is largest

These slides are based on Tom Mitchell's book 'Machine Learning'

Fig. 2. KNN Algorithm Pseudocode

5.2 Naïve Bayes Algorithm Steps:

1. Filter the data-set (storage servers): Recovery of necessary information for preparing from the servers such as database, cloud, exceed expectations sheet etc. Calculate the likelihood of each property esteem. $[n, n_c, m, p]$
2. Here for each trait we find the likelihood of event utilizing the taking after equation. (Specified within the following step). For each course (class) we ought to apply the equation.
3. Applying formula

$$P(\text{attribute value } (a_i) / \text{subject value } (v_j)) = (n_c + mp) / (n+m)$$

Where:

n = the no of preparing illustrations for which $v = v_j$

n_c = no of cases for which $v = v_j$ and $a = a_i$

p = a priori gauge for $P(a_{ij} v_j)$

m = the equivalent sample size

4. Multiply the probabilities by p : For each class, here we multiple the results of each instance with p and ultimate outcomes are used for classification. Compare the values and classify the attribute values to one of the predefined set of class.

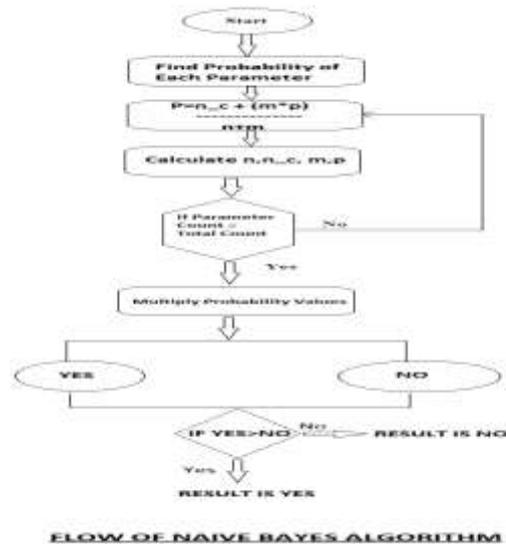


Fig. 3. NB Algorithm Flow

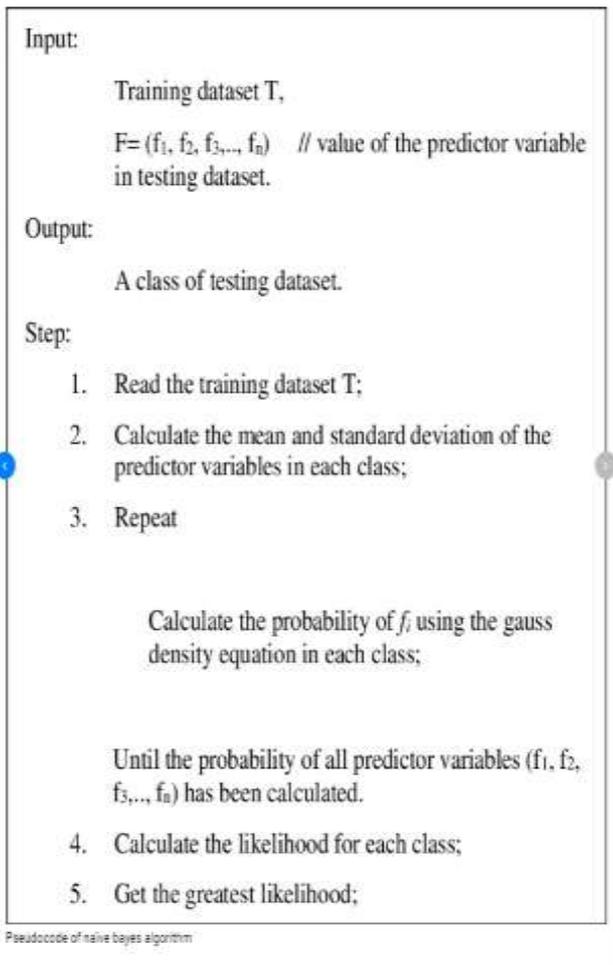


Fig. 4. NB Algorithm Pseudocode

5.3 Experimental Results:

KNN Algorithm Results:

Here we build a interactive system useful for the society. This project build using Microsoft technologies. Educational datasets trained using KNN algorithm. We are getting following results and it takes around 2000 milli seconds for prediction. We got 80.2% as accuracy which is less than bayes and hybrid results.

Constraint	KNN Algorithm
Accuracy	80.2%
Time (milli secs)	2106
Correctly Classified (precision)	80.2%
Incorrectly Classified (Recall)	19.8%

TABLE II. KNN ALGORITHM RESULTS

Naive Bayes Results:

Here Naive Bayes algorithm is programmed in such a way that, it works for dynamic datasets. Naive Bayes algorithm logic is written and it's our own library. We are getting following results and it takes around 21000 milli seconds for prediction. We are getting accuracy of 85.5% which is greater than knn but less than hybrid algorithm results.

Constraint	NB Algorithm
Accuracy	85.7%

Time (milli secs)	21006
Correctly Classified (precision)	8.2%
Incorrectly Classified (Recall)	15.3 %

TABLE III. NB RESULTS

Hybrid (NB with KNN) Algorithm results:

Here we build on the fly application helpful for the society. Educational datasets trained using Bayes based along with KNN algorithm and we got very good results. We are getting following results and it takes around 15000 milli seconds for prediction. We got the best results as accuracy is 94.5%.

Constraint	Hybrid Algorithm
Accuracy	94.5%
Time (milli secs)	15068
Correctly Classified (precision)	94.5%
Incorrectly Classified (Recall)	5.5 %

TABLE IV. HYBRID ALGORITHM RESULTS

Comparative Analysis:

As specified we got best results when we combined both algorithms with an accuracy of 94.5% compared to NB with 85.7% and KNN with 80.2%.

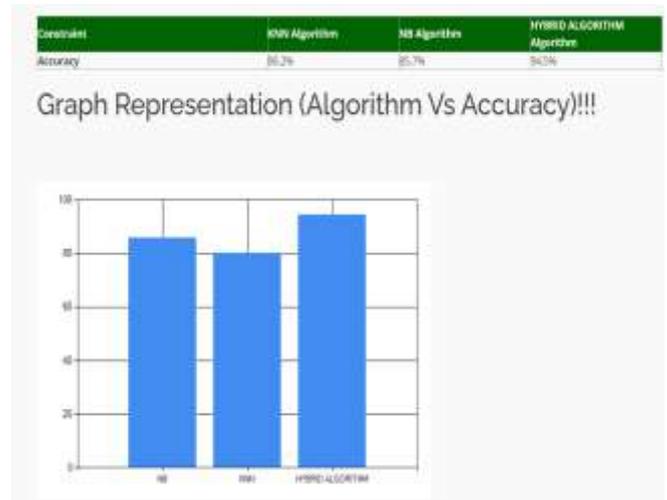


Fig. 5. COMPARISON GRAPH

VI. CONCLUSION

College students are enduring from numerous mental wellbeing issues counting mental trouble, somatization, obsessiveness, interpersonal sensitivity, uneasiness, anxiety, hostility, fear, paranoia and mental disturbance, which can have a number of bad effects on them. ML is about anticipating the long haul based on past data. We use ML techniques to predict student stress. There is also a need of a well developed real time application so that students can use the application to determine their levels of stress.

Future Enhancements:

Extra strategies just like the Bayes based classifier can be utilized to analyze the proficiency of the demonstrate. One can actualize profound learning strategies like CNN and confirm how the show performs for the given instances. A much more particular and tremendous datasets can be utilized as a preparing show since the no of reactions is constrained in our case.

REFERENCES

- [1] Paper Title: Predicting the stress level of students using Supervised Machine Learning and Artificial Neural Network (ANN), Authors: Suraj Arya, Anju, Nor Azuana Ramli, Year: 2024
- [2] Paper Title: Stress Detection in College Students Using Machine Learning Algorithm, Authors: Ms. Ancy Paul, Ms. Resija P R, Year: 2024
- [3] Paper Title: Predictive Analysis of Student Stress Level using Machine Learning, Authors: Dr. Anbarasi M, Sethu Thakkilapati, Veeragandham Rajeev Vas, Sarabu Venkata Bharath Viswath, Year: 2023
- [4] Paper Title: Stress Prediction in Indian School Students Using Machine Learning, Authors: Aanchal Bisht, Shreya Vashisth, Muskan Gupta, Ena Jain, Year: 2022
- [5] Galabawa, J. Lwaitama, A Community Secondary schools: How long is their Journey to Quality Education? Paper presented at the University of Dares Salaam, Tanzania, (2018).
- [6] Adler, A.K., and Wahl, O.F. Children's beliefs about people labelled mentally ill. American Journal of Orthopsychiatry, (2018)
- [7] Brockington, I.F.; Hall, P.; Levings, J. Murphy, C. The community's tolerance of the mentally ill. British Journal of Psychiatry, 162:pp93-99, (2017).
- [8] Corrigan, P.W.; River, L.; Lundin, R.K.; Uphoff Wasowski, K.; Campion, J.; Mathisen, J.; Goldstein, M.A. Stigmatizing attributions about mental illness. Journal of Community Psychology, (2010).
- [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 itemsets mining for analysis of life-satisfaction and loneliness of retired athletes," Cluster Computing, vol. 2, pp. 1-9, May 2017.
- [13] Huang S. C, Zhong J D, and Wen-Jua Q. I, "Statistical analysis and association rule mining of application in college students' mental health," Journal of Anyang Institute of Technology, pp. 108-111, 2014.
- [14] Long Z, Feng H, Xue D, and Xiangjun D, "Positive and Negative Association Rules Mining for Mental Health Analysis of College Students," Journal of Mathematics Science and Technology Education, vol. 13, pp. 5577-5587, 2017.
- [15] Herawan T, Vitasari P, and Abdullah Z, "Mining Interesting Association Rules of Students Suffering Study Anxieties Using SLP-Growth Algorithm," IGI Global, pp. 24-41, 2012.
- [16] Tella, Impact of motivation on student's academic achievement in mathematics among secondary school students. Eurasia journal of Science & Technology Education, vol 3, Issue 2, (2017).
- [17] Wellington, J, Educational research: Contemporary issues and practical approaches. London: Continuum. Wiki- Educator Teacher motivation. April (2016).
- [18] Ali1, S. Haider, Z., Munir1, F. Khan, H. & Ahmed, A, Factors Contributing to the Students' Academic Performance. A case study of Islamia University, SubCampus. Science and Education Journal, Vol. 1, Issue (8) pp1-10, (2017).
- [19] Evodia, S, Evaluation of the Guidance and counselling in secondary school is Tanzania. Unpublished Master dissertation. Open university of Tanzania (2015).
- [20] Galabawa, J. Lwaitama, A Community Secondary schools: How long is their Journey to Quality Education? Paper presented at the University of Dares Salaam, Tanzania, (2018).
- [21] Adler, A.K., and Wahl, O.F. Children's beliefs about people labelled mentally ill. American Journal of Orthopsychiatry, (2018)
- [22] Brockington, I.F.; Hall, P.; Levings, J. Murphy, C. The community's tolerance of the mentally ill. British Journal of Psychiatry, 162:pp93-99, (2017).
- [23] Corrigan, P.W.; River, L.; Lundin, R.K.; Uphoff Wasowski, K.; Campion, J.; Mathisen, J.; Goldstein, M.A. Stigmatizing attributions about mental illness. Journal of Community Psychology, (2010).