A REAL TIME APPLICATION FOR STUDENTS STRESS LEVEL PREDICTION IN EDUCATION SECTOR USING ML

Project Reference No.: *48S_BE_5857*

College : Malnad College of Engineering, Hassan
Branch : Department of Information Science and Engineering
Guide(s) : Dr Ananda Babu J
Student(s): Mr. Manoj Kumar V
Mr. Bharath M
Mr. Khajabandur N M
Mr. Mayur V

Keywords:

Naive Bayes, KNN, Stress, Student performance, Data Science.

Introduction:

Stress is a significant factor that causes imbalance in individuals' lives and is a major concern for psychological well-being. Colleges often evaluate student mental health but fail to uncover hidden patterns in psychological data. There is a need for a system to address student mental health, especially focusing on stress prediction. Student stress can arise from multiple sources such as workload, assignments, family, friends, attendance, and teaching. Machine learning enables prediction based on historical data and patterns. In this project, we use machine learning to predict student stress levels effectively. The system will also provide suggestions based on identified stress levels. Early detection of stress in students is a challenge for today's medical sector. Stress has become a global health issue and must be addressed proactively. Current diagnosis systems are time-consuming and less efficient. Our proposed Machine learning algorithms ensure higher accuracy and performance in stress analysis. This system benefits educational institutions and healthcare providers. We are using a dataset of around 5000 records to improve prediction accuracy.

Objectives:

1. Proposed system is a real time application
2. The model classifies the students into Stress or Stress Free

3. Proposed system gives better decision and also improvise the business.
4. Proposed system makes use of data science technique “classification rules” for predicting stress in college students.
5. Proposed system is meant for stress prediction.

Methodology:

The first step in the stress prediction process is data collection. Stress-related data is gathered from various sources and includes parameters such as Gender, Age, Financial Issues, Family Issues, Health Issues, Pressure, Regularity, and Interaction. This diverse dataset provides a comprehensive foundation for analysis. In the second step, data preparation is performed where the collected data is cleaned and only relevant, useful features are extracted. This is essential to ensure efficient processing, as using the entire dataset would be time-consuming and could include noise or irrelevant information. The data is then segmented according to the requirements of the system.

In the third step, constraints are specified. This involves selecting the most influential parameters for stress prediction. The same parameters identified during data collection are used again to define the model’s focus. The fourth step introduces machine learning algorithms using supervised learning. This approach involves training models with data that includes the correct output labels, allowing them to learn from the input-output mapping. The two algorithms used for stress prediction are Naïve Bayes and K-Nearest Neighbour (KNN).

These algorithms are chosen for several reasons: they are efficient classifiers, perform well with both small and large datasets, handle datasets with few or many parameters, and provide accurate results. The fifth step is stress prediction, where the trained model evaluates the given parameters and predicts the stress level of a student. Using both Naïve Bayes and KNN allows us to compare results and enhance overall performance. In the sixth step, we evaluate the accuracy of the system. The dataset is split into 90% training data and 10% testing data, which helps validate the performance of the model and identify potential improvements.

Finally, in step eight, visual representation is provided. The prediction results and suggestions are displayed to the user through a graphical user interface (GUI), making the system interactive and user-friendly.

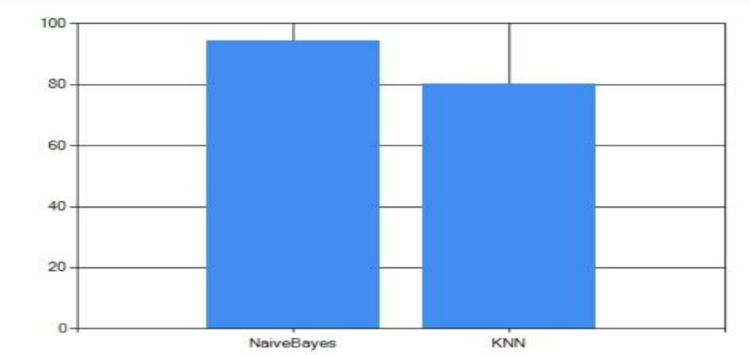
Result and Conclusion:

Here we build a real time application useful for the society. This project build using Microsoft technologies. Educational datasets trained using Naive Bayes algorithm and we got very good results with accuracy of 94.5%. 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 1500 milli seconds for prediction. KNN algorithm takes around 2000 milli seconds for prediction with accuracy of 80.2%.

COMPARE ALGORITHMS!!

Constraint	NB Algorithm	KNN Algorithm
Accuracy	94.5%	80.2%

Graph Representation (Algorithm Vs Accuracy)!!!



Project Outcome & Industry Relevance:

The project has practical implications in both the educational and healthcare sectors. By predicting student stress levels early, it helps institutions take proactive measures to support mental well-being. The system can be integrated into college or university portals to regularly assess student stress, enabling counselors and administrators to intervene before issues escalate. In the healthcare domain, it aids doctors and therapists in understanding the psychological health of students, making their diagnosis and treatment more data-driven.

Simulation/Study:

This project is primarily a simulation and study-based model rather than a physical working model. It involves the development and testing of a machine learning system using a dataset of around 5000 records. The focus is on algorithmic implementation, prediction accuracy, and interface-based output display, not on hardware components. The outcome is a software-based predictive system supported by a graphical user interface (GUI).

Project Outcomes and Learnings:

Key outcomes include the successful implementation of a stress prediction model using Naïve Bayes and KNN algorithms with high accuracy. The system is capable of classifying student stress levels and providing relevant suggestions. Through this project, we gained practical experience in data preprocessing, feature selection, model training, and validation. We also learned how to evaluate model performance and the significance of machine learning in solving real-world psychological and educational problems.

Future Scope:

In addition to the current algorithms, further exploration of Naïve Bayes and other supervised learning techniques can improve the system's performance. To boost accuracy and detect deeper data patterns, advanced methods like Convolutional Neural Networks (CNNs) can be implemented. CNNs are effective for high-dimensional data and can enhance prediction results. Testing different network architectures and activation functions may further optimize performance. The current model, trained on around 5000 records, may lack diversity. By expanding the dataset to include more varied samples, the system can generalize better across different student groups. Incorporating real-time data from wearables or mobile apps can also improve prediction timeliness. These improvements will help create a more accurate, scalable, and practical stress prediction system for real-world use in education and healthcare.