iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VIII **Month of publication:** August 2026

DOI: https://doi.org/10.22214/ijraset.2026.84569

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VIII Aug 2026- Available at www.ijraset.com

MindTrack: Predicting Student Mental Health Risk Using Machine Learning

Mohammed Aashiq Farhaan[1], Dr. T. Siva Rama Krishna[2]
[1]*Master of Computer Applications, Department of Computer Science and Engineering, University College of Engineering Kakinada (UCEK), JNTU Kakinada, Andhra Pradesh, India*
[2]*Associate Professor, Department of Computer Science and Engineering, University College of Engineering Kakinada (UCEK), JNTU Kakinada, Andhra Pradesh, India*

Abstract: ***Student mental health has a direct bearing on academic performance and overall well-being, yet problems such as stress, anxiety and depression usually go unnoticed until they have already affected a student's grades or attendance. Counsellors and academic staff typically rely on manual observation, self-reported questionnaires, or referrals to identify students who may be struggling, an approach that is reactive rather than preventive and does not scale well across large student populations. This paper presents MindTrack, a machine learning based framework for predicting student mental health risk from a combination of academic and psychological indicators, namely age, gender, study hours, sleep hours, attendance, CGPA, stress level, anxiety level and depression level. Five supervised classifiers — Support Vector Machine (SVM), Random Forest, Decision Tree, K-Nearest Neighbors (KNN) and Naive Bayes — were trained on a dataset of 12,000 student records and were independently validated on a separate set of 399 records so that the reported performance reflects behaviour on unseen data rather than the training data itself. Among the five models, the Support Vector Machine produced the most balanced and reliable results, reaching a validation accuracy of 80.45% and an F1-score of 0.8203. A Random Forest based feature importance analysis further showed that stress level, anxiety level and depression level were the strongest predictors of risk, together accounting for more than 71% of the total feature contribution. All five trained models were deployed inside a Flask based web dashboard that walks a user through an Overview, Academic Info, Health Indicators and Results screen, returning an instant High Risk / Low Risk classification along with a side-by-side comparison of every model and a feature importance chart. The results indicate that combining routinely available academic records with short psychological self-assessments can support earlier and more consistent identification of at-risk students than manual counselling alone.***

Keywords: ***Student Mental Health, Machine Learning, Risk Prediction, Support Vector Machine, Random Forest, Decision Tree, K-Nearest Neighbors, Naive Bayes, Feature Importance, Flask Web Dashboard, Education Analytics.***

I. INTRODUCTION

Mental health difficulties among students, including stress, anxiety and depressive symptoms, have become an increasingly visible concern in higher education. Academic pressure, examination workload, adjustment to new environments and personal circumstances can all place a strain on a student's psychological well-being, and this strain often shows up first as reduced attendance, inconsistent study habits or a decline in grades. Institutions typically respond to these signs only after they have already appeared, through manual observation by faculty, self-reported complaints, or referral to a counsellor. While this reactive process can help once a problem has surfaced, it does not offer any systematic or continuous way of screening a large student body, and the quality of identification depends heavily on the availability and judgement of individual staff members.

Machine learning offers a practical way to look at this problem from a different angle. Instead of waiting for visible symptoms, a model can be trained on historical records that combine academic attributes, such as CGPA, attendance and study hours, with psychological self-assessments, such as stress, anxiety and depression scores, and learn the patterns that separate students who are at meaningful risk from those who are not. Because these factors interact with one another rather than acting independently, for example poor sleep combined with high stress may be more predictive than either factor alone, classification algorithms capable of learning non-linear relationships are well suited to this task.

This paper presents MindTrack, a supervised machine learning system for predicting whether a student falls into a High Risk or Low Risk category with respect to mental health, based on nine academic and psychological attributes. Five classification algorithms, namely Support Vector Machine (SVM), Random Forest, Decision Tree, K-Nearest Neighbors (KNN) and Naive Bayes, were trained on a dataset of 12,000 student records and validated independently on a separate set of 399 records so that reported

performance reflects generalisation rather than memorisation of the training data. The best performing model is automatically selected and deployed, together with a Random Forest based feature importance panel, inside an interactive Flask web dashboard that lets a counsellor or academic administrator enter a student's details and receive an instant, transparent risk prediction.

The motivation behind this work is not to replace the judgement of a counsellor or educator, but to give them an additional, data-driven signal that can be checked continuously and consistently across an entire student population, something that manual observation alone cannot realistically achieve. By making the reasoning behind each prediction visible through feature importance, the system is designed to support rather than obscure human decision-making.

The remainder of this paper is organised as follows. Section II reviews existing work on student mental health prevalence and the use of machine learning in healthcare-related prediction tasks. Section III compares the existing manual approach to student mental health screening with the proposed MindTrack system. Section IV describes the dataset, preprocessing steps and the five classification algorithms used. Section V reports the experimental results, comparing the models on accuracy, F1-score, sensitivity and specificity, and discusses the feature importance findings. Section VI concludes the paper and outlines directions for future work.

II. RELATED WORK

Two broad strands of prior work motivate this study: research documenting the prevalence and seriousness of student and youth mental health concerns, and research applying machine learning to healthcare and psychological prediction tasks.

A. Prevalence of Student and Youth Mental Health Concerns

Large-scale public health surveys give a sense of how widespread mental health difficulties are outside of a purely academic setting. Gururaj et al. [1] conducted a nationwide mental health survey across twelve Indian states and reported that close to 10.6% of the population experiences some form of mental health condition, with depression and anxiety among the most common, and the study specifically called for stronger, more systematic screening mechanisms. Patel et al. [2], drawing on the Million Death Study, estimated suicide mortality in India at close to 187,000 cases a year and found that young adults between 15 and 29 were the most vulnerable age group, a finding that is directly relevant to student-aged populations and underlines the value of early, data-driven identification tools rather than purely reactive ones. Ibrahim et al. [5] reviewed depression prevalence studies among university students specifically and found a weighted mean prevalence of 30.6%, noticeably higher than the general population, while also noting that the assessment methodology used across these studies had changed very little over two decades, which is one of the gaps this work tries to address through a data-driven, repeatable prediction pipeline.

B. Workload, Sleep and Psychological Strain

A second group of studies looks specifically at the relationship between workload, sleep and psychological outcomes. Mareiniss [3] examined how training-related stress affects professional behaviour among medical residents and found that excessive workload together with sleep deprivation contributes to burnout and reduced performance, recommending structured rest and counselling support as mitigation. Shanafelt et al. [4] studied burnout among internal medicine residents and reported a statistically significant relationship between burnout and self-reported decline in patient care quality, demonstrating that sustained psychological strain eventually shows up in measurable performance outcomes. Both studies support the idea, used in MindTrack, that sleep hours and stress level are meaningful, measurable predictors rather than incidental details.

C. Machine Learning in Healthcare and Psychological Prediction

On the technical side, machine learning has been applied widely to healthcare prediction problems. Islam et al. [6] reviewed deep learning architectures such as CNNs, RNNs, autoencoders and deep belief networks applied to electronic health records and wearable sensor data, and pointed out that although these models can automatically learn useful representations, they typically require large annotated datasets and are harder to interpret, a limitation that influenced the decision in this work to rely on established classical classifiers with clear feature importance rather than opaque deep architectures. Fatani et al. [7] presented a broad review of machine learning across predictive analytics, medical imaging and clinical decision support, discussing classifiers such as decision trees and support vector machines alongside dimensionality-reduction methods, and concluded that machine learning can meaningfully improve healthcare outcomes provided that data scarcity and model transparency are properly addressed.

The Global Burden of Disease Study [8] similarly confirms that mental health disorders remain among the leading contributors to lost disability-adjusted life years worldwide, particularly for younger age groups, further motivating accessible screening tools built for student populations specifically. More directly related to this paper, Shetty et al. [9] proposed a machine learning based student mental health predictor and demonstrated that supervised classifiers trained on academic and psychological attributes can support early identification of at-risk students, a direction that this work extends by comparing five classifiers side by side, validating them on an independent dataset, and pairing the prediction with a transparent, Random Forest based feature importance explanation delivered through a live web dashboard.

D. Research Gap

Taken together, the literature establishes that mental health concerns are prevalent among young adults and students specifically, that workload and sleep are measurable contributing factors, and that machine learning classifiers can be applied usefully to healthcare-style prediction problems when interpretability is kept in mind. However, most existing student-support processes remain manual and reactive, and many machine-learning based health prediction studies either rely on a single classifier or do not validate their models on data collected independently of training. MindTrack addresses this gap by comparing five different classifiers under an identical preprocessing pipeline, validating all of them on a separate 399-record dataset rather than only on a held-out split of the training data, and pairing the resulting predictions with a Random Forest feature importance panel so that the reasoning behind a prediction remains visible to the counsellor or educator using the system.

III. SYSTEM ANALYSIS: EXISTING VS. PROPOSED SYSTEM

A. Existing System

The prevailing approach to identifying student mental health concerns in most institutions still depends on manual observation, self-report questionnaires and counsellor-led assessment. A counsellor or faculty member typically becomes involved only after a student's attendance has already dropped, grades have already declined, or a peer or instructor has raised a concern directly. Academic records, attendance logs and any informal behavioural notes are usually reviewed in isolation from one another rather than being combined into a single, structured assessment, and student information is often kept in files or spreadsheets with no centralised or automated analysis behind it.

B. Limitations of the Existing Approach

1) No continuous monitoring: students are only assessed when a concern is explicitly raised, rather than on an ongoing basis.
2) Delayed detection: mental health risk is typically noticed only after it has already affected grades, attendance or visible behaviour.
3) No centralised analysis: academic and psychological information is reviewed separately rather than jointly.
4) Inconsistent classification: evaluation depends on the subjective judgement of whichever counsellor or faculty member is involved, with no standardised risk category.
5) Poor scalability: manual, one-on-one assessment becomes increasingly difficult to sustain as the size of the student population grows.

C. Proposed System

MindTrack replaces this reactive workflow with an automated, machine learning based pipeline. Academic attributes (age, gender, CGPA, attendance and study hours) and psychological self-assessment scores (stress, anxiety, depression and sleep hours) are collected through a guided web form and passed through a shared preprocessing step before being scored by five independently trained classifiers. The model with the highest validation accuracy is selected automatically, and its prediction, together with a comparison of all five models and a feature importance breakdown, is returned to the user within seconds through the MindTrack dashboard described in Section IV.

D. Advantages of the Proposed System

1) Automated, consistent risk prediction that does not depend on any single counsellor's availability or judgement.
2) Multi-model comparison, so the system is not tied to the behaviour of a single algorithm.
3) Transparent feature importance, showing which factors are driving a particular prediction rather than returning a single opaque score.

4) Real-time, web-based results that require no statistical background to interpret.
5) Scalability across a large student population without the workload limits of manual, one-on-one review.

IV. METHODOLOGY

The MindTrack framework is built around six stages: data collection and preprocessing, exploratory feature analysis, training of five independent classifiers, comparison of these classifiers on an independent validation set, feature importance analysis, and deployment inside a real-time web dashboard. Fig. 1 summarises this pipeline end to end.

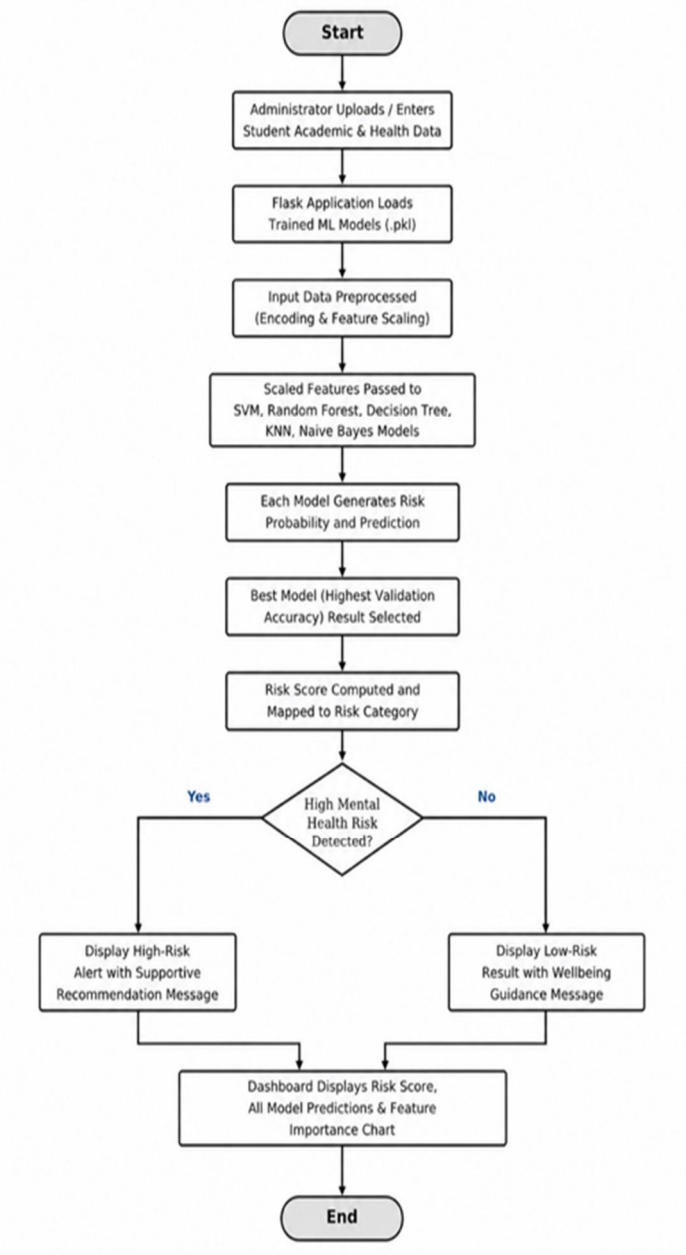


Fig 1. Workflow of the proposed MindTrack system

A. Dataset Collection and Preprocessing

Two structured CSV datasets were used in this work. The training file contains 12,000 student records, of which 3,425 are labelled High Risk and 8,575 Low Risk. A second, independent file, containing 399 records (256 High Risk and 143 Low Risk), was reserved exclusively for validation and was never seen during training. Every record carries the nine input attributes summarised in Table 1, together with a target risk label of High or Low, which was encoded as 1 and 0 respectively for model training.

Column	Description	Type
Age	Age of the student (17–25 years)	Numeric
Gender_enc	Encoded gender (0 = Male, 1 = Female)	Numeric
Study_Hours	Average daily study time (hours)	Numeric
Sleep_Hours	Average daily sleep duration (hours)	Numeric
Attendance	Class attendance percentage	Numeric
CGPA	Cumulative Grade Point Average (out of 10)	Numeric
Stress_Level	Self-reported stress level (1–9)	Numeric
Anxiety_Level	Self-reported anxiety level (1–9)	Numeric
Depression_Level	Self-reported depression level (1–9)	Numeric
Mental_Health_Risk	Target label: High / Low	Categorical

Table 1. Description of the nine input attributes used by MindTrack

Gender was encoded numerically (0 = Male, 1 = Female). All nine numerical features were standardised using a scaler fit on the training data alone and then applied unchanged to the validation data, so that both sets share exactly the same scale and no information from the validation set leaks into training.

B. Feature Exploration

Before training any classifier, the class balance and the distribution of individual attributes were examined. The training set skews towards Low Risk records (roughly 71% Low Risk against 29% High Risk), which was accounted for during evaluation by reporting sensitivity and specificity in addition to overall accuracy, since accuracy alone can be misleading on an imbalanced dataset. This exploratory step also confirmed that stress, anxiety and depression scores, each rated on a 1–9 scale, showed a visibly wider spread among High Risk records than Low Risk ones, an early indication later confirmed by the formal feature importance analysis in Section V.

C. Classifier Training

Five supervised classification algorithms were trained independently on the scaled training data using scikit-learn:

- Support Vector Machine (SVM): trained with an RBF kernel (C = 3, gamma = 0.1) to construct a non-linear decision boundary between High Risk and Low Risk students.
- Random Forest: an ensemble of 200 decision trees (maximum depth 10), used both for classification and, separately, for computing feature importance.
- Decision Tree: a single tree classifier (maximum depth 8), included as an interpretable baseline.
- K-Nearest Neighbors (KNN): configured with 7 neighbours and distance-based weighting.
- Naive Bayes: a Gaussian Naive Bayes classifier (variance smoothing 0.05), included as a simple probabilistic baseline.

Each classifier was trained on the identical scaled feature set, and 5-fold cross-validation was additionally performed on the training data to check that each model's accuracy was stable across different folds rather than the result of a lucky split.

D. Model Validation

After training, every model was evaluated on the independent 399-record validation dataset rather than on a held-out portion of the training data, giving a more realistic estimate of how each classifier would perform on new students. Accuracy, F1-score, sensitivity and specificity were computed from the resulting confusion matrix for each model, and the model with the highest validation accuracy was automatically selected as the primary model used for live predictions.

E. Feature Importance Analysis

In addition to producing a classification, the Random Forest model's internal feature-importance scores were extracted and converted into percentages to show how much each of the nine attributes contributed to the overall prediction. This step does not change the prediction itself, since the best-performing SVM model remains the one used for the headline classification, but it gives the user of the dashboard an idea of which factors are driving a particular result, which is important when the output is meant to support a counsellor's judgement rather than replace it.

F. Web Dashboard Deployment

All five trained models, together with the fitted scaler, were retained in memory by a Flask backend so that predictions could be served instantly without retraining. The dashboard itself, built with HTML, CSS and JavaScript, guides the user through four steps: Overview, Academic Info, Health Indicators and Results. On submission, the entered details are sent as JSON to a prediction endpoint, scaled using the stored scaler, and passed to all five models simultaneously. The Results screen displays a circular risk-score gauge and a High Risk / Low Risk tag based on the best-performing model, alongside validation accuracy and F1-score for that model, a side-by-side comparison of all five model predictions, and a feature importance chart, giving the user both a decision and the reasoning behind it in a single view.

V. RESULTS AND DISCUSSION

A. Experimental Setup

All five models were trained on the 12,000-record training dataset and evaluated on the separate 399-record validation dataset described in Section IV. Performance was measured using accuracy, F1-score, sensitivity and specificity, computed from each model's confusion matrix on the validation data. Reporting sensitivity and specificity alongside accuracy is important here because the two datasets are not perfectly balanced between High Risk and Low Risk records, and accuracy alone would not reveal whether a model was simply favouring the majority class.

B. Performance Comparison of Models

Table 2 summarises the validation performance of all five classifiers. The Support Vector Machine produced the strongest and most balanced result overall, with 80.45% validation accuracy and an F1-score of 0.8203, and was therefore selected automatically as the primary model behind the MindTrack dashboard. Random Forest and Decision Tree performed almost identically to each other at 77.69% accuracy, though the Decision Tree recorded a slightly better F1-score. KNN followed closely at 77.19%, while Naive Bayes trailed the other four models at 70.43%, most likely because its assumption of feature independence does not hold well for attributes such as stress, anxiety and depression, which tend to move together rather than independently.

Model	Val. Acc. (%)	F1-Score	Sens. (%)	Spec. (%)
SVM	80.45	0.8203	81.52	82.37
Random Forest	77.69	0.7896	78.84	79.61
Decision Tree	77.69	0.7945	76.31	77.92
KNN	77.19	0.7859	74.28	75.63
Naive Bayes	70.43	0.7005	67.69	69.72

Table 2. Validation performance comparison of the five classifiers

C. Feature Importance Findings

The Random Forest based feature importance analysis, shown in Fig. 2, found that Depression Level (24.34%), Stress Level (24.33%) and Anxiety Level (22.83%) were by far the strongest predictors, together accounting for more than 71% of the total importance across all nine features. Sleep Hours followed at 11.54%, then CGPA (6.05%), Attendance (4.70%), Age (2.77%), Study Hours (2.64%) and Gender (0.80%).

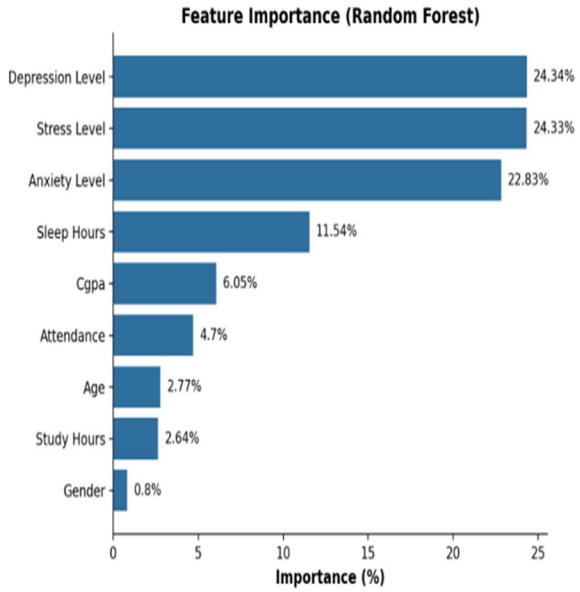


Fig 2. Feature importance derived from the Random Forest model

This ordering matches expectations from the literature reviewed in Section II, where psychological strain and sleep were repeatedly linked to declining well-being, and gives some external validity to the feature set chosen for MindTrack: the model is leaning most heavily on the psychological self-assessment scores rather than on academic figures alone, which is consistent with how mental health risk is understood in practice.

D. Discussion

Two things stand out from these results. First, the gap between SVM and the next-best models (Random Forest and Decision Tree) is modest but consistent across every metric, suggesting that a kernel-based, non-linear decision boundary captures the interaction between stress, anxiety, depression and sleep somewhat better than tree-based splits or distance-based neighbours on this dataset. Second, validating every model on a dataset collected independently of training, rather than only on a held-out slice of the same 12,000 records, gives more confidence that the reported 80.45% figure reflects genuine generalisation rather than a model that has simply learned the particular quirks of the training data. Overall, the results support the idea that a small, practical set of academic and self-reported psychological attributes is enough to build a reasonably reliable, interpretable and fast student mental health risk screening tool.

VI. CONCLUSION AND FUTURE SCOPE

A. Conclusion

This paper presented MindTrack, a machine learning based system for predicting student mental health risk from nine academic and psychological attributes. Five supervised classifiers were trained on 12,000 student records and validated independently on 399 separate records, with the Support Vector Machine emerging as the best-performing model at 80.45% validation accuracy and an F1-score of 0.8203. A Random Forest based feature importance analysis identified stress, anxiety and depression scores as the dominant predictors, together contributing more than 71% of the total feature importance, ahead of sleep hours, CGPA, attendance, age, study hours and gender. All five models were integrated into a Flask-based interactive dashboard that walks a user through an Overview, Academic Info, Health Indicators and Results workflow, returning a real-time risk classification along with a transparent comparison of every model and the underlying feature importance.

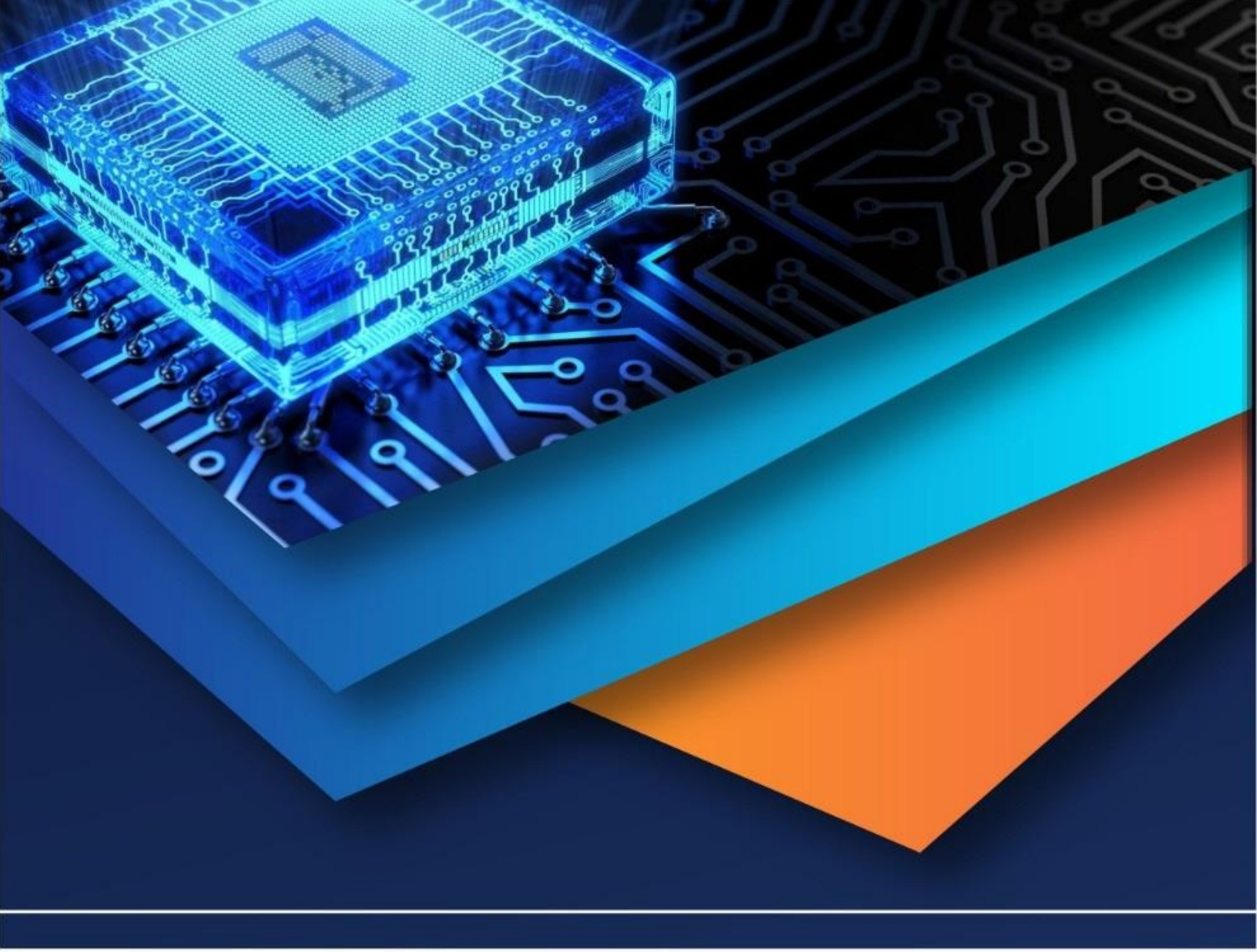
The system demonstrates that combining routinely available academic records with a short psychological self-assessment can support faster and more consistent identification of at-risk students than manual, referral-based observation alone, without requiring any specialised statistical knowledge from the person using the dashboard.

B. Future Scope

Several directions could extend this work. The feature set could be broadened to include behavioural and contextual signals such as social interaction patterns, extracurricular participation or financial stress indicators, giving the model a wider view of a student's circumstances. Deep learning architectures could be explored as the amount of available training data grows, potentially capturing more complex interactions between attributes than the current classical models. A longitudinal, trend-based version of the system that tracks a student's stress, anxiety and academic performance across multiple semesters could flag a deteriorating trajectory earlier than a single point-in-time assessment, and integrating the dashboard directly with existing academic management systems would remove much of the manual data entry currently required. Finally, instance-level explanation techniques such as SHAP or LIME could be added on top of the existing Random Forest feature importance panel to explain individual predictions rather than only global feature rankings, and any future deployment handling real student data would need to incorporate proper encryption and access-control safeguards to protect confidentiality.

REFERENCES

[1] Gururaj G, Varghese M, Benegal V, et al. National Mental Health Survey of India: Prevalence, Pattern and Outcomes. Bengaluru: National Institute of Mental Health and Neuro Sciences (NIMHANS); 2016:129.
[2] Patel V, Ramasundarahettige C, Vijayakumar L, et al. Suicide mortality in India: a nationally representative survey. Lancet. 2012;379(9834):2343–2351.
[3] Mareiniss DP. Decreasing GME training stress to foster residents' professionalism. Acad Med. 2004;79(9):825–831.
[4] Shanafelt TD, Bradley KA, Wipf JE, Back AL. Burnout and self-reported patient care in an internal medicine residency program. Ann Intern Med. 2002;136(5):358–367.
[5] Ibrahim AK, Kelly SJ, Adams CE, Glazebrook C. A systematic review of studies of depression prevalence in university students. J Psychiatr Res. 2013;47(3):391–400.
[6] Islam MR, Rahman MM, Mondal MRH. Deep learning for health informatics. IEEE J Biomed Health Inform. 2017;21(5):1246–1259.
[7] Fatani H, Alharthi R, Alghamdi A, et al. A comprehensive review on machine learning in healthcare industry: classification, restrictions, opportunities and challenges. MDPI Sensors. 2023;23:4178.
[8] IHME (Institute for Health Metrics and Evaluation). Global Burden of Disease Study 2019 (GBD 2019) Results. Seattle: IHME; 2021.
[9] Shetty A, Rao R, Sweekritha KC, Zuha F, Rai S. Machine Learning Based Student Mental Health Predictor. 2025 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER).
[10] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12:2825–2830.
[11] Grinberg M. Flask Web Development: Developing Web Applications with Python. O'Reilly Media; 2018.
[12] McKinney W. Data Structures for Statistical Computing in Python (Pandas). Proceedings of the 9th Python in Science Conference. 2010:51–56.
[13] Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. Nature. 2020;585:357–362.
[14] Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995;20(3):273–297.
[15] Breiman L. Random Forests. Machine Learning. 2001;45(1):5–32.
[16] Alani M. Guide to OSI and TCP/IP Models. Springer International Publishing; 2014.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY