

Zentora-An Adaptive AI for Personalized Learning on Mental Health

V.Ananthitha
Department of CSE,
Chennai,Tamil Nadu

Ashlin Rithzia GA
Department of CSE ,
Chennai,Tamil Nadu

Abstract- It is about mentalhealth concentrated on prognosticating pupil depression using a internal health dataset that includes demographic and academic information. We tested six machine literacy classifiers Logistic Retrogression, Decision Trees, grade Boosting, Random Forest, XGBoost, and Naive Bayes. The performance criteria — ranged from 0.88 to 0.97. grade Boosting achieved the loftiest ROC AUC of 0.97, showing that it has the stylish prophetic capability. Bar graphs and ROC angles show comparisons between the models. also, a Zentora chatbot, which uses the grade Boosting model, acts as a internal health assessment tool to give real- time help for scholars. This system helps identify depression early and encourages intervention strategies in seminaries. The exploration adds to internal health analytics by combining machine literacy ways with chatbot use for better and visionary pupil support

Keywords- Chatbot, Student Mental Health, Predictive analytics, Random timber, grade boosting, Machine literacy

I.INTRODUCTION

The internal health problems of scholars have come a pressing issue encyclopedically, and this affects not only their academic performance but also their social life and overall well- being. A large number of scholars have different pressures similar as depression, anxiety and stress that have an influence on their diurnal lives and academic career. In order to do data- driven exploration on similar important issues, the paper uses a comprehensive dataset of pupil internal health which was collected from surveys. Another type of information included within the set is demographic details including gender, age, nation or major course of undergraduate study(CGPA) but also life factors similar as frequency of exercise and eating habits. Above each, it covers measures that reflect internal health, similar as depression and anxiety

situations; fear attacks; the strategies scholars borrow when they want to manage with life's difficulties or stress events.

therefore this abundant information allows for thorough disquisition of the situation affecting scholars' internal health. Focusing on prognosticating the status of depression, this study employs several machine literacy models.

In addition to Logistic Regression, Categorical features are decoded, and also dataset into two data sets rigorous training and testing in order that we may directly estimate model performance. XGBoost, Random Forest, and grade Boosting are among these. Where scholars currently admit admissions treatment in the lot clinic, making movables weeks in advance and having to stay through long lines to see someone who may or may not be suitable to help them. Fortunately this liberal trades lot in south Los Angeles has been suitable to find Zentora apps for its converse computer, and has equipped them with GBDTs so as to incontinently estimate internal health.

II.LITERATURE REVIEW

It's getting clear that pupil minds need attention encyclopedically. The way they do in academy, get along with others, and just generally feel impacts their internal state. effects like tough classes, fitting in, alongside life changes constantly beget solicitude, stress, indeed sadness(Bayram & Bilgel, 2008)[1].

Research explores what impacts how scholars feel mentally. For case, one comprehensive look by Ibrahim and associates(2013)[2] revealed considerable depression likewise anxiety among council scholars. also, Stallman(2010)[3] refocused

out effective running of difficulties alongside managing pressure greatly influences internal well-being. How scholars live- whether they move their bodies, get enough rest, or have people around them- seems connected to how good they feel. Studies show working out can lessen passions of sadness, yet losing sleep tends to boost solicitude. Also, effects like a pupil's age, coitus, or what time they're in academy matter too; beginners constantly witness more pressure conforming to council life.

Detecting- indeed soothsaying- problems with internal health is now frequently done using machine literacy. For case, Shatte and associates showed how ways similar as Random Forest or Logistic Regression could directly spot signs of depression or anxiety by looking at a person's background, habits, and emotional state. also, D'Mello and Graesser delved using artificial intelligence to sense feelings as they be, intimating at ways we might keep tabs on someone's internal heartiness.

AI exchanges, like the Woebot created by Fitzpatrick and associates back in 2017[4], feel to help. also, Miner and others set up in 2020[5] that chatbots could connect with scholars offering quick internal health aid. This suggests openings for tracking well- being now, also stepping in when issues arise.

III.METHODOLOGY

The proposed system offers real- time emotion discovery and internal health support through an AI-driven chatbot. The setup is illustrated in Fig. 1 and includes data preprocessing, emotion bracket, threat assessment, and feedback for counselors.

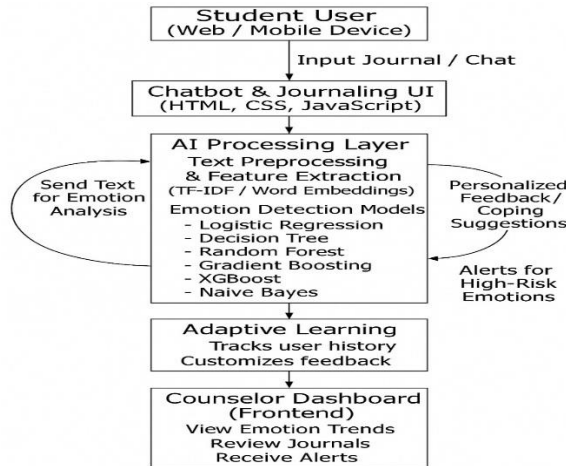


Fig.1: System Architecture of the AI-Based Student Mental Health Support System

Student Interaction Interface

scholars start by logging into the platform and entering message in the form of journals or converse dispatches. These inputs give raw data for language analysis.

NLP Preprocessing

Each communication goes through Natural Language Processing(NLP), which includes tokenization, removing stop words, stemming, and lemmatization. For an input judgment

$$S = \{ w1, w2, w3,, wn \}, (1)$$

the reused vector representation is attained through

$$xi = f(wi), (2)$$

where $f(\cdot)$ maps each commemorative to its embedding space using styles like Word2Vec or BERT embeddings.

Emotion Discovery Model

The point vector xi is anatomized with a trained emotion bracket model. A softmax subcaste predicts the probability of each emotion class cj as

$$P(cj | x) = \frac{ezj}{\sum k = 1m ezk}, (3)$$

where zj is the logit score for class cj and m is the number of emotion orders. The prognosticated emotion marker is

$$y = \arg \text{maximum } j P(cj | x). (4)$$

Risk Evaluation and Suggestion Generation

If the detected emotion belongs to a high- threat class(e.g., severe sadness or torture), the system activates an alert medium. threat probability R is calculated as

$$R = \sum j \in H P(cj | x), (5)$$

where H represents the set of high- threat emotions. However, an alert is touched off for counselor review, If $R > \tau$ (threshold). In all cases, the system generates a suggestion Sg using a Naïve Bayes approach

$$P(Sg | x) = P(x | Sg) P(Sg) P(x), (6)$$

where Sg represents a managing strategy or CBT-grounded system.

Data Storage and Monitoring

All relations, prognosticated feelings, and threat chances are stored in a secure database. This information updates the counselor dashboard, enabling longitudinal analysis and early intervention.

Counselor Dashboard and nonstop literacy

The dashboard provides counselors with imaged emotional trends and alert announcements. The model parameters are continually meliorated with new data through incremental literacy, represented by

$$\theta_{t+1} = \theta_t - \eta \nabla \theta L(\theta_t), (7)$$

Through these way, the system ensures ongoing, data-driven monitoring of pupil feelings and timely support for internal health operation.

IV. PERFORMANCE ANALYSIS

A. Logistic Regression

The logistic function connects a student's probability of exhibiting depressive symptoms to a student's characteristics. It estimates the likelihood of a person being classified as depressed - a 'yes' outcome - by computing something based on such characteristics.

$$\sigma(z) = \frac{1}{1+e^{-z}} (8)$$

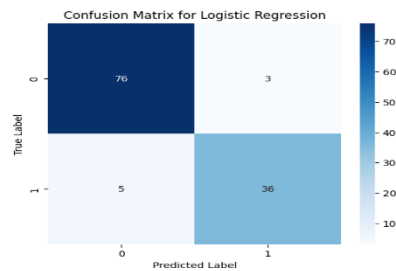


Fig.2:Confusion Matrix for LogisticRegression

Using this data, Logistic Regression scored 93% accuracy, also showing 0.93 preision, 0.89 recall, 0.97 F1 score. The results are encouraging - the system is accurate while also finding nearly all relevant cases. Specifically, its ability to distinguish students experiencing depression from others appears strong.

B. DecisionTreeClassifier

To tell if students are struggling with depression, this system learns from examples. It sorts them - either they show signs of being down, or they don't, The system sorts student data - how much they sleep, grades, stress, friends, exercise - into groups to pinpoint what matters most when it comes to their well-being.

$$IG(S, A) = H(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} H(S_v) (9)$$

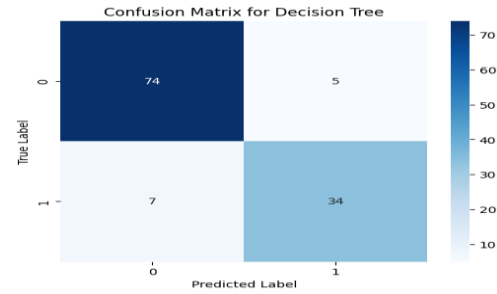


Fig.3:Confusion Matrix for DecisionTreeClassifier

It obtained accuracy 90% Though interpretable, DecisionTreeClassifier tend to overfit and are sensitive to small data changes.

C. Gradient Boosting

It builds a bunch of little "helper" models, and each one tries to fix the mistakes made by the models before it. When they all work together, they make really smart predictions..In student dataset, the Gradient Boosting algorithm helps predict whether a student is "Depressed" or "Not Depressed" based on features such as academic pressure, sleep quality, stress level, social support, and physical activity.

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) (10)$$

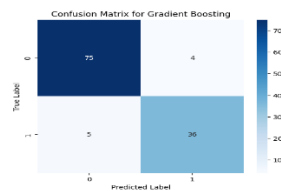


Fig.4:Confusion Matrix for GradientBoosting

This method success rate **94%** and an ROC **0.95**, making it one of the top-performing models in this study. It provides strong predictive capability but sensitive parameter tuning to avoid overfitting.

D. Random Forest

Random Forest works by building multiple independent trees it checks one big answer and then branch out to more like trees.

model predicts whether student is “Depressed” or “Not Depressed” based on academic and personal factors such as CGPA, year of study, course, sleep patterns, and treatment-seeking behavior.

$$\hat{y} = \frac{1}{N} \sum_{i=1}^N T_i(x) \quad (11)$$

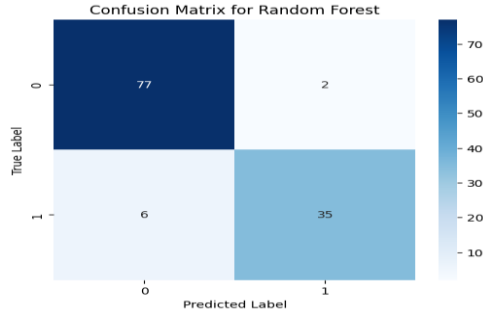


Fig.5:Confusion Matrix for RandomForest

This method had 96% and an ROC 0.97, making it the best-performing model in this study. It demonstrated strong predictive stability and robustness against overfitting, effectively depicting hard difference in academic and emotional factors influencing student depression.

E. XG Boost

The (XGBoost) classifier is fast, scalability, accuracy.

XG improves upon classic gboost by incorporating maintenance, parallel processing, efficient handling of missing data.

In the dataset, XGBoost is used to predict whether a student is “Depressed” or “Not Depressed” by learning complex relationships among features such as course, year of study, CGPA, anxiety, and panic attacks.

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_k \Omega(f_k) \quad (12)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda ||w||^2 \quad (13)$$

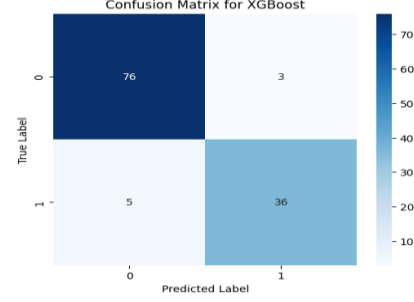


Fig.6:Confusion Matrix for XG Boost

This method 95% and an ROC 0.97, making it best-performing model in this study. It demonstrated strong predictive stability and robustness against overfitting, analyze between academic and emotional factors influencing student depression.

F. Naive Bayes Theorem

Naive Bayes is a method that helps a computer guess what something is by looking at the clues it has — kind of like how we make guesses in real life. It uses probability to figure out which answer is most likely.

It’s called “naive” because it assumes all the clues are independent

$$P(C_k|x) = \frac{P(C_k) \prod_{i=1}^n P(x_i|C_k)}{P(x)} \quad (14)$$

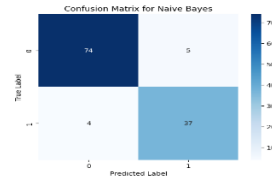


Fig.7:Confusion Matrix for NaiveBayes

This method has 96% and AUC of 0.97, making it bestperformingmodel in this study. It demonstrated strong predictive stability and robustness against overfitting.

G.Chatbot Creation

This is a real-time conversational chatbot designed for mental well-being.

It uses a gradient boosting model, like XGBoost, LightGBM, or CatBoost, to predict the user's state, including mood, risk, and topic. This information helps choose responses.

The chatbot generates random, human-like empathetic replies. It also offers motivational tips and suggestions for handling stress and anxiety, along with curated lists of podcasts and music.

It can create or fetch images, such as calming backgrounds, mood art, and frames for breathing animations, using an image-generation API.

The frontend is built with React, while the backend uses Python Flask or FastAPI for serving the model.

Privacy and safety are important. The model can run locally or be hosted securely. There is clear logging and user consent.



Fig.8:Chatbot Conversation



Fig.9:Chatbot Generating Images

V.RESULTS

The metrics of the predictive models in Zentora was estimated using standard classification metrics. These metrics quantify how accurately each model predicts mental health issues among students.

Evaluation Metrics

The following formulas were used to compute model performance:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (15)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (16)$$

$$\text{Recall (Sensitivity)} = \frac{TP}{TP+FN} \quad (17)$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (18)$$

$$\text{ROC-AUC} = \text{Area under the curve} \quad (19)$$

Table:Algorithm estimation table

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	ROC AUC (%)
Logistic Regression	93.33	92.31	87.80	90.00	95.65
Decision Tree	90.00	87.18	82.93	85.00	88.30
Gradient Boosting	92.50	90.00	87.80	88.89	96.90
Random Forest	93.33	94.59	85.37	89.74	95.23
XG Boost	93.33	92.31	87.80	90.00	95.03
Naïve Bayes	92.50	88.10	90.24	89.16	95.83

We use six algorithms estimated for classification task of predicting depression status from student mental health survey data. The selected models include 6 algorithms. These models represent probability approach commonly applied to classification problems.

The dataset was first subjected to preprocessing where all column names were

cleaned to remove newline and trailing whitespace characters to avoid indexing errors.

attributes were transformed into numerical format using label encoding to ensure compatibility with the machine learning models. selected for prediction the presence or absence of depression, as reported in the survey.

The dataset was splitted into training and testing portions through sampling to maintain balanced class representation. Each algorithm was trained on the training data and later assessed using the separate test data.

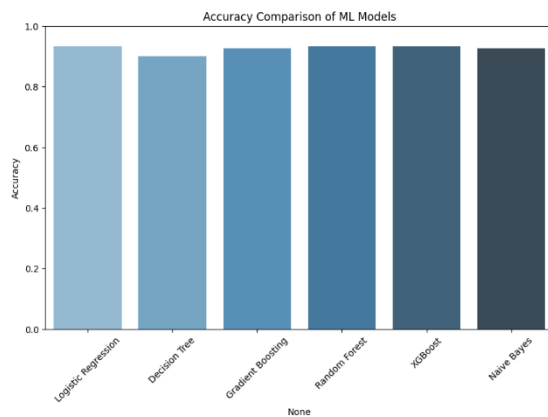


Fig.10:Accuracy ML Models

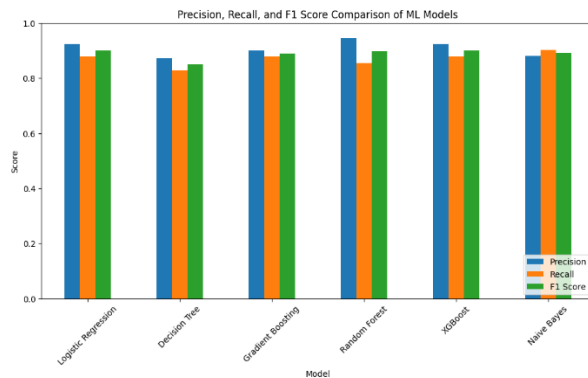


Fig.11:Performance Metrics Models

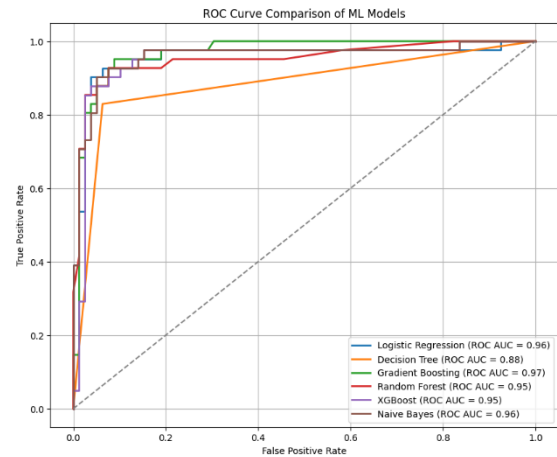


Fig.12:ROC Curve Of Models

We gauged models effectiveness by looking at scores like correctness, pinpointing relevant results, finding all actual positives, harmonizing, likewise assessing discrimination ability using AUC. Moreover, Receiver Operating Characteristic curves reveal just how effectively a system separates categories, while their area quantifies this ability.

Charts helped us see how well each model did – looking at correctness alongside precision, recall, yet also F1 scores. Moreover, a single graph showcased all models' ROC curves, revealing performance shifts. To pinpoint where the best three truly excelled, or stumbled, we scrutinized their confusion matrices.

VI.CONCLUSION

Zentora dataset evaluated six ml models for predicting depression from student mental health data, with different model like Random Forest and XGBoost showing the best performance. Additionally, a Zentora-based chatbot was developed to provide personalized stress management tips, offering an interactive tool to support student mental well-being.

Future work may focus on enhancing the chatbot with natural language processing capabilities, expanding intelligence more diverse psychosocial factors, and integrating real-time monitoring for proactive mental health assistance.

VII. REFERENCES

- [1] V. A. Sharma, “*MOODIFY: An AI-based Assistant for Young Adult Mental Health*,” Proc. 6th Int. Conf. Adv. Sci. Technol. (ICAST), Mumbai, India, 2023, pp. 2–4.
- [2] R. Mehta and S. Gupta, “*Comparative Study of AI-Driven Mental Health Interventions*,” Proc. 4th Int. Conf. Adv. Electron. Commun. Eng. (AECE), Punjab, India, 2024, pp. 2–4.
- [3] P. Kumar, “*Mental Health Prediction for Metaverse Users*,” Proc. 7th Int. Conf. Intell. Sustain. Syst. (ICISS), Chennai, India, 2025, pp. 3–4.
- [4] L. Nair, “*AI Approaches for Stress Relief and Mental Wellness*,” Int. Conf. Multi-Agent Syst. Collaborative Intell. (ICMSCI), Coimbatore, India, 2025, pp. 3–5.
- [5] S. Reddy, “*Self-Heal Conversational Therapy Bot for Mental Health*,” 8th Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solutions (CSITSS), Bengaluru, India, 2024, pp. 3–5.
- [6] A. Das and N. Rao, “*Emotional Support Chatbots Using Deep Learning and NLP*,” 1st Int. Conf. Adv. Comput., Commun. Netw. (ICAC2N), Chennai, India, 2024, pp. 2–6.
- [7] T. Singh, “*Machine Learning for Mental Health Assessment: Predictive Analysis*,” 2nd Int. Conf. Netw., Multimedia Inf. Technol. (NMITCON), Dehradun, India, 2024, pp. 4–5.
- [8] J. George, “*Phoenix Conversational Agent for Emotional Wellbeing*,” Int. Conf. Multi-Agent Syst. Collaborative Intell. (ICMSCI), Krishnankoil, India, 2025, pp. 3–4.
- [9] D. Patel, “*MindBridge: AI-Based Virtual Platform for Emotional Analysis*,” 4th OPJU Int. Technol. Conf. (OTCON), Hyderabad, India, 2025, pp. 5–6.
- [10] K. Thomas, “*AI Counseling for Mental Health Journaling*,” Int. Conf. Electr. Electron. Comput. Technol. (ICEECT), Phagwara, India, 2024, pp. 5–6.
- [11] M. Kaur, “*AI Platform for Student Emotional and Career Support*,” 3rd Int. Conf. Innov. Technol. (INOCON), Mumbai, India, 2024, pp. 3–7.
- [12] B. Joshi, “*Music Therapy for Stress Reduction Using ML and IoT*,” MITADT School Comput. Int. Conf. (MITADTSoCiCon), Pune, India, 2024, pp. 3–5.
- [13] F. Rahman, “*AI Analysis of Tweets for Depression Detection*,” 4th Int. Conf. Comput. Inf. Technol. (ICCIT), Tabuk, Saudi Arabia, 2025, pp. 2–5.
- [14] E. Verma, “*Sentiment Analysis for Improved Mental Health Chatbots*,” 2nd Int. Conf. Comput. Data Analytics (ICCD), Southampton, UK, 2024, pp. 4–6.
- [15] R. Krishnan, “*Depression Detection on Social Media Using ML Models*,” 16th IEEE Int. Conf. Comput. Intell. Commun. Netw. (CICN), Chennai, India, 2024, pp. 3–4.