JENTIK

OPEN ACCESS

Jurnal Pendidikan Teknologi Informasi dan Komunikasi
https://ejournal.1001tutorial.com/index.php/jentik

Predicting Student Depression Using the Naive Bayes Model on the Student Depression Dataset from Kaggle

Rebina Putri Sonjaya[1*], Andre Rangga Gintara[2], Lala Septem Riza[3], Muhammad Nursalman[4], Eki Nugraha[5], Didin Wahyudin[6]

[1,2,3,4,5,6]Universitas Pendidikan Indonesia
Bandung, Indonesia
rebinaputris@upi.edu*

Article Information:
Received June 15, 2025
Revised June 22, 2025
Accepted June 23, 2025

Keywords:
Dataset Kaggle; Depression Detection; Machine Learning; Naïve Bayes; Student Mental Health

Abstract

Background of Study: The increasing prevalence of depression among college students highlights the urgent need for effective early detection strategies to promote mental well-being within higher education environments.

Aims and Scope of Paper: This study aims to develop a predictive model for student depression using the Naive Bayes classification algorithm, with a focus on identifying key contributing factors from student-related data.

Methods: The research utilizes the Student Depression dataset from Kaggle, containing structured survey data on academic stress, sleep duration, financial stress, GPA, and family mental health history. Data preprocessing included feature selection, handling of missing values, and normalization. The dataset was split into training and testing sets at a 75:25 ratio. Model training was conducted using the R programming language with the application of Laplace smoothing.

Result: The Naive Bayes model achieved an accuracy of 77.66%, a specificity of 84.21%, and a sensitivity of 68.42%, indicating strong predictive performance, particularly in identifying depressive cases. Financial and academic stress were identified as the most influential factors.

Conclusion: Despite its simplicity, the Naive Bayes algorithm proves to be an effective tool for initial screening of students at risk of depression, offering valuable support for educational institutions in delivering timely mental health interventions.

A. Introduction

Today, global concern over the high prevalence of depression among college students is increasing. A report from the WHO shows that almost all psychiatrists experienced in dealing with college students agree that about 10 % of members of higher education institutions are likely to experience an emotional disturbance each year, which seriously hinders their academic or work activities. Previous research has also stated that depression has an impact on a large number of college students (Lorentzen et al., 2020) (Liu et al., 2023), and the prevalence is higher compared to the general population (Campbell et al., 2022) (Adlaf et al., 2014). Early detection of depression is essential for timely intervention, appropriate support, and reducing the serious impacts that can arise (Moon et al., 2021) (Windarwati et al., 2022) (Sayed et al., 2022). However, many college students tend to delay seeking psychological help until their depressive symptoms reach moderate to severe levels. This is often due to low mental health literacy and the perception that their problems are not serious enough to require professional help (Kustimah et al., 2023).

To address these challenges, the use of technology-based self-detection tools is becoming increasingly important so that universities can identify at-risk students and proactively intervene. Advances in machine

How to Cite : Sonjaya, R. P., Gintara, A. R., Riza, L. S., Nursalman, M., Nugraha, E., & Wahyudin, D. (2025). Predicting Student Depression Using the Naive Bayes Model on the Student Depression Dataset from Kaggle. *JENTIK : Jurnal Pendidikan Teknologi Informasi Dan Komunikasi*, *4*(1). https://doi.org/10.58723/jentik.v4i1.448
ISSN : 2963-1963
Published by : CV Media Inti Teknologi

learning (ML) have opened up great opportunities in developing data-driven prediction systems, which are able to process complex relationships between variables with high accuracy (Lee et al., 2018) (Hatton et al., 2019) (Shatte et al., 2019) (Gil et al., 2022).

A number of ML algorithms have been used in depression prediction, including Support Vector Machine, Decision Tree, Neural Network, Latent Dirichlet Allocation, and Clustering (Shatte et al., 2019). However, most research is still limited to the use of complex models that are often less transparent and difficult to interpret.

One of the main challenges in depression detection research is limited access to relevant and structured data. In this study, the Student Depression dataset from Kaggle was used, which contained survey data related to symptoms of depression, stress, anxiety, as well as demographic factors and student lifestyle. This structured, survey-based dataset allows for clearer feature interpretation and reduces the need for complex preprocessing required in social media–based approaches.

Although there have been many studies of machine learning–based depression predictions, most of them focus on the general population or use data from social media that require complex text extraction. This kind of structured survey-based approach to students is still relatively rare, even though it can be more easily implemented by educational institutions.

In this context, the Naive Bayes Classifier algorithm offers a simple yet effective alternative, assuming conditional independence between features making it suitable for limited datasets and quick predictions (Cruz et al., 2023). Its computational efficiency, ease of parameter estimation, and interpretability make it ideal for deployment by non-technical users such as campus counsellors (Stiglic et al., 2020). Although often considered a baseline, Naive Bayes shows competitive performance, especially on structured survey datasets: for example, (Cruz et al., 2023) reported 78.03 % accuracy in predicting depression among 519 university students, and (Haque et al., 2021) reported up to 85 % accuracy using survey-derived features.

Therefore, this study leverages Naive Bayes for its practicality and transparency, aiming to develop a depression prediction model in college students using actual survey datasets from Kaggle. With a focus on early detection, method efficiency, interpretability, and practical application in educational settings, the study is expected to contribute to efforts to prevent and treat depression in higher education environments.

B. Research Methods

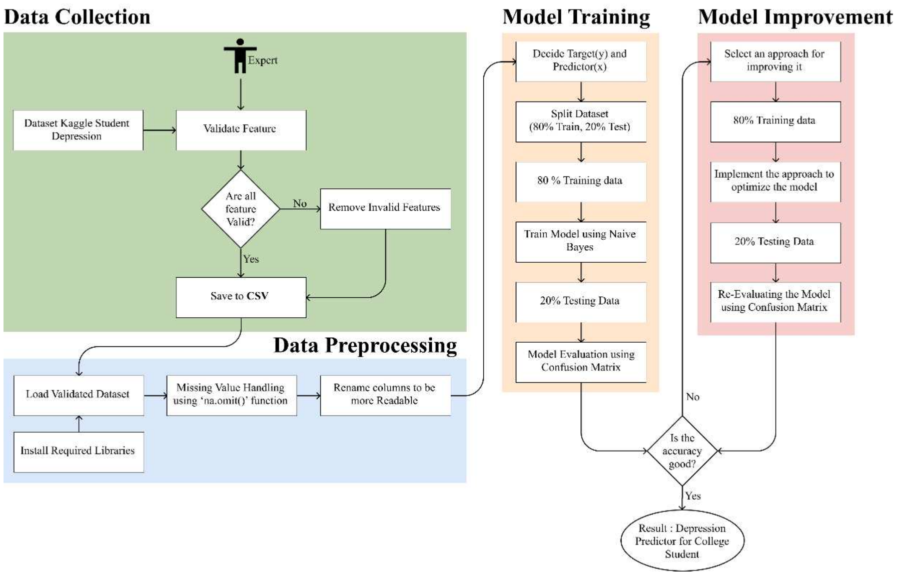

Figure 1. Computing Model

The computational model in this study is divided into 5 processes: data collection, data pre-processing, modeling, model improvement and final results. A diagram of the computational model can be seen in Figure 1.

1. Data Collection

Data was collected from the Kaggle platform with a dataset titled Student Depression, which contains various features related to students' mental states. Furthermore, feature validation is performed by an expert (in this context it can be a psychology student) to ensure only the relevant features are retained. The expert conducted validation based on psychological relevance, such as whether the feature reflected emotional, behavioral, or cognitive indicators commonly associated with depressive symptoms. If there is a feature that is considered invalid or irrelevant, it is removed before the data is saved in CSV format. This expert-driven feature screening helps maintain the clinical validity of the dataset prior to preprocessing and modeling (Shatte et al., 2019).

2. Data Preprocessing

This stage aims to prepare the dataset for effective model training by ensuring data consistency, completeness, and proper formatting. The steps involved are as follows:

a. Loading Dat

The dataset Student_Depression_Dataset.csv read into R using the read.csv() function. The stringsAsFactors = FALSE argument is used to ensure that text fields are not automatically converted into factors, giving more control at a later stage.

```
df <- read.csv("Student_Depression_Dataset.csv", stringsAsFactors =
FALSE)
```

b. Column Name Cleanup (Sub step 3a is coded)

The column names in the dataset are cleaned using the clean_names() function of the janitor package. This process changes the name of the column to a standard format (e.g., all lowercase, spaces replaced with an underscore _), such as academic_pressure, work_study_hours, etc. This makes it easy to write code and be consistent.

```
df <- clean_names(df)
```

c. Column Selection (Sub-step 3b in the code)

Based on the initial instructions, the relevant columns are selected:

- Categorical Predictor Features: academic_pressure, sleep_duration, degree, work_study_hours, financial_stress, family_history_of_mental_illness.
- Features of Numerical Predictor: cgpa.
- Identity columns (saved but not for models): gender, age.
- Variabel Target: depression.

Other columns that are not included in this list are ignored for the modeling process. The R code checks to make sure all the required fields are in the dataset.

d. Handling Missing Values (Sub-step 3c in the code)

Iterations are carried out on each column that has been selected. If a column is character-typed, the value of an empty string ("") or string "na" (no matter the case case) is converted to a standard NA value of R.

Tabel 1. Example of Missing Value Handling

Raw Entry (Before)	Processed Value (After)
""	NA
"na"	NA
"NA"	NA

After that, all rows containing at least one NA value in the selected columns are removed from the dataset using na.omit(). The code also checks if there is any data left after this process.

e. Data Transformation (3d sub-steps in code)

Conversion of Categorical Columns to Factors: Columns that are categorical (gender, academic_pressure, sleep_duration, degree, work_study_hours, financial_stress, family_history_of_mental_illness, and depression as targets) are converted to data factor types in R. Prior to conversion, text values are cleaned of spaces at the beginning/end (trimws()) and converted to uppercase (toupper()) to ensure consistency (e.g., "Yes", "yes", "YES" are treated the same). The code checks whether the depression target variable has at least two levels required for classification.

Tabel 2. Example of Categorical Value Transformation

Original	Transformed
" less than 5 hours "	LESS THAN 5 HOURS
"yes"	YES

CGPA Scale Conversion: The cgpa column that originally had a scale of 0-10 is converted to a scale of 0-4 to match the GPA standard. First, it is ensured that the cgpa column is numerical type. If there is a comma as a decimal separator, it is converted to a period. Rows with invalid cgpa values (can't be changed to numerical) will be deleted. The numerical cgpa value is then divided by 2.5. Clipping is performed, which is the cgpa value that after conversion exceeds 4 will be set to 4, and those less than 0 will be set to 0. This cgpa column is retained as a numerical feature for the Naïve Bayes model.

Tabel 3. Example of CGPA Scale Conversion (0–10 to 0–4)

Original CGPA	Converted (0–4 scale)
8.5	3.4
9.2	3.68
10.5	4.0 (clipped)

Preparation of Final Data for Modeling: A new dataframe (modeling_df) is created that contains only the final predictor features (model_features_final which includes categorical features and numerical cgpa) and target depression variables. All categorical features and target variables are ensured to be factor-type, and factor levels that have no observations are removed using droplevels().

3. Data Splitting

Datasets that have gone through pre-processing (modeling_df) are then divided into two subsets:

a. Training Set: Used to "teach" or train the Naïve Bayes model.
b. Test Set: Used to test how well a trained model can make predictions on new data that it has never seen.

The partitioning process is done using the sample.split() function from the caTools package. The split ratio is set to 75% for the training data and 25% for the test data (SplitRatio = 0.75). This 75:25 division is a commonly accepted standard in machine learning experiments and offers a good balance between training efficiency and testing reliability (Shin et al., 2024).

Stratification based on the target variable (depression) is applied to ensure that both subsets maintain the same class distribution, which is especially important when working with potentially imbalanced data. This helps prevent biased model evaluation and improves generalization to unseen data.

set.seed(123) is used before partitioning to ensure that the random process in the data selection is reproducible, so that if the code is re-run, the result of the data partition will remain the same. This practice ensures the experiment is reliable and can be repeated or audited consistently by other researchers (Berrar, 2018).

Checks were carried out to ensure that the training data and test data were not empty and both still had representations of both target classes (depression). This step is essential to confirm that no class is excluded from either subset, which could otherwise invalidate the model training or testing process.

4. Building the Naïve Bayes Model

The Naïve Bayes model is constructed using the training data (train_data) as follows:

This section not only outlines the technical construction of the model but also includes critical enhancements such as Laplace smoothing and robust error handling with tryCatch, which improve stability and generalizability in practical applications.

a. Model Formula: The formula used is depression ~ .. This means that we want to predict the depression variable using all the other variables present in the train_data as predictor features. The formula depression ~ . instructs the model to predict the depression target using all other variables.
b. Model Training: The naiveBayes() function of the e1071 package is used to train the model.
   - data = train_data: Specifies the data used for training.
   - laplace = 1: Apply Laplace smoothing with parameter 1. This is a technique to address the zero probability problem that can occur if there is a combination of features and classes that never appear in the training data, thus making the model more stable.
   - The numerically typed cgpa feature will be automatically handled by the naiveBayes function assuming that its values are Gaussian-distributed (normal) for each target class. The model will calculate the average and standard deviation of the cgpa for each class from the training data.
c. TryCatch blocks are used to handle potential errors during the model training process and provide a warning message if there is a factor feature with only one unique level, which would make it unsuitable for training.
d. The results of the trained model are stored in naive_bayes_model object and then printed to the console to display the model summary.

```
formula_nb <- depression ~ .
naive_bayes_model <- NULL
tryCatch({
  naive_bayes_model <- naiveBayes(formula = formula_nb, data =
train_data, laplace = 1)
}, error = function(e) {
  print(paste("Error Naive Bayes:", e$message))
  for(col_name in setdiff(names(train_data), "depression")){
    if(is.factor(train_data[[col_name]]) &&
nlevels(train_data[[col_name]]) < 2 && nrow(train_data) > 0){
      print(paste("Peringatan: Fitur faktor", col_name, "hanya
memiliki 1 level unik."))
    }
  }
})
if (is.null(naive_bayes_model)) stop("Gagal melatih model Naive
Bayes.")
print("Output Model Naive Bayes:")
print(naive_bayes_model)
```

Laplace smoothing is particularly important in classification problems with small datasets or imbalanced class distributions. In practice, it avoids assigning a zero probability to any category simply because it does not appear in the training data. For example, if a certain category of a predictor variable is present in the test data but was not seen during training for a particular class, the model without Laplace smoothing would assign zero probability to the entire prediction. By adding a small value (in this case, 1) to all feature counts, Laplace smoothing ensures that every possible feature-class combination receives a non-zero probability, making the model more reliable and generalizable.

In addition, the use of the tryCatch block ensures that errors during model training (such as the presence of factor variables with only one unique level) do not cause the entire process to fail. Instead, the code explicitly checks each factor predictor and prints a warning if it is found to be unsuitable for training. This proactive handling of potential data quality issues enhances model stability and provides clear feedback to the developer, facilitating debugging and dataset refinement.

This approach ensures both model robustness and interpretability, even in the presence of imperfect or imbalanced data.

5. Making Predictions

Once the model has been successfully trained, it is used to make predictions on the test data (test_data): The predict() function is used with naive_bayes_model and newdata = test_data to generate a depression class prediction for each observation in the test data. The prediction results are stored in the predictions variable.

```
predictions <- predict(naive_bayes_model, newdata = test_data)
```

6. Model Evaluation

The performance of the model was evaluated by comparing the predictions with the actual depression values of the test data:

a. Factor Level Adjustment: To ensure consistency during evaluation, the factor level of the actual target variable (test_data$depression) and the prediction variable is equalized using all_possible_levels taken from train_data$depression. This is important for confusionMatrix() to function correctly. This step is especially crucial in R, where mismatched factor levels between prediction and reference values can cause errors or misalignment in evaluation metrics.
b. Handling of NA on Factors (if any): Checks are carried out to see if any NA values appear after adjustment of the factor level. If there are, rows with those NA values will be removed from the evaluation, although this should not occur if the data is already clean. This check helps prevent misleading metrics due to invalid rows being included in the evaluation.
c. Confusion Matrix: The confusionMatrix() function from the caret package is used to create a contingency table (confusion matrix) that compares the model's predictions with the actual values. This function also accounts for unbalanced classes by providing balanced accuracy, a robust metric in such contexts.
d. Performance Metrics: The output of confusionMatrix includes various important metrics namely Accuracy, 95% CI (Confidence Interval) for Accuracy, No Information Rate (NIR), P-Value [Acc > NIR], Kappa, Sensitivity (Recall), Specificity, Post Pred Value (Precision), Neg Pred Value, Balanced Accuracy: The average of Sensitivity and Specificity.
e. The accuracy of the model is specifically extracted and printed to the console.
These comprehensive steps for evaluating classification performance help ensure the reliability and validity of the model in detecting depression cases accuratel.

C. Results and Discussion

1. Results

The Naïve Bayes model has been trained using 75% of the data from Student_Depression_Dataset.csv datasets that have gone through the pre-processing stage. Features used in modeling include academic_pressure, sleep_duration, degree, work_study_hours, financial_stress, family_history_of_mental_illness (as a categorical feature), and cgpa (as a numerical feature after scaling conversion to 0-4). Laplace smoothing with a value of 1 is applied during model training to improve stability.

1.1. Summary of the Naïve Bayes Model

a. A-priori probability of Target Class (depression):

```
A-priori probabilities:
Y
        0         1
0.4144721 0.5855279
```

Figure 2. A priori probability.

- Class '0' (Non-Depressed): 0.4144721 (about 41.45%)
- Class '1' (Depression): 0.5855279 (about 58.55%) This indicates that in the training data, there are slightly more cases classified as 'Depression' (class '1').

b. Conditional Probability of Features: The model calculates the probability of each category/value of a feature appearing, given a specific depression state (class '0' or '1'). Some of the important findings of conditional probability are:

- Academic_Pressure

```
   academic_pressure
Y              0            1            2            3            4            5
  0 0.0004609357 0.3372896981 0.2233233464 0.2554736114 0.1054390413 0.0780133671
  1 0.0002447581 0.0568654646 0.0952108999 0.2751081015 0.2401892796 0.3323814963
```

Figure 2. Probability academic_pressure.

Higher levels of academic stress (e.g., level '5') are more commonly associated with class '1' (Depression) with a probability of 0.332, compared to class '0' (Non-Depression) which has a probability of 0.078 for the same level. In contrast, lower levels of academic stress (e.g., level '1') were more often associated with class '0' (P=0.337) than class '1' (P=0.057).

- Sleep_Duration

```
   sleep_duration
Y      5-6 HOURS    7-8 HOURS LESS THAN 5 HOURS MORE THAN 8 HOURS       OTHERS
  0 0.2340670739 0.2528523683      0.2552725596      0.2571165149 0.0006914832
  1 0.2147519582 0.2680319843      0.3264523499      0.1901109661 0.0006527415
```

-

Figure 3. Probability sleep_duration

The sleep duration of "LESS THAN 5 HOURS" was more often found in class '1' (Depression, P=0.326) than in class '0' (Non-Depression, P=0.255).

- Financial_Stress

```
   financial_stress
Y            1          2          3          4          5
  0 0.30148669 0.24674427 0.19107987 0.15131958 0.10936960
  1 0.09921671 0.13454634 0.18937663 0.24510444 0.33175587
```

Figure 4. Probability financian_stress

Financial stress at the highest level (e.g., level '5') was significantly more common in class '1' (Depression, P=0.332) than in class '0' (Non-Depressive, P=0.109). In contrast, the lowest level financial stress (e.g., level '1') was more common in class '0' (P=0.301) than class '1' (P=0.099).

- Family_History_of_Mental_Illness

```
   family_history_of_mental_illness
Y          NO       YES
  0 0.5463454 0.4536546
  1 0.4957969 0.5042031
```

Figure 5. Probability family_history_of_mental_illness

The presence of a family history of mental illness (YES) is slightly more often observed in class '1' (Depression, P=0.504) than in class '0' (Non-Depression, P=0.454).

- **CGPA**

```
   cgpa
Y        [,1]      [,2]
  0 3.038219 0.5977549
  1 3.071477 0.5833965
```

Figure 6. CGPA Probability

The average GPA (scale 0-4) for class '0' (Non-Depression) was 3.038 (standard deviation 0.598), while for class '1' (Depression) it was 3.071 (standard deviation 0.583). The difference in average GPA between these two groups is relatively small.

1.2. Evaluation Results on Test Data (from print(confusion_matrix_results))

a. Confusion Matrix:

```
          Reference
Prediction    0    1
         0 1978  645
         1  913 3439
```

Figure 7. Confusion Matrix Results

(Assuming '0' = Not Depression, '1' = Depression)

- True Negative (TN): 1978
- False Positive (FP): 645
- False Negative (FN): 913
- True Positive (TP): 3439

b. Overall Model Accuracy: 77.66%

```
               Accuracy : 0.7766
                 95% CI : (0.7667, 0.7864)
    No Information Rate : 0.5855
    P-Value [Acc > NIR] : < 2.2e-16
```

Figure 8. Model Accuracy Results

- 95% Confidence Interval for Accuracy: (0.7667, 0.7864)
- No Information Rate (NIR): 0.5855 (Model accuracy well above NIR)
- P-Value [Acc > NIR]: < 2.2e-16 (Statistically significant model accuracy)

c. Statistics Kappa

```
Kappa : 0.5335
```

Figure 9. Kappa Statistics Results

- 0.5335 (Indicates moderate agreement)

d. Performance Metrics Per Class (with '0' as the Positive Class for this metric)

```
         Sensitivity : 0.6842
         Specificity : 0.8421
      Pos Pred Value : 0.7541
      Neg Pred Value : 0.7902
          Prevalence : 0.4145
      Detection Rate : 0.2836
Detection Prevalence : 0.3761
   Balanced Accuracy : 0.7631
```

Figure 10. Performance metrics per class

- Sensitivity (Recall for class '0' - No Depression): 0.6842 (68.42%)
- Specificity (Recall for class '1' - Depression): 0.8421 (84.21%)
- Post Pred Value (Precision for class '0' - Not Depressed): 0.7541 (75.41%)
- Neg Pred Value (Precision for class '1' - Depression): 0.7902 (79.02%)
- Balanced Accuracy: 0.7631 (76.31%)

2. Discussion

The Naïve Bayes model developed to predict the depression status of students based on the Student_Depression_Dataset.csv showed a fairly good performance with an overall accuracy of 77.66% on the test data. This accuracy is significantly higher than the random guessing rate (No Information Rate 0.5855), indicating that the model is successfully capturing relevant patterns in the data.

From the conditional probability analysis generated by the model, several factors show a stronger association with depressive status:

- Financial Stress and Academic Stress: These two variables demonstrate the most notable differences between depressed and non-depressed students. Higher levels of both stressors are more frequently found in students predicted as depressed. This supports findings from prior research, such as (Jin, 2025), who also used this Kaggle dataset (with a Decision Tree model), and from (Danahy et al., 2024), who highlighted financial burdens like debt and limited access to social activities as stress triggers.
- Sleep Duration: Sleep deprivation (less than 5 hours) showed a strong correlation with depression, aligning with findings by (Alqurashi et al., 2022) and (Dinis & Bragança, 2018) who noted that both insufficient and excessive sleep are linked to poor mental health.
- Family History of Mental Illness: While not as strong as stress variables, this factor still contributes to increased likelihood of depression, in line with the known genetic and environmental risks of depression.
- GPA: After normalization to a 0–4 scale, GPA did not show a significant predictive difference between depressed and non-depressed students, indicating it may be less useful than psychosocial or lifestyle indicators.

The confusion matrix revealed that the model's Specificity (84.21%) is higher than its Sensitivity (68.42%), suggesting it is more effective in identifying true positive depression cases than avoiding false positives. This is suitable for early screening purposes, where missing a case of depression is more harmful than a false alarm. Similar to (Haque et al., 2021), model performance can vary depending on the dataset's structure and the model's clinical objectives. A Kappa value of 0.5335 indicates a moderate degree of agreement between the model's prediction and the actual condition, indicating that the model provides better results than a chance deal.

2.1 Implications

Naïve Bayes model provides a useful basis for understanding the factors associated with depression in college students in the context of the datasets used. While its accuracy is not perfect, it can be an early tool in identifying students who may be at risk and need further attention. The limitations of the Naïve Bayes model, such as the assumption of independence between features, may limit its ability to capture more complex interactions between variables. Future research may explore other machine learning algorithms or ensemble techniques for the potential for improved accuracy and deeper understanding.

2.2 Research Contribution

This study contributes to the growing body of research on mental health prediction among college students by validating the utility of a simple, interpretable algorithm Naïve Bayes on a structured survey dataset. Unlike many machine learning studies that rely on unstructured social media text or black-box models, this research emphasizes practical implementation using interpretable features. By showing that Naïve Bayes can achieve reasonable accuracy, particularly in detecting true depression cases, the study supports its potential as a lightweight and low-resource screening tool in academic institutions.

2.3 Limitations

Although the Naïve Bayes algorithm performed adequately in this study, it has inherent limitations. Its core assumption of feature independence may oversimplify the complex interrelationships that often exist between psychological, social, and behavioral variables. Additionally, the dataset used was cross-sectional and self-reported, which could introduce bias or limit the generalizability of findings over time and across different cultural contexts. The lack of temporal tracking also prevents the model from distinguishing between short-term emotional responses and more persistent depressive symptoms. Another limitation lies in the exclusive use of structured categorical and numeric survey data, which may not capture the nuances of students' mental health experiences, unlike multimodal data such as sensor data, diaries, or open responses

2.4 Suggestions

Future research is encouraged to address these limitations by incorporating more complex or hybrid machine learning models that can learn interactions and hierarchies among variables—such as Random Forests, Gradient Boosting Machines, or Deep Learning approaches (Agarwal et al., 2025) (Rony et al., 2024). Incorporating Explainable AI (XAI) may also increase transparency and trust in automated predictions, especially in educational and clinical settings (Atlam et al., 2025).

Researchers should consider utilizing longitudinal data to monitor depressive symptoms over time, which would allow the development of predictive models that are sensitive to early warning signs (Chikersal et al., 2021). Lastly, ethical considerations such as student privacy, informed consent, and cultural adaptation of AI models must be taken into account to ensure the model's responsible deployment in different educational environments, including Indonesia (Saeidnia et al., 2024) (Alavi et al., 2025).

D. Conclusion

This study aims to develop and evaluate a depression prediction model in college students using the Naïve Bayes algorithm by utilizing Student_Depression_Dataset.csv dataset. After going through a series of stages including pre-processing of data, model training, and evaluation, several important conclusions can be drawn:

1. Naïve Bayes Model Performance: The developed model showed quite good performance, with an accuracy of 77.66%, a specificity of 84.21%, and a sensitivity of 68.42%. A Kappa value of 0.5335 indicates a moderate degree of agreement between the prediction and the actual conditions, indicating that the model performs better than random predictions.
2. Factors Associated with Depression: Based on the conditional probability analysis of the Naïve Bayes model, several factors show a significant association with the depressive status of college students are Financial stress and academic stress being the dominant factors that appear most often, Short sleep duration (less than 5 hours) is also an important indicator that leads to a depressive condition. Family history of mental illness contributed, although not as strong as the previous two factors. GPA (CGPA) shows a non-significant difference between the depressed and non-depressed groups, so its contribution as a differentiator is relatively low.
3. Practical Implications: These findings confirm that although the Naïve Bayes algorithm is relatively simple, it can be used as an initial tool to detect the risk of depression in college students. With an understanding of significant factors such as financial stress, academic stress, and sleep duration, educational institutions can develop more targeted and preventive mental intervention and support strategies.

Overall, this study confirms the potential use of machine learning, particularly Naïve Bayes algorithms, in an effort to understand and predict depression among college students, which can be the basis for the development of more effective prevention and intervention strategies in higher education settings.

E. Acknowledgment

The authors would like to express their sincere gratitude to the Department of Computer Science Education, Universitas Pendidikan Indonesia, for providing the academic support and research facilities used in this study. Special thanks are extended to the supervising lecturers—LSR, MN, EN, and DW—for their invaluable guidance, encouragement, and constructive feedback throughout the research process. The authors also acknowledge the contributions of colleagues and classmates who provided insight and support during the preparation of this manuscript.

F. Author Contribution Statement

RPS and ARG contributed to data collection, preprocessing, and initial model implementation. LSR, MN, EN, and DW provided supervision, guidance on research design, and validation of results. All authors contributed to the discussion of results, manuscript revision, and approved the final version of the manuscript.

References

Adlaf, E. M., Gliksman, L., Demers, A., & Newton-Taylor, B. (2014). The prevalence of elevated psychological distress among canadian undergraduates: Findings from the 1998 Canadian campus survey. *Journal of the American College Health Association*, *50*(2), 67–72. https://doi.org/10.1080/07448480109596009

Agarwal, P., Mundaragi, R. S., Kohad, R. S., Allada, R., Bharadwaj, S. R., & Shobha, T. (2025). Predicting Student Depression Using Machine Learning. *International Journal of Innovative Science and Research Technology*, *10*(1), 940–945. https://doi.org/10.5281/zenodo.14737958

Alavi, M., Le Lagadec, D., & Cleary, M. (2025). Challenges of Cross-Cultural Validation of Clinical Assessment Measures: A Practical Introduction. *Journal of Advanced Nursing*, 1–9. https://doi.org/10.1111/jan.16906

Alqurashi, Y. D., Al Qattan, A. H., Al Abbas, H. E., Alghamdi, M. A., Alhamad, A. A., Al-Dalooj, H. A., Yar, T., Al Khathlan, N. A., Alqarni, A. S., & Salem, A. M. (2022). Association of sleep duration and quality with depression among university students and faculty. *Acta Biomedica*, *93*(5), 1–8. https://doi.org/10.23750/abm.v93i5.13002

Atlam, E. S., Rokaya, M., Masud, M., Meshref, H., Alotaibi, R., Almars, A. M., Assiri, M., & Gad, I. (2025). Explainable artificial intelligence systems for predicting mental health problems in autistics. *Alexandria Engineering Journal*, *117*(December 2024), 376–390. https://doi.org/10.1016/j.aej.2024.12.120

Berrar, D. (2018). Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, *1–3*, 542–545. https://doi.org/10.1016/B978-0-12-809633-8.20349-X

Campbell, F., Blank, L., Cantrell, A., Baxter, S., Blackmore, C., Dixon, J., & Goyder, E. (2022). Factors that influence mental health of university and college students in the UK: a systematic review. *BMC Public Health*, *22*(1), 1–22. https://doi.org/10.1186/s12889-022-13943-x

Chikersal, P., Doryab, A., Tumminia, M., Villalba, D. K., Dutcher, J. M., Liu, X., Cohen, S., Creswell, K. G., Mankoff, J., David Creswell, J., Goel, M., & Dey, A. K. (2021). Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing: A machine learning approach with robust feature selection. *ACM Transactions on Computer-Human Interaction*, *28*(1), 1–41. https://doi.org/10.1145/3422821

Cruz, T. F., Flores, E. E. C., & Quispe, S. J. C. (2023). Prediction of Depression Level in University Students through a Naive Bayes based Machine Learning Model. *ArXiv Preprint*, *1*. https://doi.org/10.48550/arXiv.2307.14371

Danahy, R., Loibl, C., Montalto, C. P., & Lillard, D. (2024). Financial stress among college students: New data about student loan debt, lack of emergency savings, social and personal resources. *Journal of Consumer Affairs*, *58*(2), 692–709. https://doi.org/10.1111/joca.12581

Dinis, J., & Bragança, M. (2018). Quality of sleep and depression in college students: A systematic review. *Sleep Science*, *11*(4), 290–301. https://doi.org/10.5935/1984-0063.20180045

Gil, M., Kim, S. S., & Min, E. J. (2022). Machine learning models for predicting risk of depression in Korean college students: Identifying family and individual factors. *Frontiers in Public Health*, *10*. https://doi.org/10.3389/fpubh.2022.1023010

Haque, U. M., Kabir, E., & Khanam, R. (2021). Detection of child depression using machine learning methods. *PLoS ONE*, *16*(12 December 2021), 1–13. https://doi.org/10.1371/journal.pone.0261131

Hatton, C. M., Paton, L. W., McMillan, D., Cussens, J., Gilbody, S., & Tiffin, P. A. (2019). Predicting persistent depressive symptoms in older adults: A machine learning approach to personalised mental healthcare. *Journal of Affective Disorders*, *246*(September 2018), 857–860. https://doi.org/10.1016/j.jad.2018.12.095

Jin, Y. (2025). Decision Tree-Based Modeling in Mental Health Early Warning System for Higher Education Students. *Journal of Combinatorial Mathematics and Combinatorial Computing*, *127b*(July 2024), 1013–1034. https://doi.org/10.61091/jcmcc127b-057

Kustimah, K., Hanifah, H., Devy Kumalasari, A., & Meindy, N. (2023). When do College Students Seek Mental Health Services? *The Open Psychology Journal*, *16*, 1–6. https://doi.org/10.2174/18743501-v16-e230420-2022-112

Lee, Y., Ragguett, R. M., Mansur, R. B., Boutilier, J. J., Rosenblat, J. D., Trevizol, A., Brietzke, E., Lin, K., Pan, Z., Subramaniapillai, M., Chan, T. C. Y., Fus, D., Park, C., Musial, N., Zuckerman, H., Chen, V. C. H., Ho, R., Rong, C., & McIntyre, R. S. (2018). Applications of machine learning algorithms to predict therapeutic outcomes in depression: A meta-analysis and systematic review. *Journal of Affective Disorders*, *241*(March), 519–532. https://doi.org/10.1016/j.jad.2018.08.073

Liu, Y., Yu, H., Shi, Y., & Ma, C. (2023). The effect of perceived stress on depression in college students: The role of emotion regulation and positive psychological capital. *Frontiers in Psychology*, *14*(March), 1–10. https://doi.org/10.3389/fpsyg.2023.1110798

Lorentzen, V., Handegård, B. H., Moen, C. M., Solem, K., Lillevoll, K., & Skre, I. (2020). CORE-OM as a routine outcome measure for adolescents with emotional disorders: Factor structure and psychometric properties. *BMC Psychology*, *8*(1), 1–14. https://doi.org/10.1186/s40359-020-00459-5

Moon, N. N., Mariam, A., Sharmin, S., Islam, M. M., Nur, F. N., & Debnath, N. (2021). Machine learning approach to predict the depression in job sectors in Bangladesh. *Current Research in Behavioral Sciences*, *2*(May), 100058. https://doi.org/10.1016/j.crbeha.2021.100058

Rony, J. H., Syeed, M. M. M., Khan, R. H., Fatema, K., Hossain, M. S., & Uddin, M. F. (2024). Predicting Depression Among University Students: A Comparative Assessment of ML & DL models using XAI. *Conference Proceeding - 23rd International Symposium on Communications and Information Technologies, ISCIT 2024*, *December*, 175–180. https://doi.org/10.1109/ISCIT63075.2024.10793544

Saeidnia, H. R., Hashemi Fotami, S. G., Lund, B., & Ghiasi, N. (2024). Ethical Considerations in Artificial Intelligence Interventions for Mental Health and Well-Being: Ensuring Responsible Implementation and Impact. *Social Sciences*, *13*(7). https://doi.org/10.3390/socsci13070381

Sayed, T. A., Mahmoud, O. A. A., & Hadad, S. (2022). Early versus late onset depression: sociodemographic and clinical characteristics. *Middle East Current Psychiatry*, *29*(1). https://doi.org/10.1186/s43045-022-00227-8

Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine*, *49*(9), 1426–1448. https://doi.org/10.1017/S0033291719000151

Shin, J., Moon, H., Chun, C.-J., Sim, T., Kim, E., & Lee, S. (2024). Enhanced Data Processing and Machine Learning Techniques for Energy Consumption Forecasting. *Electronics*, *13*(19), 1–27. https://doi.org/10.3390/electronics13193885

Stiglic, G., Kocbek, P., Fijacko, N., Zitnik, M., Verbert, K., & Cilar, L. (2020). Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(5), 1–12. https://doi.org/10.1002/widm.1379

Windarwati, H. D., Lestari, R., Wicaksono, S. A., Kusumawati, M. W., Ati, N. A. L., Ilmy, S. K., Sulaksono, A. D., & Susanti, D. (2022). Relationship between stress, anxiety, and depression with suicidal ideation in adolescents. *Jurnal Ners*, *17*(1), 36–41. https://doi.org/10.20473/jn.v17i1.31216