



## **Comparative Analysis of KNN and Naive Bayes for Adolescent Mental Health Detection**

**\*Mahesa Darma Satria<sup>1</sup>, Yogi Saputra<sup>2</sup>**

<sup>1,2</sup> Magister of Informatics, Institut Informatika dan Bisnis Darmajaya, Indonesia

### **ARTICLE INFO**

#### **Article history:**

Received May 31, 2026

Revised Jun 27, 2026

Accepted Jun 30, 2026

#### **Keywords:**

*machine learning;  
mental health detection;  
k-nearest neighbor;  
naive bayes;  
imbalanced dataset*

### **ABSTRACT**

This study is motivated by the rapid rise of social media use among adolescents, which may adversely affect mental health by increasing the risk of stress, anxiety, and depression; thus, early detection is essential to prevent more severe outcomes. It aims to compare the performance of K-Nearest Neighbor (KNN) and Naive Bayes in detecting adolescent mental health risks and to evaluate the impact of data balancing using the Synthetic Minority Over-sampling Technique (SMOTE). A quantitative experimental design was applied, including data preprocessing, model implementation, and evaluation using 10-fold cross-validation with accuracy, precision, recall, F1-score, and AUC as performance metrics. The results show that Naive Bayes provides more stable performance with higher accuracy and precision, while KNN combined with SMOTE significantly improves recall, particularly for minority classes, indicating a trade-off between precision and recall in model selection. This study contributes a comprehensive analysis of the role of data balancing in classification performance within mental health contexts. Future work should explore ensemble and deep learning approaches and utilize larger, more diverse datasets to enhance generalizability.

*This is an open access article under the [CC BY-SA](#) license.*



### **Corresponding Author:**

Mahesa Darma Satria,

Magister of Informatics, Institut Informatika dan Bisnis Darmajaya

Jl. Zainal Abidin Pagar Alam No. 93, Gedong Meneng, Rajabasa, Bandar Lampung City, Lampung 35141, Indonesia

Email: [mahesa.student@darmajaya.ac.id](mailto:mahesa.student@darmajaya.ac.id)

## **INTRODUCTION**

The rapid advancement of digital technology and the widespread adoption of social media among adolescents have introduced significant challenges in the domain of mental health, particularly in the early detection of disorders such as depression and anxiety [1]. Empirical evidence suggests that intensive exposure to social media is positively associated with an increased risk of psychological distress, driven by factors including social pressure, self-comparison, and unhealthy patterns of digital interaction [2]. Concurrently, the ongoing digital transformation has generated vast amounts of behavioral data, offering new opportunities to assess individuals' mental states in a more objective and data-driven manner [3]. Within this context, machine learning approaches have emerged as highly relevant, given their capability to process large-scale datasets and uncover complex patterns that are difficult to identify through conventional analytical methods [4]. Accordingly, the development of data-driven models for the early detection of mental health conditions based on social media usage has become an essential step toward

enabling more effective preventive interventions.

In recent years, a growing body of research has explored the application of machine learning algorithms for mental health classification, including the use of K-Nearest Neighbor (KNN) and Naive Bayes methods [5]. KNN is recognized as a distance-based approach that effectively captures similarity between data instances, whereas Naive Bayes represents a probabilistic model known for its computational efficiency and robustness in handling high-dimensional data [6]. A study by S. Uddin et al. highlights that the performance of classification algorithms is strongly influenced by data characteristics, such as feature distribution and dataset complexity [7]. Furthermore, existing literature consistently indicates that no single algorithm universally outperforms others across all scenarios, underscoring the importance of comparative analysis in identifying the most suitable model for a given context [7]. Therefore, a systematic comparison between KNN and Naive Bayes for adolescent mental health detection represents a relevant and necessary line of inquiry.

Despite these advancements, a notable research gap persists, particularly in addressing the issue of imbalanced data, which is commonly encountered in mental health datasets. Prior studies have predominantly emphasized accuracy as the primary evaluation metric, often overlooking more informative measures such as recall and F1-score, which are crucial for assessing performance on minority classes [8][9]. This limitation is critical, as the failure to correctly identify individuals experiencing mental health disorders may lead to severe psychological consequences. In addition, data balancing techniques, such as Synthetic Minority Over-sampling Technique (SMOTE), have not been extensively examined in comparative studies involving KNN and Naive Bayes. Consequently, there is a need for research that not only compares algorithmic performance but also systematically evaluates the impact of data balancing techniques on classification outcomes.

The novelty of this study lies in its integrative comparative approach, combining the performance evaluation of KNN and Naive Bayes algorithms with the application of SMOTE to address data imbalance issues [10]. Unlike prior work that relies primarily on accuracy, this study incorporates precision, recall, and F1-score as key evaluation metrics, particularly emphasizing performance on minority classes [11]. Moreover, it provides an in-depth analysis of the trade-off between precision and recall, a critical consideration in the design of early mental health detection systems. By leveraging behavioral data derived from social media usage, this research also contributes to a deeper understanding of how data characteristics influence machine learning performance in real-world scenarios. This approach is expected to offer new insights into the development of more robust and data-adaptive mental health detection models.

Based on the identified background and research gaps, this study aims to conduct a comparative analysis of KNN and Naive Bayes algorithms for detecting adolescent mental health risks using social media usage data. Specifically, the study seeks to address the following research questions: (1) how do KNN and Naive Bayes perform in detecting adolescent mental health conditions based on standard classification evaluation metrics, and (2) how does the application of SMOTE influence the performance improvement of both algorithms when handling imbalanced datasets. The findings of this study are expected to contribute to the advancement of more accurate and adaptive early detection models for mental health.

## **RESEARCH METHODS**

This research applies a quantitative experimental design to examine and compare the ability of machine learning models to identify adolescent mental health risk based on social media usage behavior. This design was chosen because machine learning enables the analysis of complex behavioral data and has been widely used in digital mental health studies to support predictive assessment [12].

The methodological procedure in this study is organized into several sequential stages. These stages include dataset preparation, preprocessing, class imbalance handling, classification model construction, and model performance evaluation.

### Datasets

The dataset used in this research contains behavioral attributes related to adolescents' social media activities and mental health indicators. The variables include social media usage duration, sleep duration, stress level, anxiety level, and physical activity. The output variable is *depression\_label*, which is treated as a binary classification target, where 0 represents the non-risk class and 1 represents the depression-risk class.

The distribution of the target class indicates a strong imbalance between the two categories, with the number of depression-risk instances being much lower than the number of non-risk instances. Such a distribution is a common issue in mental health classification and may cause the model to favor the majority class if no balancing strategy is applied [13].

### Data Preprocessing

The preprocessing stage is carried out to ensure the quality, consistency, and readiness of the dataset prior to model training. The target variable is defined by assigning *depression\_label* as the classification label, while categorical features are transformed into numerical representations using dummy encoding. Subsequently, feature scaling is performed through Z-score normalization to standardize the range of all variables:

$$Z = \frac{x - \mu}{\sigma} \quad (1)$$

This normalization process is particularly crucial for distance-based algorithms such as K-Nearest Neighbor (KNN), as differences in feature scales can significantly affect distance computations and, consequently, the overall model performance [14].

### Handling Imbalanced Data

To address the issue of class imbalance, this study employs the Synthetic Minority Over-sampling Technique (SMOTE). This technique generates synthetic samples for the minority class by interpolating between existing instances and their nearest neighbors. Mathematically, the process can be expressed as follows:

$$x_{new} = x_i + \delta(x_{nn} - x_i) \quad (2)$$

This approach has been shown to be effective in enhancing the model's ability to learn patterns from the minority class, particularly in mental health datasets that typically exhibit skewed distributions [15]. To prevent data leakage, SMOTE is applied exclusively to the training data within the cross-validation framework.

### Classification Models

This study employs two primary classification algorithms, namely K-Nearest Neighbor (KNN) and Naive Bayes (NB), which are widely used in comparative machine learning studies.

#### K-Nearest Neighbor (KNN)

K-Nearest Neighbor is an instance-based learning algorithm that classifies data points based on their proximity to other instances in the feature space. The distance metric used in this study is the Euclidean distance:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

The parameter  $k$  is set to 3, and a weighted voting scheme is applied to improve prediction stability. KNN is highly sensitive to data distribution and feature scaling; therefore, preprocessing and data balancing play a crucial role in enhancing its performance.

### Naive Bayes (NB)

Naive Bayes is a probabilistic classification algorithm based on Bayes' Theorem, which assumes independence among features:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \quad (4)$$

Despite its simplifying assumption, Naive Bayes remains effective across a wide range of classification tasks, particularly in high-dimensional and complex datasets [16].

### Model Evaluation

Model evaluation is conducted using a 10-fold cross-validation approach to ensure robust and generalizable results. This technique allows each subset of the data to serve as a testing set in turn, thereby reducing evaluation bias and improving reliability.

The evaluation metrics employed in this study include accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). In the context of imbalanced datasets, recall and F1-score are considered primary indicators, as they provide a more representative assessment of the model's ability to correctly identify instances from the minority class [17]. The F1-score is computed as follows:

$$F1 = 2 \frac{Precision \cdot Recall}{Precision + Recall} \quad (5)$$

The use of this metric is strongly recommended in classification tasks involving imbalanced data distributions, as it captures the trade-off between precision and recall in a balanced and informative manner.

### Experimental Design

The experimental workflow in this study is presented in Figure 1. The process begins with data collection, in which adolescent mental health data related to social media usage were obtained through questionnaire-based measurement. After the data were collected, preprocessing was performed through data cleaning, missing value handling, categorical encoding, and feature normalization.

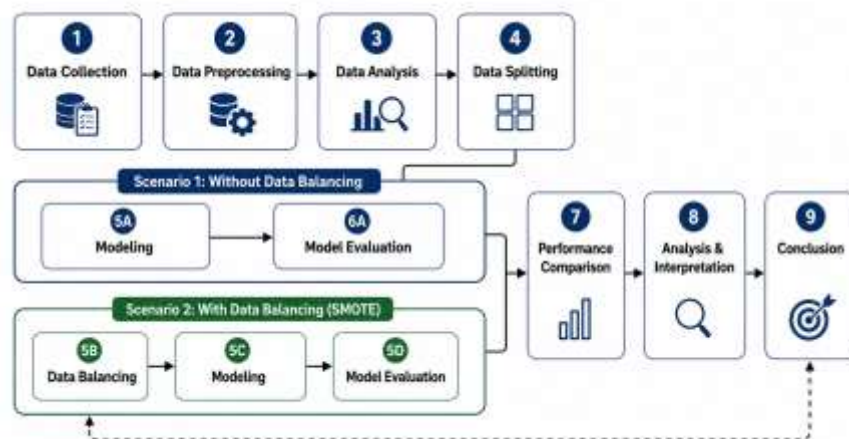


Figure 1. Workflow of the experimental design for comparing KNN and Naive Bayes under imbalanced and balanced data conditions.

After preprocessing, exploratory data analysis was conducted to identify the characteristics of the dataset, including the distribution of the target class. The dataset was then evaluated using a 10-fold cross-validation scheme to ensure that model testing was

not dependent on a single training-testing split.

The experiment was divided into two main scenarios. In the first scenario, the original imbalanced dataset was directly classified using KNN and Naive Bayes. The results of both models were then measured using accuracy, precision, recall, F1-score, and AUC.

In the second scenario, SMOTE was applied to the training data to increase the representation of the minority class. After balancing, the same algorithms, namely KNN and Naive Bayes, were trained and evaluated using the same performance metrics. This design makes it possible to compare model behavior before and after data balancing.

Finally, the performance results from both scenarios were analyzed to identify the effect of SMOTE on each algorithm. The interpretation focused on changes in accuracy, precision, recall, F1-score, and AUC, especially the trade-off between precision and recall in detecting depression-risk cases.

## RESULTS AND DISCUSSION

As the first stage of result presentation, this study displays a portion of the dataset to provide an overview of the attributes used in the classification process. The dataset contains several behavioral and psychological indicators related to adolescents, including social media usage duration, sleep duration, stress level, anxiety level, physical activity, and depression label.

Table 1. Sample of Raw Dataset Used in the Study

| social media hours | sleep hours | stress level | anxiety level | physical activity | depression label |
|--------------------|-------------|--------------|---------------|-------------------|------------------|
| 5.2                | 6.5         | High         | Medium        | Low               | 1                |
| 2.1                | 7.8         | Low          | Low           | High              | 0                |
| 6.0                | 5.9         | High         | High          | Low               | 1                |
| .....              | .....       | .....        | .....         | .....             | .....            |
| 3.5                | 7.0         | Medium       | Medium        | Medium            | 0                |
| 4.8                | 6.2         | High         | High          | Low               | 1                |

The target class distribution shows a highly unequal proportion between the two categories. The non-depression class accounts for most of the observations, while the depression class appears in a much smaller number. This condition confirms that the dataset is imbalanced and requires an appropriate strategy to reduce potential classification bias.

Table 2. Class Distribution of the Dataset

| Class              | Count | Percentage |
|--------------------|-------|------------|
| Non-Depression (0) | 818   | 97.38%     |
| Depression (1)     | 22    | 2.62%      |

The evaluation results without applying data balancing techniques indicate that Naive Bayes outperforms KNN. As shown in Table 3, Naive Bayes achieves an accuracy of 98.81%, with perfect precision (100.00%), a recall of 54.55%, and an F1-score of 70.59%. The high AUC value (0.996) further demonstrates the model's excellent ability to distinguish between classes. In contrast, KNN without SMOTE exhibits substantially lower performance in recall (18.18%) and F1-score (29.63%), despite maintaining relatively high accuracy (97.74%). This finding suggests that KNN is less effective in identifying minority class instances under imbalanced data conditions.

Table 3. Performance Comparison of Classification Models Without SMOTE

| Model       | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) | AUC   |
|-------------|--------------|---------------|------------|--------------|-------|
| Naive Bayes | 98.81        | 100.00        | 54.55      | 70.59        | 0.996 |

|           |       |       |       |       |       |
|-----------|-------|-------|-------|-------|-------|
| KNN (k=3) | 97.74 | 80.00 | 18.18 | 29.63 | 0.797 |
|-----------|-------|-------|-------|-------|-------|

Following the application of SMOTE, notable changes in model performance are observed, particularly for KNN. As presented in Table 4, Naive Bayes shows only a marginal improvement in recall (55.56%), but experiences a significant decrease in precision (45.45%), resulting in a lower F1-score (50.00%) and AUC (0.880). Conversely, KNN demonstrates a substantial increase in recall to 90.91%, although its precision decreases to 24.39%. This trade-off leads to an F1-score of 38.32% and an AUC of 0.961, indicating improved capability in detecting minority class instances, albeit at the cost of increased misclassification.

Table 4. Performance Comparison of Classification Models with SMOTE

| Model               | Accuracy (%) | Precision (%) | Recall (%) | F1-score (%) | AUC   |
|---------------------|--------------|---------------|------------|--------------|-------|
| Naive Bayes + SMOTE | 97.00        | 45.45         | 55.56      | 50.00        | 0.880 |
| KNN (k=3) + SMOTE   | 92.38        | 24.39         | 90.91      | 38.32        | 0.961 |

Further analysis using the confusion matrix for the KNN model with SMOTE reveals that the model successfully identifies the majority of depression cases, correctly classifying 20 out of 22 instances. However, it also produces a considerable number of misclassifications in the non-depression class, with 62 instances incorrectly predicted as depression. This result highlights the trade-off between precision and recall, where increased sensitivity toward the minority class leads to a decline in overall prediction precision.

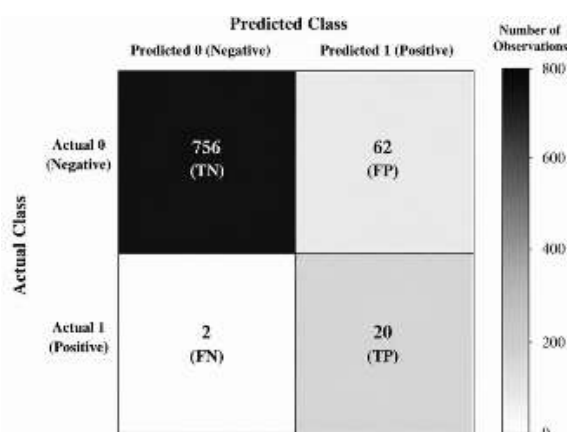


Figure 2. Confusion matrix of the KNN model with SMOTE, showing the distribution of prediction outcomes across actual and predicted classes.

Overall, the findings indicate that Naive Bayes provides superior performance in terms of accuracy and model stability, whereas KNN combined with SMOTE is more effective in enhancing the detection of minority class instances. These results emphasize that model selection should be aligned with the specific objectives of the analysis, particularly in early mental health detection, where sensitivity is often prioritized over overall classification accuracy.

Based on the results presented in the previous section, this study demonstrates that the performance of classification algorithms is strongly influenced by the characteristics of data distribution, particularly class imbalance, which is commonly observed in digital mental health datasets [18]. The finding that Naive Bayes achieves high accuracy and precision indicates that this probabilistic algorithm tends to optimize predictions toward the majority class. This observation is consistent with I. H. Sarker, who argues that Naive Bayes exhibits high stability when applied to imbalanced data distributions [19]. Similarly, research by Didit Johar Maulana et al. reports that while Naive Bayes maintains strong performance in terms of accuracy, it has limitations in effectively detecting minority classes [20].

This limitation is reflected in the relatively low recall values obtained by the Naive Bayes model, indicating reduced sensitivity in identifying depression cases as the minority class. This finding aligns with the study conducted by Cheng Liu et al., which highlights that conventional machine learning models often struggle to detect rare mental health cases in imbalanced datasets [21]. Therefore, despite its superior global accuracy, Naive Bayes may be less suitable for early detection applications that require high sensitivity.

In contrast, the significant improvement in recall observed in the KNN model after the application of SMOTE suggests that distance-based algorithms are highly dependent on data distribution within the feature space [5]. This result is supported by findings from Muhammad Mujahid et al., who state that oversampling techniques such as SMOTE enhance the representation of minority classes, thereby improving the performance of instance-based algorithms like KNN [22]. Furthermore, E. F. Agyemang et al. confirm that KNN demonstrates substantial improvements in classification performance after data balancing, particularly in medical classification tasks [23].

However, the increase in recall for KNN is accompanied by a notable decline in precision, illustrating the classical trade-off between sensitivity and specificity. This phenomenon has been widely documented in the literature, where improvements in minority class detection are often associated with an increase in false positive rates [11]. A study by S. Riyanto et al. emphasizes that in imbalanced datasets, metrics such as F1-score provide a more reliable evaluation than accuracy alone, as they better capture the balance between precision and recall [24]. Accordingly, the performance of KNN with SMOTE in this study should be interpreted primarily in terms of its detection capability (recall) and metric balance (F1-score), rather than accuracy alone.

From a theoretical perspective, the findings reinforce the notion that no single algorithm universally outperforms others across all dataset conditions; instead, performance is highly dependent on data characteristics and analytical objectives, as highlighted by S. Uddin et al. Moreover, these results align with the bias-variance trade-off theory, where simpler models such as Naive Bayes tend to exhibit high bias but low variance, whereas KNN demonstrates higher variance and is more sensitive to data distribution [25]. Additionally, this study shows that data balancing techniques such as SMOTE function as moderating factors that can enhance the performance of certain algorithms, although their impact is not universally beneficial across all models.

From a practical standpoint, the findings have important implications for the development of early detection systems for mental health based on digital data. In the context of initial screening, models with high recall, such as KNN with SMOTE, are more suitable, as they minimize the risk of failing to identify individuals who may be experiencing depression [26]. This is consistent with recommendations from Nasirul Mumenin et al., who emphasize the importance of sensitivity in digital mental health systems to prevent delayed intervention [27]. Conversely, for applications requiring high confidence in prediction outcomes, such as clinical decision support systems, models with high precision, such as Naive Bayes, are more appropriate [28].

Overall, this study demonstrates that the combination of classification algorithms with data balancing techniques (e.g., SMOTE, ADASYN, oversampling, and undersampling) plays a crucial role in improving accuracy, recall, and F1-score in depression and mental health detection models [29]. Furthermore, these findings reinforce the previously observed trade-off between precision and recall, providing a more comprehensive understanding of model behavior under imbalanced data conditions. The results also open opportunities for future research to explore more advanced approaches, such as ensemble learning and deep learning, in order to achieve a more optimal balance between precision and recall.

## CONCLUSION

This study concludes that algorithm selection and imbalance handling strategies have a substantial influence on the performance of machine learning models for adolescent mental health risk detection based on social media usage behavior. The results indicate that model effectiveness cannot be generalized across all data conditions because each algorithm responds differently to class distribution and evaluation objectives. Naive Bayes demonstrates more stable performance, particularly in terms of accuracy and precision, making it suitable for classification scenarios that require reliable overall prediction. Meanwhile, K-Nearest Neighbor combined with SMOTE shows a strong improvement in recall for the minority class, indicating its relevance for early screening purposes where identifying potential risk cases is more important than minimizing false positive predictions.

The main contribution of this study lies in its comparative evaluation of two classification algorithms under imbalanced and balanced data conditions. By examining the effect of SMOTE, this research provides a clearer understanding of how data balancing can change model behavior, especially in mental health classification tasks where the number of risk cases is often much smaller than non-risk cases. Theoretically, this study reinforces the importance of considering data distribution, bias-variance characteristics, and evaluation metrics when selecting classification models. Practically, the findings can support the development of data-driven early detection systems by helping researchers and practitioners choose models according to specific objectives, whether for broad initial screening or for prediction scenarios that require higher precision.

However, this study has several limitations. First, the dataset contains a very small number of minority-class instances, which may limit the generalizability of the findings. Second, the study only evaluates two basic machine learning algorithms, namely KNN and Naive Bayes, without involving more complex models such as ensemble learning or deep learning. Third, the variables used in this research are limited to general behavioral indicators and do not include richer data sources, such as text-based social media content, emotional expressions, or real-time digital activity patterns.

Future research is therefore recommended to use larger and more diverse datasets so that model performance can be validated across broader adolescent populations. Further studies should also compare additional algorithms, including Random Forest, Gradient Boosting, ensemble models, and deep learning architectures, to obtain a more comprehensive understanding of model performance in mental health detection. In addition, the integration of multimodal data sources may improve prediction quality and provide a more complete representation of adolescent mental health conditions. Overall, this study can serve as an initial reference for developing more adaptive, sensitive, and practical machine learning-based mental health screening systems.

## REFERENCES

- [1] H. Lahti, M. Kokkonen, L. Hietajärvi, N. Lyyra, and L. Paakkari, "Social media threats and health among adolescents: evidence from the health behaviour in school-aged children study," *Child and Adolescent Psychiatry and Mental Health* 2024 18:1, vol. 18, no. 1, pp. 62-, May 2024, doi: 10.1186/S13034-024-00754-8.
- [2] F. Mougharbel et al., "Heavy social media use and psychological distress among adolescents: the moderating role of sex, age, and parental support," *Front. Public Health*, vol. 11, p. 1190390, Jun. 2023, doi: 10.3389/FPUBH.2023.1190390/FULL.
- [3] J. L. Kalman et al., "Digitalising mental health care: Practical recommendations from the European Psychiatric Association," *European Psychiatry*, vol. 67, no. 1, p. e4, 2024, doi: 10.1192/J.EURPSY.2023.2466.

- [4] K. F. Yuen, X. Wang, T. Kyriazos, and M. Poga, "Application of Machine Learning Models in Social Sciences: Managing Nonlinear Relationships," *Encyclopedia* 2024, Vol. 4, Pages 1790-1805, vol. 4, no. 4, pp. 1790-1805, Nov. 2024, doi: 10.3390/ENCYCLOPEDIA4040118.
- [5] C. El Morr et al., "Predictive Machine Learning Models for Assessing Lebanese University Students' Depression, Anxiety, and Stress During COVID-19," *J. Prim. Care Community Health*, vol. 15, Jan. 2024, doi: 10.1177/21501319241235588.
- [6] N. Ukey, Z. Yang, B. Li, G. Zhang, Y. Hu, and W. Zhang, "Survey on Exact kNN Queries over High-Dimensional Data Space," *Sensors* 2023, Vol. 23, Page 629, vol. 23, no. 2, p. 629, Jan. 2023, doi: 10.3390/S23020629.
- [7] S. Uddin and H. Lu, "Dataset meta-level and statistical features affect machine learning performance," *Scientific Reports* 2024 14:1, vol. 14, no. 1, pp. 1670-, Jan. 2024, doi: 10.1038/s41598-024-51825-x.
- [8] M. Xiao, "Electroencephalogram Emotion Recognition via AUC Maximization," Aug. 2024, Accessed: Apr. 30, 2026. [Online]. Available: <https://arxiv.org/pdf/2408.08979>
- [9] S. F. Taskiran, B. Turkoglu, E. Kaya, and T. Asuroglu, "A comprehensive evaluation of oversampling techniques for enhancing text classification performance," *Scientific Reports* 2025 15:1, vol. 15, no. 1, pp. 21631-, Jul. 2025, doi: 10.1038/s41598-025-05791-7.
- [10] A. G. Putrada, I. D. Wijaya, and D. Oktaria, "Overcoming Data Imbalance Problems in Sexual Harassment Classification with SMOTE," *International Journal on Information and Communication Technology (IJOICT)*, vol. 8, no. 1, pp. 20-29, Aug. 2022, doi: 10.21108/IJOICT.V8I1.622.
- [11] A. Islam, S. B. Belhaouari, A. U. Rehman, and H. Bensmail, "KNNOR: An oversampling technique for imbalanced datasets," *Appl. Soft Comput.*, vol. 115, p. 108288, Jan. 2022, doi: 10.1016/J.ASOC.2021.108288.
- [12] İ. Baydili, B. Tasci, and G. Tasci, "Deep Learning-Based Detection of Depression and Suicidal Tendencies in Social Media Data with Feature Selection," *Behavioral Sciences* 2025, Vol. 15, Page 352, vol. 15, no. 3, p. 352, Mar. 2025, doi: 10.3390/BS15030352.
- [13] D. Elreedy et al., "A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning," *Machine Learning* 2023 113:7, vol. 113, no. 7, pp. 4903-4923, Jan. 2023, doi: 10.1007/S10994-022-06296-4.
- [14] M. Pagan, M. Zarlis, and A. Candra, "Investigating the impact of data scaling on the k-nearest neighbor algorithm," *Computer Science and Information Technologies*, vol. 4, no. 2, pp. 135-142, Jul. 2023, doi: 10.11591/CSIT.V4I2.P135-142.
- [15] V. Wijaya and N. Rachmat, "Comparison of SVM, Random Forest, and Logistic Regression Performance n Student Mental Health Screening," *JEECS (Journal of Electrical Engineering and Computer Sciences)*, vol. 9, no. 2, pp. 173-184, Dec. 2024, doi: 10.54732/JEECS.V9I2.9.
- [16] S. S. Prasetyowati and Y. Sibaroni, "Unlocking the potential of Naive Bayes for spatio temporal classification: a novel approach to feature expansion," *Journal of Big Data* 2024 11:1, vol. 11, no. 1, pp. 106-, Aug. 2024, doi: 10.1186/S40537-024-00958-X.
- [17] M. Abdelhamid and A. Desai, "Balancing the Scales: A Comprehensive Study on Tackling Class Imbalance in Binary Classification," Sep. 2024, Accessed: Apr. 30, 2026. [Online]. Available: <https://arxiv.org/pdf/2409.19751>
- [18] E. Gentili et al., "Machine learning from real data: A mental health registry case study," *Computer Methods and Programs in Biomedicine Update*, vol. 5, p. 100132, Jan. 2024, doi: 10.1016/J.CMPBUP.2023.100132.

- [19] I. H. Sarker, A. S. M. Kayes, and P. Watters, "Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage," *Journal of Big Data* 2019 6:1, vol. 6, no. 1, pp. 57-, Jul. 2019, doi: 10.1186/S40537-019-0219-Y.
- [20] D. J. Maulana, S. Saadah, and P. E. Yunanto, "Kmeans-SMOTE Integration for Handling Imbalance Data in Classifying Financial Distress Companies using SVM and Naïve Bayes," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 8, no. 1, pp. 54-61, Feb. 2024, doi: 10.29207/RESTI.V8I1.5140.
- [21] C. Liu, L. Ji, J. Lu, J. Ma, and X. Sui, "College student mental health assessment: Predictive models based on machine learning and feature importance analysis," *Molecular & Cellular Biomechanics*, vol. 22, no. 3, pp. 1477-1477, Feb. 2025, doi: 10.62617/MCB1477.
- [22] M. Mujahid et al., "Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering," *Journal of Big Data* 2024 11:1, vol. 11, no. 1, pp. 87-, Jun. 2024, doi: 10.1186/S40537-024-00943-4.
- [23] E. F. Agyemang et al., "Addressing Class Imbalance Problem in Health Data Classification: Practical Application From an Oversampling Viewpoint," *Applied Computational Intelligence and Soft Computing*, vol. 2025, no. 1, p. 1013769, Jan. 2025, doi: 10.1155/ACIS/1013769.
- [24] S. Riyanto, I. S. Sitanggang, T. Djatna, and T. D. Atikah, "Comparative Analysis using Various Performance Metrics in Imbalanced Data for Multi-class Text Classification," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 6, pp. 1082-1090, Jun. 2023, doi: 10.14569/IJACSA.2023.01406116.
- [25] E. Ahmed, "Student Performance Prediction Using Machine Learning Algorithms," *Applied Computational Intelligence and Soft Computing*, vol. 2024, no. 1, p. 4067721, Jan. 2024, doi: 10.1155/2024/4067721.
- [26] Z. Sabouri, N. Gherabi, M. Nasri, M. Amnai, H. El Massari, and I. Moustati, "Prediction of Depression via Supervised Learning Models: Performance Comparison and Analysis," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 19, no. 09, pp. 93-107, Jul. 2023, doi: 10.3991/IJOE.V19I09.39823.
- [27] N. Mumenin et al., "Screening depression among university students utilizing GHQ-12 and machine learning," *Heliyon*, vol. 10, no. 17, Sep. 2024, doi: 10.1016/j.heliyon.2024.e37182.
- [28] D. Widyawati, A. Faradibah, P. Lestari, and L. Belluano, "Comparison Analysis of Classification Model Performance in Lung Cancer Prediction Using Decision Tree, Naive Bayes, and Support Vector Machine," *Indonesian Journal of Data and Science*, vol. 4, no. 2, pp. 78-86, Jul. 2023, doi: 10.56705/IJODAS.V4I2.76.
- [29] T. R. Shaha et al., "Feature group partitioning: an approach for depression severity prediction with class balancing using machine learning algorithms," *BMC Medical Research Methodology* 2024 24:1, vol. 24, no. 1, pp. 123-, Jun. 2024, doi: 10.1186/S12874-024-02249-8.