Classification of Anxiety Levels of IGD Patients at RSU Royal Prima Medan Using Support Vector Machine (SVM) Algorithm

Kharisma Gunanta Ginting[1*], Nugroho Prasetyo[2], Al Vino Gunawan[3], Magdalena Sihombing[4], Adli Abdillah Nababan[5]

[1,2,3,4,5]Informatics Engineering Department, Universitas Prima Indonesia, Medan, North Sumatra, Indonesia
E-mail: [1*]gunantakharisma03@gmail.com, [2]nugsfree@gmail, [3]vinoal683@gmail.com, [4]magdalenasihombing17@gmail, [5]adliabdillahnababan@unprimdn.ac.id,

(Received: 5 May 2025, revised: 23 May 2025, accepted: 26 May 2025)

Abstract

The level of patient anxiety in the Emergency Department (ED) is an important indicator that affects the diagnosis and medical management process. However, the classification of anxiety levels is often hampered by data imbalance, which can reduce the accuracy of predictive models. This study aims to develop a patient anxiety level classification model in the ED using the Support Vector Machine (SVM) algorithm with the application of the Synthetic Minority Oversampling Technique (SMOTE) to address the class imbalance issue. The data used consists of 734 patient samples divided into 80% training data and 20% testing data, including physiological parameters such as systolic and diastolic blood pressure, respiratory rate, heart rate, and demographic data such as age and gender. The preprocessing process includes imputing missing values and normalizing numerical features so that the model can learn optimally. Performance evaluation shows that the use of SMOTE increases classification accuracy from 95% to 97%, as well as improving precision, recall, and F1-score metrics at almost all anxiety levels. Visualization of the relationships between numerical features also reveals significant correlation patterns between physiological variables and patient anxiety levels. The results of this study confirm the effectiveness of SMOTE in addressing data imbalance, thus producing a more accurate anxiety level classification model that can serve as an aid in clinical decision-making. Thus, the developed model has significant clinical utility potential as a diagnostic aid that can accelerate and improve the accuracy of patient management in the emergency department, thereby supporting the overall improvement of healthcare service quality.

Keywords: SMOTE, Support Vector Machine, Anxiety Classification, Oversampling, Emergency Installation.

I. INTRODUCTION

Royal Prima Hospital Medan, as one of the largest private hospitals in North Sumatra, was established with the vision to excel in healthcare services, education, and research, and with the mission to provide quality healthcare services, enhance human resource competencies, and establish partnerships for the benefit of public health [1]. One important part of this hospital is the Emergency Room (ER), which serves patients who are sick and injured, functioning as the main entry point for patients by performing triage based on priority. The ER has also seen a 30% increase in visits globally, including 4.402.205 patients in Indonesia, which accounts for 13.3% of the total patients in general hospitals [2]. Locally, Royal Prima Medan recorded an average of 4.199 patient visits per month in the ER, indicating a high service load at this facility [3]. However, it is important to note that depression and anxiety in patients can have a negative impact on their health conditions, such as decreased quality of life, delays in receiving treatment, non-compliance with treatment plans, as well as worsening prognosis and survival [4].

Anxiety is a condition characterized by feelings of discomfort, worry, and tension that are vague and subjective without a clear trigger, and is influenced by biological, psychological, and social factors, showing a close relationship between physical and mental health [5] [6]. For individuals experiencing anxiety disorders, feelings of anxiety tend to arise continuously, are difficult to control, and are often accompanied by other physical symptoms [7]. This becomes more real when patients enter healthcare facilities, where fears and anxieties often arise, especially when they receive care in the Emergency Department [8].

SVM algorithm is one of the supervised machine learning algorithms based on statistical learning theory [9]. The Support Vector Machine (SVM) algorithm has a basic principle of linearly separable classification, but it has also been developed to handle non-linear problems. In the process, this algorithm learns from input data to determine object labels and has been

DOI: 10.34148/teknika.v14i2.1243

CC BY SA

successfully applied in various biological applications to predict disease categories. SVM has the ability to determine the distance between two sets of data from different classes as well as the hyperplane that optimally separates them, where the linear formulation of SVM has a hyperplane located in the middle of the input data space [10] [11].

In recent years, anxiety classification using machine learning techniques has seen significant advancements. Recent studies show that machine learning approaches, including SVM algorithms and deep learning, are able to improve the accuracy of anxiety level detection and classification by effectively utilizing physiological and psychological data [12] [13]. Research by Foysal et al. (2024) highlights a comprehensive analysis of the stress factors affecting anxiety, which is relevant for the development of better predictive models [12]. In addition, a recent review of the transition from machine learning to deep learning in the field of medicine emphasizes the great potential of this technology in enhancing the diagnosis and management of mental conditions such as anxiety [13].

II. RESEARCH METHODOLOGY

A. Data Collection

The data used to classify the level of anxiety of patients in the emergency department at the hospital was collected through medical records and direct observations of patients who came to the emergency department (Figure 1). The data obtained includes various variables relevant to the level of anxiety of the patients, such as:

1. Demographic Data: Including age, gender, and health history of the patient that may affect the level of anxiety.
2. Medical Parameters: Such as systolic, diastolic, respiratory rate, and heart rate that can indicate the physical condition of the patient related to anxiety.
3. Arrival Time and Emergency Conditions: Information about the patient's arrival time to the ER and the urgency level of the case, which can affect the patient's perception and level of anxiety.

This data is then processed and cleaned to ensure its quality and accuracy. After the pre-processing stage, the structured data is ready to be used for training and testing the SVM model. This process involves selecting relevant features, normalizing the data, and splitting the data into training and testing sets to achieve optimal classification results.

The level of anxiety in patients is classified into four categories based on the intensity of the symptoms experienced, namely Level 0 to Level 3. Level 0 (Normal) indicates that the patient is without significant signs of anxiety, thus not affecting the clinical process meaningfully. Level 1 (Mild Anxiety) describes patients experiencing mild anxiety symptoms that remain within tolerable limits and do not substantially disrupt clinical functions. At Level 2 (Moderate Anxiety), patients exhibit more pronounced anxiety that begins to affect the diagnosis and medical management process, thus requiring more attention. Meanwhile, Level 3 (Severe Anxiety) represents a condition where the patient experiences severe anxiety that can significantly hinder medical processes and requires special intervention to ensure smooth management.

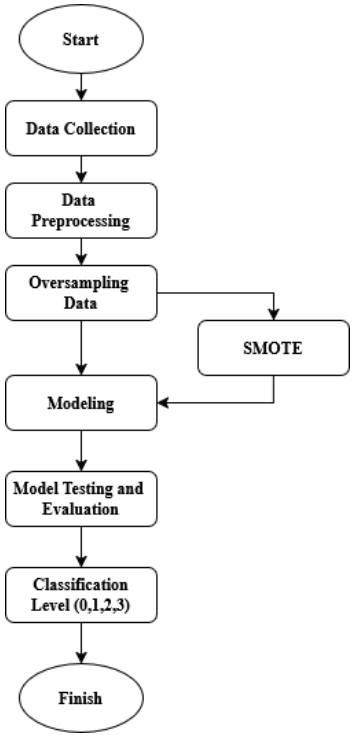


Figure 1. Research Working Procedures

B. Data Preprocessing

Research data was collected from direct observations at the emergency installation of Royal Prima General Hospital Medan, covering physiological parameters such as systolic and diastolic blood pressure, respiratory rate, heart rate, as well as demographic data such as age and gender (Figure 2). In the initial stage, an examination of the data was conducted to identify missing or incomplete values. These values were then filled using the median method to avoid significantly affecting the data distribution.

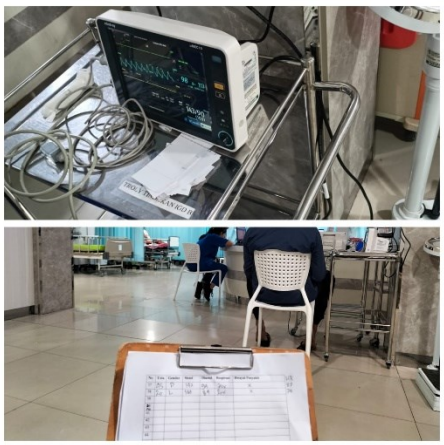


Figure 2. Observation of patient data with Mindray uMEC 10

Next, the categorical variables are converted into numeric format through an encoding process to be processed by the machine learning algorithm. All numeric features are then normalized using StandardScaler to ensure that each feature has a uniform scale so that the SVM model can perform optimally. Since the dataset has class imbalance in the anxiety level classification target, the SMOTE (Synthetic Minority Over-sampling Technique) is used to balance the number of samples between classes by creating synthetic samples in the minority class. This is important to prevent the model from being biased towards the majority class and to be able to recognize patterns from all classes effectively.

Preprocessing begins with the examination and handling of missing or incomplete values in the dataset. Missing values are filled using the median imputation method to maintain data consistency. Next, encoding is performed on categorical variables so that they can be processed by machine learning algorithms. For the gender variable, males are coded as 1 and females as 0. Meanwhile, for the hypertension history variable, patients with a history of hypertension are coded as 1 and those without a history of hypertension as 0.

Numerical features such as systolic/diastolic blood pressure, respiratory rate, heart rate, and age are normalized using standardization techniques to have a uniform scale so that the model can learn optimally. As part of the further data processing strategy, this research will compare two approaches: one dataset will be processed using the SMOTE (Synthetic Minority Over-sampling Technique) to address class imbalance on the target label; while the other set will not use SMOTE. The results of both approaches will then be compared to determine the impact of balancing on the accuracy performance of the SVM model in classifying the anxiety levels of emergency room patients.

C. Modeling

Support Vector Machine (SVM) utilizes labeled training data to determine the hyperplane that serves as the classification boundary between two classes of objects, similar to supervised learning models [14]. In the application of SVM, the training process is conducted using a dataset where the relationships between features and their group memberships are known, with the aim of forming a classifier. Once the classifier is formed, it is used to test new data during the validation phase. Ideally, validation is performed using an independent dataset; however, this method requires a large number of samples, which is often difficult to obtain. As an alternative, if the number of samples is limited, cross-validation can be used to assess the performance of the SVM classification by dividing the data into several subsets that are alternately used for training and testing. [15]. The kernel function is a function that serves to transform data into a higher dimension. The purpose of this transformation is to improve the structure of the data, making the process of separation between classes easier. In this context, the general formula for linear SVM can be expressed as Formula 1 follows:

$$f(x) = sign(w.x + b) \quad (1)$$

In this context, f(x) is the predictive function that determines the class of the input data, where (w) is the normal vector of the hyperplane, (x) is the input feature vector, and (b) is the bias or intercept that determines the position of the hyperplane [16].

SMOTE is a synthetic oversampling technique used to address the problem of data imbalance in classification. This technique works by creating new synthetic data for the minority class based on its nearest neighboring samples [17]. SMOTE creates synthetic data by interpolating between minority samples and their nearest neighbors. This method effectively improves model performance, with the chosen k value influencing the prediction results [18]. This algorithm is also capable of reducing the time required in the process of searching for optimal parameter values in the SMOTE algorithm. The mathematical formula for forming synthetic data in SMOTE is explained as Formula 2 follows:

$$DS_{syn} = DS_i + (DS_{knn} - DS_i) \times rand(0,1) \quad (2)$$

In the SMOTE formula, DS_{syn} is the new synthetic data created by interpolating between the original minority samples (DS_i) and their nearest neighbors (DS_{knn}). This process uses a random number between 0 and 1 (rand(0,1)) to determine the position of the new data along the line connecting the two points [19].

Kernel Radial Basis Function (RBF), also known as the Gaussian kernel, is a kernel function that measures the similarity between two data points based on their Euclidean distance. Mathematically, the RBF kernel is defined as shown in Formula 3:

$$K(X_1, X_2) = exp(-\frac{||X_1 - X_2||^2}{2\sigma^2}) \quad (3)$$

where σ is the bandwidth parameter that controls the scale of the distance effect. The value of this kernel ranges from 0 to 1, with a maximum value of 1 when both points are identical, and approaching 0 as the distance between points increases [20] [21].

III. RESULT AND DISCUSSION

The Support Vector Machine (SVM) model used in this analysis employs the default Radial Basis Function (RBF) kernel, as configured in the scikit-learn library. This kernel is chosen for its superior ability to handle data with non-linear patterns, which is highly relevant to the characteristics of the dataset containing physiological and demographic features such as systolic and diastolic blood pressure, heart rate, age, and other variables that interact in complex ways.

RBF kernel works by mapping data to a higher-dimensional space so that patterns that cannot be linearly separated in the original space can be better separated. The RBF kernel has a gamma parameter that can be adjusted to control the range of influence of each data point, allowing the model to be tailored to avoid underfitting or overfitting.

DOI: 10.34148/teknika.v14i2.1243
CC BY SA

Table 1. Classification Report

Technique	Level	Precision	Recall	F1-score	Support	Average Accuracy	Amount of data
Without SMOTE	0	0.94	0.98	0.96	82	0.95	147
	1	0.92	0.80	0.86	15		
	2	1.00	1.00	1.00	11		
	3	0.97	0.95	0.96	39		
With SMOTE	0	0.95	0.99	0.97	82	0.97	147
	1	0.93	0.87	0.90	15		
	2	1.00	1.00	1.00	11		
	3	1.00	0.95	0.97	39		

The Table 1 presents the evaluation of the classification model's performance under two conditions, namely without the application of SMOTE and with the application of SMOTE. In the condition without SMOTE, the model demonstrates fairly good performance with an average accuracy of 0.95. For the majority class (Level 0), the model achieves a precision value of 0.94 and a recall of 0.98, indicating that the model is able to classify this class very accurately and almost does not miss any samples that actually belong to this class. However, in the minority class such as Level 1, the precision and recall values are 0.92 and 0.80 respectively, indicating that the model is still less sensitive in detecting all samples of this class, leading to some false negatives.

After the application of SMOTE, there was a significant performance increase, especially in the minority class. The precision and recall values at Level 1 increased to 0.93 and 0.87, indicating that the model became more sensitive and able to capture more minority class samples without significantly increasing classification errors. At Level 3, precision increased from 0.97 to 1.00 and F1-score from 0.96 to 0.97, showing a better balance between false positives and false negatives. Level 2 maintained a perfect performance with precision, recall, and F1-score values all at 1.00. Overall, the average accuracy of the model increased to 0.97, indicating that the use of SMOTE not only helped improve performance in the minority class but also maintained or even improved performance in the majority class.

The high precision values across all classes indicate that the model rarely makes errors in classifying samples into inappropriate classes. The increase in recall for the minority class after the implementation of SMOTE indicates that the model has become better at capturing all samples from that class, thereby reducing the risk of false negatives that can negatively impact real-world applications, especially in medical contexts or risk assessment. The increase in F1-score for the minority class also reflects a better balance between precision and recall, which is crucial in classification cases with imbalanced class distributions. The consistent support in both conditions shows that the test data used is consistent, so the observed performance differences purely stem from model changes due to the SMOTE technique.

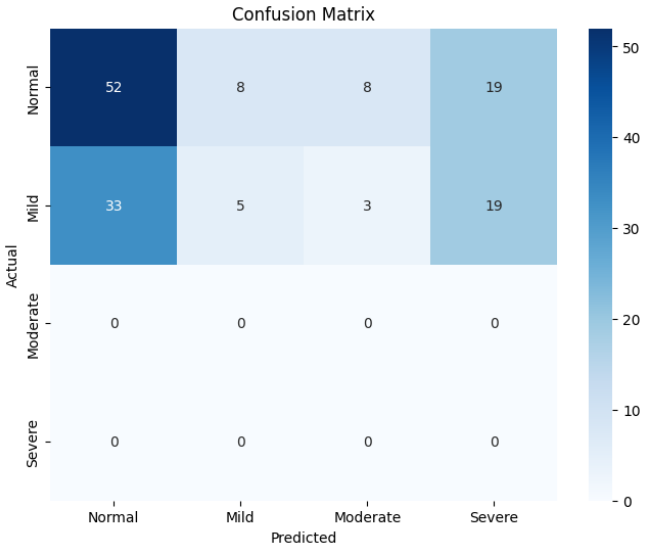


Figure 3. Confussion Matrix Without SMOTE

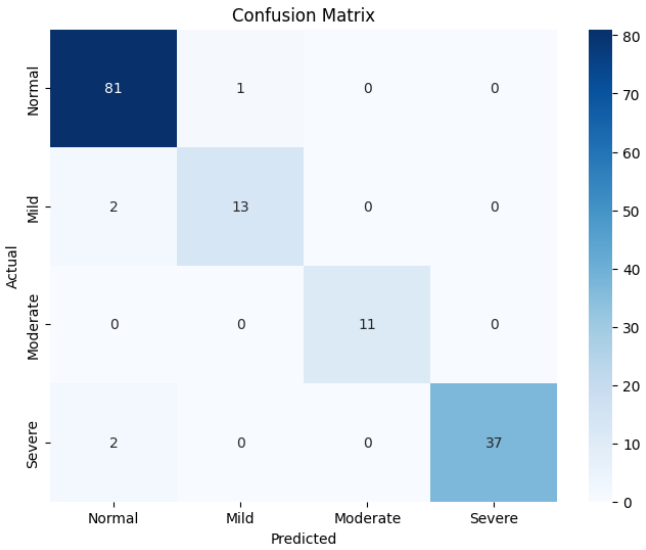


Figure 4. Confussion Matrix With SMOTE

The confusion matrix presented in Figure 3 (without SMOTE) and Figure 4 (with SMOTE) provides a clear picture of the impact of applying SMOTE on the performance of the classification model in identifying levels of anxiety. SMOTE (Synthetic Minority Over-sampling Technique) is a method

DOI: 10.34148/teknika.v14i2.1243
CC BY SA

used to address class imbalance by adding synthetic data to the minority class, thus making the class distribution more balanced and allowing the model to learn more effectively.

In the condition without the application of SMOTE, the confusion matrix shows that the model's performance is suboptimal, especially in recognizing minority classes. The model correctly classified 52 samples in the Normal class, but there were significant misclassifications: 8 samples were classified as Light, 8 as Medium, and 19 as Heavy. This indicates a model bias towards the majority class (Normal), resulting in considerable prediction errors. In the Light class, the model was only able to correctly classify 5 samples, while 33 samples were misclassified as Normal, 3 as Medium, and 19 as Heavy. The Medium and Heavy classes were not detected at all, indicating very high data imbalance and the model's inability to recognize those classes.

After the implementation of SMOTE, there was a significant improvement in performance across all classes, especially the minority class. In the Normal class, the number of correct predictions increased to 81 samples with only 1 misclassification to the Mild class. The Mild class also showed improvement with 13 samples being classified correctly and only 2 misclassifications to the Normal class. The Medium class, which was previously undetected, is now successfully classified with 11 samples correctly without any prediction errors, indicating that SMOTE has successfully addressed the issue of data imbalance, allowing the model to recognize this class well. In the Heavy class, the model was able to classify 37 samples correctly, although there were 2 misclassifications to the Normal class. This indicates that while there are still some minor errors, the model overall is able to recognize the Heavy class well.

This performance improvement is also reflected in other evaluation metrics such as precision, recall, and F1-score, which show higher values under conditions with SMOTE compared to those without SMOTE, as well as an increase in average accuracy from 0.95 to 0.97. This emphasizes that the application of SMOTE not only enhances the model's sensitivity to the minority class but also maintains or even improves performance on the majority class.

Figure 5 presents a comparison of class distribution on the full dataset, training data before SMOTE, and training data after the application of SMOTE. This graph provides a clear picture of the class imbalance present in the original dataset and how SMOTE successfully addresses this issue.

In the first graph, the class distribution in the full dataset shows a significant imbalance, with the Normal class dominating with 410 samples, followed by the Heavy class with 197 samples, while the Light and Medium classes have far fewer samples, with 74 and 53 samples respectively. This imbalance can cause the classification model to be biased towards the majority class, resulting in suboptimal performance on the minority class.

Next, the second graph shows the class distribution in the training data before the application of SMOTE. There is a decrease in the number of samples in each class caused by the data splitting process, but the pattern of imbalance is still clearly visible. The Normal class still has the largest number of samples at 328, while the Heavy class has 158, and the Light and Medium classes have only 59 and 42 samples, respectively. This condition indicates that the training data is still very imbalanced, which could potentially cause the model to struggle to accurately recognize minority classes.

In the third graph, after the application of SMOTE, the class distribution in the training data becomes perfectly balanced, with each class having the same number of samples, which is 328 samples. SMOTE (Synthetic Minority Over-sampling Technique) works by generating synthetic data for the minority classes so that the number of samples in each class becomes equal. This is very important to enhance the model's ability to learn from all classes proportionally without bias towards the majority class.

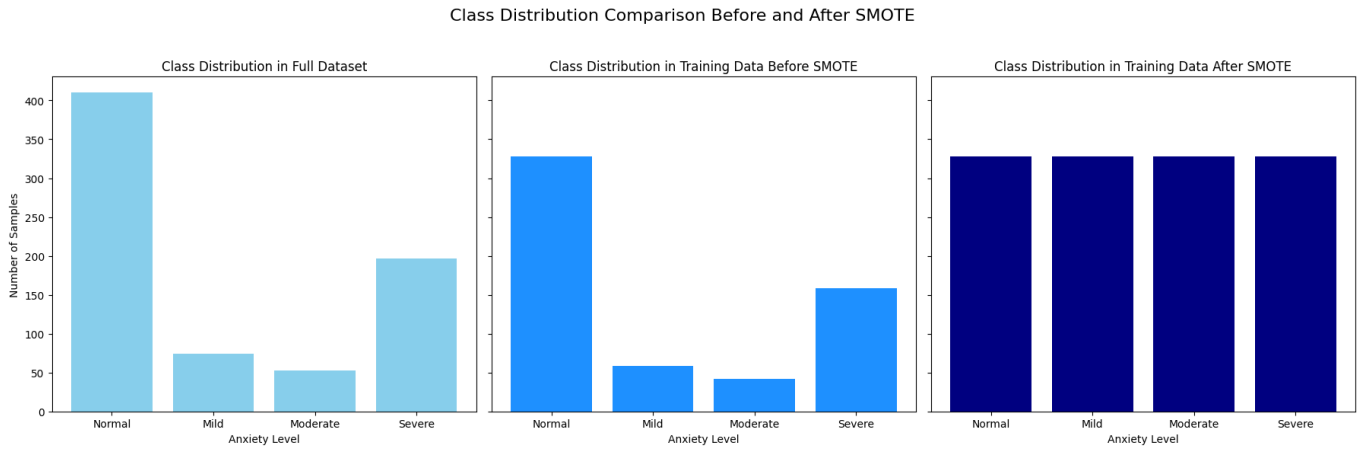


Figure 5. Comparison of Class Distribution Before and After SMOTE

DOI: 10.34148/teknika.v14i2.1243

CC BY SA

The successful application of SMOTE that balances the class distribution directly contributes to the improved performance of the classification model, as seen in the confusion matrix and the previous evaluation metrics. With a balanced class distribution, the model can better recognize patterns in the minority class, thereby enhancing accuracy, recall, and F1-score for those classes. This is crucial in the context of anxiety level classification, where mistakes in identifying the minority class can have serious impacts on diagnosis and treatment.

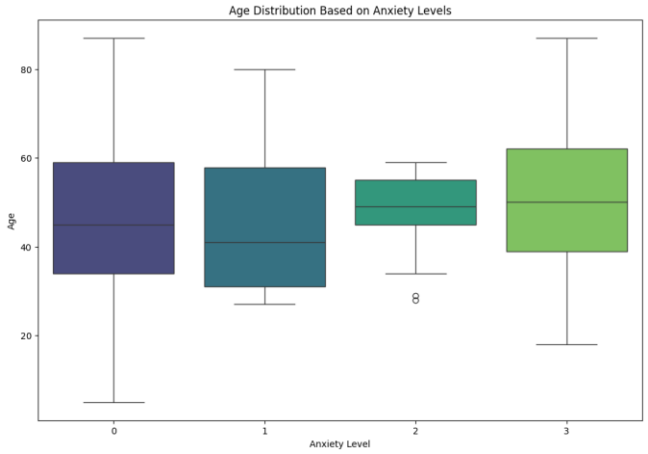


Figure 6. Age Distribution Based on Anxiety Levels

Figure 6 presents a box plot of the age distribution based on four levels of anxiety: Normal (0), Mild (1), Moderate (2), and Severe (3). This analysis uses a descriptive statistical approach to illustrate the age characteristics at each anxiety level. At the Normal (0) level, the age distribution shows a quite broad range, with a minimum value of around 5 years and a maximum close to 90 years. The median age is around 45 years, with a relatively large interquartile range (IQR), indicating high age variation in this group. This indicates that individuals with normal anxiety levels are evenly distributed across various age ranges, from children to the elderly.

Light Level (1) has a slightly narrower age distribution compared to the Normal level, with a median age of about 41 years. The age range at this level is also quite broad, starting from around 27 years to 80 years, but the smaller IQR indicates a more concentrated age focus on the young adult to middle-aged groups. This may suggest that mild anxiety is more commonly found in that age group.

At the Moderate level (2), the age distribution tends to be narrower and the median age is higher, around 48 years. The age range starts from about 34 years to 59 years, with some outliers below the age of 30. This indicates that moderate anxiety is more dominant in the middle-aged group, with little age variation outside that range.

Severe Level (3) shows a fairly wide age distribution with a median age of around 50 years, slightly higher than other levels. The age range starts from about 18 years to 85 years, with a fairly large IQR. This indicates that severe anxiety can occur in various age groups, but tends to be more common in adults to the elderly.

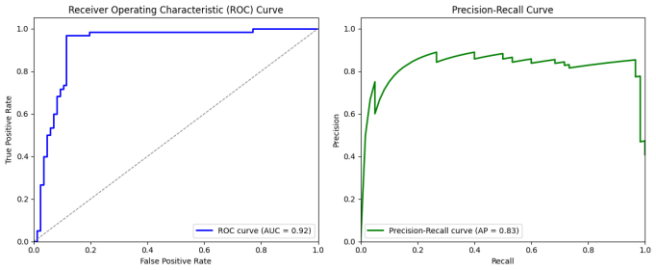


Figure 7. ROC Curve & Precision-Recall Curve

Figure 7 presents two curves evaluating the performance of the classification model, namely the Receiver Operating Characteristic (ROC) Curve and the Precision-Recall Curve, which each provide a comprehensive overview of the model's ability to classify levels of anxiety.

In the ROC curve, the horizontal axis represents the False Positive Rate (FPR) while the vertical axis shows the True Positive Rate (TPR). The resulting ROC curve demonstrates the model's excellent performance with an Area Under the Curve (AUC) of 0.92. An AUC value approaching 1 indicates that the model has a high ability to distinguish between positive and negative classes, thus resulting in a relatively low prediction error rate. In other words, the model is capable of accurately identifying individuals with anxiety levels, minimizing classification errors that could lead to inappropriate interventions.

Meanwhile, the Precision-Recall curve illustrates the trade-off between precision and recall of the model. This curve is very important, especially in the context of imbalanced datasets, such as cases of anxiety levels that have different class distributions. The Precision-Recall curve in Figure 7 shows an Average Precision (AP) of 0.83, indicating that the model is not only able to capture most positive cases (high recall), but also maintains prediction accuracy (high precision). This indicates that the model is effective in reducing both false positives and false negatives, which is crucial in medical or psychological applications to avoid diagnostic errors.

Figure 8 presents a visualization of the pairplot that illustrates the relationship between numerical and categorical variables in the dataset used for classifying anxiety levels. This pairplot displays scatter plots for each pair of features, as well as univariate distributions (histograms and density plots) on the diagonal, with data points colored according to the anxiety level class (0: Normal, 1: Mild, 2: Moderate, 3: Severe). This visualization provides a comprehensive overview of the data distribution patterns and correlations among features that are important for modeling.

From the diagonal histogram, it can be seen that the age feature has a relatively even distribution with a peak in the young adult to middle-aged range, consistent with the previous discussion of age distribution. The gender feature is a binary categorical variable, with a relatively balanced distribution between males and females. Vital features such as systolic, diastolic, respiration, and heart rate show considerable variation, with several distribution peaks indicating the presence of subgroups within the data.

.

DOI: 10.34148/teknika.v14i2.1243
CC BY SA

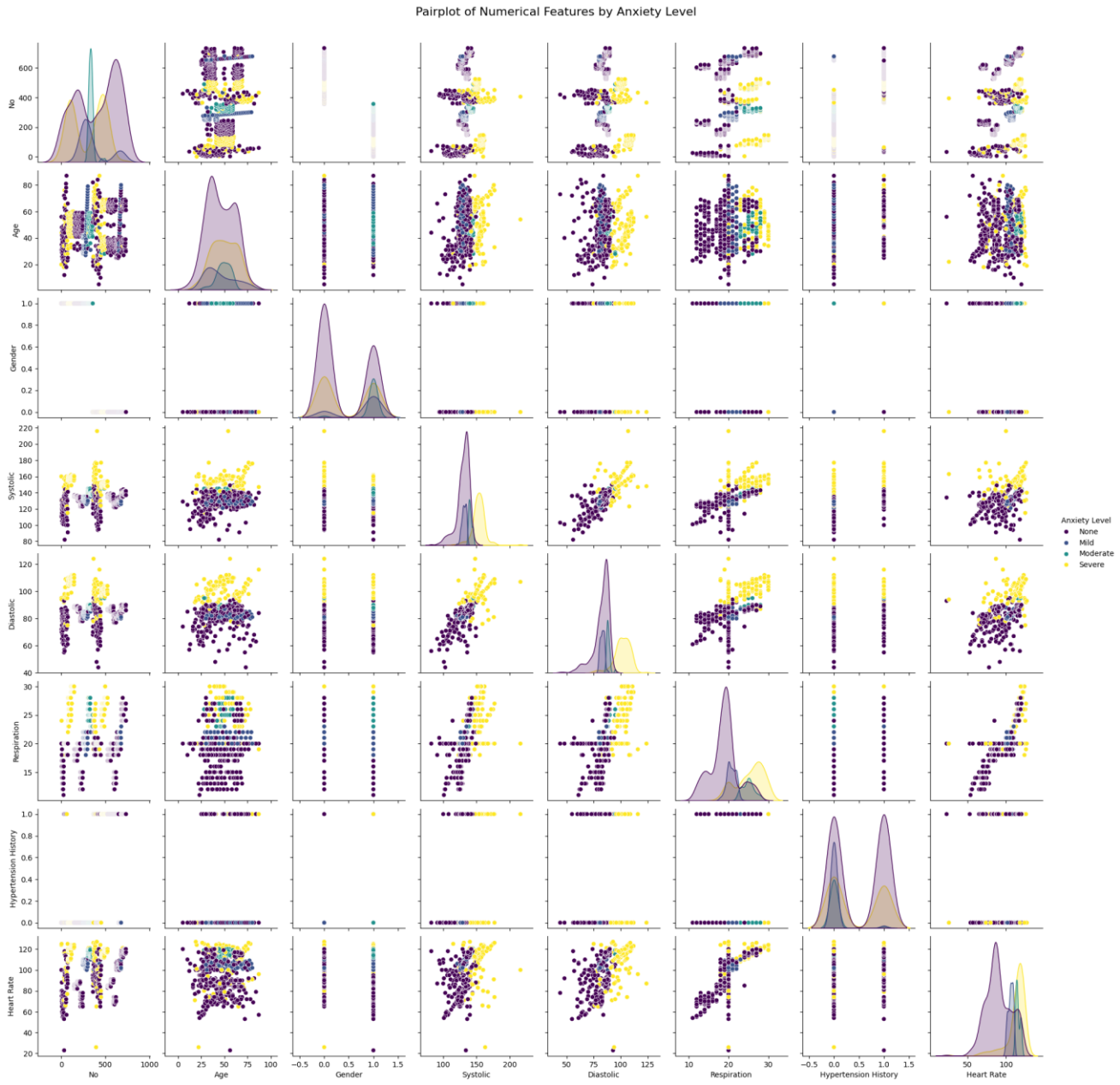


Figure 8. Pairplot

The scatter plot between numerical features reveals a strong positive correlation between systolic and diastolic blood pressure, which is a common physiological relationship. Additionally, there is a moderate correlation between heart rate and respiration, which may reflect the physiological response to anxiety levels. This relationship is important to consider in modeling because these features are interrelated and can influence predictions

The color of the dots representing anxiety levels shows different clustering patterns across several features. For instance, the class with severe anxiety levels (yellow) tends to concentrate at higher systolic and diastolic value ranges, as well as higher heart rates, compared to the normal class (purple). This indicates that increased blood pressure and heart rate are associated with higher anxiety levels, which is consistent with the medical literature.

In the age feature, the point distribution shows that mild and moderate levels of anxiety are spread over a narrower age range compared to normal and severe levels, which supports previous findings regarding age distribution. The gender feature does not show a clear separation pattern among classes, indicating that gender may not be a key predictor variable in this classification.

This pairplot visualization provides important insights that several features have strong correlations and different distribution patterns across anxiety classes. This indicates the need for preprocessing techniques such as normalization and handling multicollinearity so that the model can learn

DOI: 10.34148/teknika.v14i2.1243
TEKNIKA, Volume 14(2), July 2025, pp. 255-263
ISSN 2549-8037, EISSN 2549-8045

effectively. Furthermore, the clustering patterns observed in vital features suggest that the classification model can effectively utilize these features to differentiate anxiety levels.

This study uses observational data from patients in the Emergency Department of Royal Prima Hospital Medan, which introduces a potential bias due to the data source being limited to a single location and population. This limits the representativeness of the data and reduces the research findings' ability to be generalized to a broader population with different characteristics.

The SMOTE technique is applied to address class imbalance in patient anxiety data, which improves the model's performance in recognizing the minority class. However, the use of synthetic data from SMOTE carries the risk of introducing bias because the patterns generated may not fully reflect the natural variation of the original data, resulting in the model potentially experiencing overfitting and being less adaptive to new real data.

IV. CONCLUSION

The Support Vector Machine (SVM) model with a Radial Basis Function (RBF) kernel employed in this study demonstrated robust performance in classifying anxiety levels based on physiological and demographic data. The choice of the RBF kernel was motivated by its capability to effectively handle non-linear data patterns by projecting the input features into a higher-dimensional space, thereby enhancing the model's classification power. A critical hyperparameter, gamma, was carefully tuned to balance the model's bias-variance trade-off, preventing both underfitting and overfitting.

The evaluation was conducted under two experimental conditions: without and with the application of the Synthetic Minority Over-sampling Technique (SMOTE) to mitigate class imbalance. Without SMOTE, the model achieved an average accuracy of 95%, with high precision and recall for the majority class. However, the recall for the minority class (moderate anxiety level) was limited to 80%, indicating suboptimal sensitivity and a risk of false negatives. The introduction of SMOTE significantly improved the model's performance, particularly for the minority class, where precision and recall increased to 93% and 87%, respectively. For the high anxiety class, precision reached 100% with an F1-score of 97%, reflecting a well-balanced trade-off between false positives and false negatives. Overall, the model's accuracy improved to 97%, demonstrating that SMOTE effectively enhances classification performance without compromising the majority class.

The consistent increase in precision suggests a reduction in misclassification errors, while the improved recall for the minority class is particularly important in the context of mental health, where timely and accurate identification of anxiety levels can facilitate early intervention. To maximize the practical utility of this classification model in healthcare settings, it is recommended to integrate the SVM with RBF kernel and SMOTE technique into Hospital Information Systems (SIRS), especially within the triage module or initial patient assessment in Emergency Departments. This integration would enable real-time anxiety level prediction based on patients' physiological and demographic data, thereby supporting healthcare professionals in making prompt and informed decisions.

Future research directions include exploring alternative machine learning algorithms such as Random Forest, Gradient Boosting, Neural Networks, and Deep Learning to potentially enhance accuracy and sensitivity beyond what SVM offers. Validating the model across diverse hospital environments and geographical regions is essential to assess its generalizability and adaptability to varied population characteristics.

REFERENCES

[1] RSU Royal Prima Medan, "Profile of Royal Prima Hospital," [Online]. Available: https://royalprima.com/. [Accessed 3 March 2025].

[2] F. Lainsamputty and M. Wuisang, "Hubungan antara Kecemasan dan Karakteristik Pasien di Instalasi Gawat Darurat Rumah Sakit di Sulawesi Tengah," Journal of Islamic Medicine, vol. VI, no. 01, pp. 28-38, 2022, doi: https://doi.org/10.18860/jim.v6i1.15331.

[3] L. Fan, "Effect of Brand Image, Service Quality and Facilities on Hospitalization Patients' Satisfaction at Rs Royal Prima Medan," International Journal of Research in Engineering and Science (IJRES), vol. XI, no. 3, pp. 346-350, 2023, doi: http://dx.doi.org/10.52155/ijpsat.v34.1.4529.

[4] H. Yu , X. Mujun, L. Xin, Z. Ping, F. Limei, L. Aming and L. Xinquan, "Short Sleep Duration as a Risk Factor for Depression, Anxiety and Fatigue in Patients with Leukemia," Neuropsychiatric Disease and Treatment , no. 18, p. 1573–1582 , 2022, doi: https://doi.org/10.2147/ndt.s362229

[5] A. P. d. M. Simplício, P. C. d. A. Viola, L. C. d. C. Lavôr and e. a. , "Unhealthy Dietary Pattern Associated with Common Mental Disorders in Adults and Older Adults: A Population-based Study," Bentham Science Publishers, vol. XX, pp. 1155-1164, 2024, doi: https://doi.org/10.2174/1573401319666230503155748

[6] T. Kechine, T. Ali, T. Worku, L. Abdisa and T. A. Yadeta, "Anxiety and Associated Factors Among Clients on Highly Active Antiretroviral Therapy (HAART) in Public Hospitals of Southern Ethiopia: A Multi-Center Cross-Sectional Study," Psychology Research and Behavior Management, vol. XV, p. 3889–3900, 2022, doi: https://doi.org/10.2147/prbm.s385630.

[7] T. Anjarsari, I. R. I. Astutik and U. Indahyanti, "Deteksi Dini Gangguan Kecemasan Menggunakan Metode Naïve Bayes," JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika, vol. VII, no. 04, pp. 1198 - 1210, 2022, doi: https://doi.org/10.29100/jipi.v7i4.3197.

[8] M. Andrian, N. Aristiya, S. Rahayu and A. Novita, "Efektifitas Pemberian Terapi Slow Deep Breathing dan Terapi Dzikir Terhadap Penurunan Tingkat Kecemasan Pada Pasien di IGD," Jurnal Ners, vol. VIII, no. 02, pp. 1042 - 1048, 2024.

[9] D. M. Abdullah and A. M. Abdulazeez, "Machine Learning Applications based on SVM Classification: A Review," Qubahan Academic Journal , vol. I, no. 02, pp. 81-90, 2021, doi: https://doi.org/10.48161/qaj.v1n2a50.

[10] N. W. Utami and N. N. Saptiari, "Penerapan Data Mining Untuk Klasifikasi Penyebab Kematian Menggunakan Algoritma Support Vector Machine," Jurnal Ilmiah Ilmu Terapan Universitas Jambi, vol. IV, no. 02, pp. 234-240, 2020, doi: https://doi.org/10.22437/jiituj.v4i2.13268.

[11] M. Furqan, R. Kurniawan and K. I. HP, "Evaluasi Performa Support Vector Machine Classifier Terhadap Penyakit Mental," Jurnal Sistem Informasi Bisnis , vol. X, no. 02, pp. 203-210, 2020, doi: https://doi.org/10.21456/vol10iss2pp203-210.

[12] R. d. Filippis and A. A. Foysal, "Comprehensive analysis of stress factors afecting students: a machine learning approach," Discover Artifcial Intelligence, 2024, doi: https://doi.org/10.1007/s44163-024-00169-6.

[13] C. Chakraborty, M. Bhattacharya, S. Pal and S.-S. Lee, "From machine learning to deep learning: Advances of the recent data-driven paradigm shift in medicine and healthcare," Elsevier B.V, 2023, doi: https://doi.org/10.1016/j.crbiot.2023.100164.

[14] R. R. Pérez and J. Bajorath, "Evolution of Support Vector Machine and Regression Modeling in Chemoinformatics and Drug Discovery," Journal of Computer-Aided Molecular Design, p. 355–362, 2022, doi: https://doi.org/10.1007/s10822-022-00442-9.

[15] M. Yamamoto, E. Bagarinao, I. Kushima and e. a. , "Support vector machine-based classification of schizophrenia patients and healthy controls using structural magnetic resonance imaging from two independent sites," PLoS ONE, 2020, doi: https://doi.org/10.1371/journal.pone.0239615.

[16] S. Rabbani, D. Safitri, N. Rahmadhani and e. a. , "Comparative Evaluation of SVM Kernels for Sentiment lassification in Fuel Price Increase Analysis," MALCOM: Indonesian Journal of Machine Learning and Computer Science, vol. III, no. 2, pp. 153-160, 2023, doi: https://doi.org/10.57152/malcom.v3i2.897.

[17] R. Wardoyo, I. M. A. Wirawan and I. G. A. Pradipta, "Oversampling Approach Using Radius-SMOTE for Imbalance Electroencephalography Datasets," Emerging Science Journal, pp. 382-398, 2022, doi: https://doi.org/10.28991/ESJ-2022-06-02-013.

[18] H. Hairani, T. Widiyaningtyas and D. D. Prasetya, "Addressing Class Imbalance of Health Data: a Systematic Literature Review on Modified Synthetic Minority Oversampling Technique (SMOTE) Strategies," International Journal On Informatics Visualization, pp. 1310-1318, 2024, doi: https://dx.doi.org/10.62527/joiv.8.3.2283.

[19] R. Sujitha and B. Paramasivan, "Optimal progressive classification study using SMOTE-SVM for stages of lung disease," Automatika, pp. 807-814, 2023, doi: https://doi.org/10.1080/00051144.2023.2218167.

[20] S. Sreenivasa, "Radial Basis Function (RBF) Kernel: The Go-To Kernel," Towards Data Science, 12 October 2020, Available: https://towardsdatascience.com/.

[21] GeeksforGeeks (Tim Editorial), "Radial Basis Function Kernel – Machine Learning," GeeksforGeeks, 06 May 2024, Available: https://www.geeksforgeeks.org/.