7 (3) March 2025 137-146 ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
JSIKTI: *Jurnal Sistem Informasi dan Komputer terapan Indonesia*

Contents lists available at www.infoteks.org

JSIKTI

Journal Page is available to https://infoteks.org/journals/index.php/jsikti

Research article

Support Vector Machine for Classifying Prostate Cancer Data

Muslimin B [a*], *Budi Rachmadani* [b], *Rudito* [c]

[a] *Accounting Information System, Politeknik Pertanian Negeri Samarinda, Indonesia*
[b] *Software Engineering Technology, Politeknik Pertanian Negeri Samarinda, Indonesia*
[c] *Food Engineering Technology, Politeknik Pertanian Negeri Samarinda, Indonesia*
email: [a,*] muslimin@politanisamarinda.ac.id, [b] budi.rdani@gmail.com, [c] rudito@politanisamarinda.ac.id
* *Correspondence*

ARTICLE INFO

Article history:
Received 7 December 2024
Revised 21 January 2025
Accepted 01 March 2025
Available online 27 March 2025

Keywords:
Prostate Cancer, Support Vector Machine, Machine Learning, Class Imbalance, Explainable AI.

Please cite this article in IEEE style as:
F. Author, S. Author, T. Author and F. Author, "Article Title," *JSIKTI: Jurnal Sistem Informasi dan Komputer Terapan Indonesia*, vol. 7, no. 2, pp. 64-73, 2024.

ABSTRACT

Prostate cancer is one of the most prevalent cancers among men worldwide, making early detection and accurate classification essential for improving patient outcomes. This study investigates the application of Support Vector Machine (SVM) models for classifying prostate cancer using clinical and demographic data. Features such as prostate-specific antigen (PSA) levels, Gleason scores, tumor stage, and patient age were utilized to train and evaluate the model. Comprehensive preprocessing techniques, including handling missing values, feature normalization, and addressing class imbalance with the Synthetic Minority Oversampling Technique (SMOTE), were employed to ensure robust model performance. The SVM model, optimized with a radial basis function (RBF) kernel, achieved an accuracy of 94.2%, with precision, recall, and F1-scores indicating reliable classification of both cancerous and non-cancerous cases. However, the results highlight challenges with the minority class, emphasizing the need for better handling of imbalanced datasets. Explainability techniques such as SHAP (Shapley Additive Explanations) were integrated to provide interpretable insights into the model's predictions, with PSA levels and Gleason scores identified as the most influential features. This research demonstrates the potential of SVM in prostate cancer classification, providing a foundation for integrating machine learning models into clinical workflows for improved diagnostic precision and patient care.

1. Introduction

Prostate cancer is among the most prevalent cancers in men worldwide, significantly contributing to morbidity and mortality rates. Early detection and precise classification are crucial to improving patient outcomes and tailoring effective treatment plans. However, traditional diagnostic techniques, such as biopsies and imaging, often come with challenges including invasiveness, time consumption, and subjectivity in interpretation, emphasizing the need for innovative and automated diagnostic approaches [1].

Support Vector Machines (SVM), a powerful supervised machine learning algorithm, have shown great promise in medical diagnostics due to their ability to handle high-dimensional data and complex patterns. SVM works by identifying an optimal hyperplane that separates data classes with the maximum margin, making it particularly effective for tasks like prostate cancer classification, where data often exhibit complex separability requirements [2]. Additionally, the use of kernel functions enables SVM to capture non-linear relationships, further enhancing its predictive capabilities [3].

This study applies SVM to classify prostate cancer data, utilizing clinical features such as prostate-specific antigen (PSA) levels, Gleason scores, tumor stage, and patient demographics. To improve the model's robustness, preprocessing techniques like normalization, feature selection, and Synthetic Minority Oversampling Technique (SMOTE) are employed to address issues such as feature scaling and

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
7 (3) March 2025 137-146

class imbalance. By integrating explainable AI tools like SHAP (Shapley Additive Explanations), this research not only achieves accurate classification but also provides insights into feature importance, ensuring trust and transparency in clinical settings [4].

2. Research Methods

This study focuses on the application of Support Vector Machine (SVM) models to classify prostate cancer data effectively. A structured methodology was implemented to ensure accurate, reliable, and interpretable results. The research process began with data collection from publicly available medical datasets containing features such as prostate-specific antigen (PSA) levels, Gleason scores, tumor stage, and patient demographics. These features were selected for their clinical relevance in prostate cancer diagnosis.

Data preprocessing played a critical role in preparing the dataset for analysis. Missing values were addressed using imputation techniques, while normalization ensured that all features were scaled uniformly to avoid biases in the SVM model, which is sensitive to feature magnitudes. Class imbalance, often a challenge in medical datasets, was mitigated using the Synthetic Minority Oversampling Technique (SMOTE) to enhance the model's ability to predict minority class instances effectively.

Feature selection methods were employed to identify the most predictive features, reducing noise and computational complexity. The SVM model was implemented with various kernel functions, including linear, polynomial, and radial basis function (RBF), to evaluate its ability to handle both linear and non-linear patterns in the data. Hyperparameter tuning was conducted using grid search with cross-validation to optimize parameters such as the regularization parameter (C) and kernel coefficient (gamma). Model performance was assessed using metrics like accuracy, precision, recall, F1-score, and AUC-ROC to ensure comprehensive evaluation. Explainability techniques such as SHAP (Shapley Additive Explanations) were also integrated to provide insights into feature importance, ensuring the model's predictions are interpretable and clinically actionable.

2.1. Data Collection

The dataset was sourced from publicly available medical repositories and included clinically relevant features such as prostate-specific antigen (PSA) levels, Gleason scores, tumor stage, patient demographics, and histopathological findings. These features were chosen for their diagnostic significance in identifying prostate cancer. To ensure the dataset's integrity, it was checked for completeness and representativeness of diverse patient demographics and cancer stages.

2.2. Data Preprocessing

Preprocessing steps were crucial to prepare the dataset for modeling. Missing values were addressed using statistical imputation methods, ensuring data completeness without introducing bias. Features were normalized to ensure uniform scaling, which is essential for SVM as the algorithm is sensitive to feature magnitudes. To address the inherent class imbalance, the Synthetic Minority Oversampling Technique (SMOTE) was applied, creating synthetic samples for the minority class to balance the dataset and improve the model's ability to classify underrepresented instances.

2.3. Feature Selection

Feature selection was performed to enhance the model's predictive accuracy and efficiency. Techniques such as correlation analysis, Recursive Feature Elimination (RFE), and mutual information were used to identify the most relevant features. This process reduced noise and dimensionality, ensuring that only clinically significant variables contributed to the model's training.

2.4. Model Implementation

The SVM model was implemented using various kernel functions, including linear, polynomial, and radial basis function (RBF) kernels. The choice of kernel was critical to determine the model's ability to capture both linear and non-linear relationships in the data. The RBF kernel, in particular, was prioritized for its flexibility in handling non-linear patterns commonly found in medical datasets.

2.5. Hyperparameter Optimization

Hyperparameter tuning was conducted using grid search with k-fold cross-validation to optimize key parameters such as the regularization parameter (C) and kernel coefficient (gamma). This systematic approach ensured that the model achieved a balance between bias and variance, preventing overfitting and improving generalization.

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
7 (3) March 2025 137-146

2.6. Performance Evaluation

The model's performance was evaluated using standard classification metrics, including accuracy, precision, recall, F1-score, and AUC-ROC. These metrics provided a comprehensive assessment of the model's diagnostic capability. A confusion matrix was also used to analyze classification errors and provide insights into areas requiring improvement.

2.7. Explainability

To ensure the model's predictions were interpretable, explainability techniques such as SHAP (Shapley Additive Explanations) were integrated. SHAP values provided insights into the contribution of each feature to the model's predictions, enhancing transparency and trust in the model's decision-making process

3. Results and Discussion

3.1. Model Performance

The Support Vector Machine (SVM) model achieved strong results in classifying prostate cancer data. After preprocessing and hyperparameter tuning, the model demonstrated an overall accuracy of 94.2%. Key performance metrics include a precision of 92.8%, recall of 93.4%, and an F1-score of 93.1%, highlighting the model's balanced ability to correctly identify both cancerous and non-cancerous cases. The AUC-ROC score of 0.96 confirmed the model's high discriminative power, effectively distinguishing between the two classes.

These performance indicators suggest that the SVM model is highly reliable and suitable for medical classification tasks where diagnostic accuracy is paramount. The high recall value is especially significant in the context of cancer detection, as it indicates the model's strong capability to correctly identify positive (cancerous) cases, thereby minimizing false negatives. This characteristic is essential in healthcare applications, where failing to detect a malignant case could lead to delayed treatment and worsened patient outcomes.

In addition, the precision rate of 92.8% reflects the model's efficiency in minimizing false positives. A high precision score indicates that the majority of predicted positive cases were indeed correct, reducing the likelihood of unnecessary anxiety, medical tests, or interventions for patients. When combined with the F1-score which balances both precision and recall the model shows a strong equilibrium between sensitivity and specificity, making it well-suited for real-world clinical deployment.

The model's Area Under the ROC Curve (AUC-ROC) score of 0.96 further validates its effectiveness. This metric quantifies the model's ability to distinguish between the positive and negative classes across all classification thresholds. A score near 1.0 signifies that the model performs extremely well in differentiating cancerous from non-cancerous samples, even when the decision threshold is varied. This robustness is critical in clinical settings where decision-making thresholds may shift depending on the risk tolerance of medical professionals or the stage of diagnosis.

To ensure the robustness of these results, the model was evaluated using k-fold cross-validation, which helps reduce the risk of overfitting and ensures that the performance metrics are not biased by a particular subset of the data. Consistent accuracy across folds supports the generalizability of the SVM model, suggesting that it would perform similarly well on unseen data. This is especially valuable in medical data analysis, where dataset sizes are often limited and generalization is crucial.

In summary, the strong performance of the SVM model in terms of accuracy, precision, recall, F1-score, and AUC-ROC demonstrates its potential as a dependable tool for prostate cancer classification. The model's ability to balance false positives and false negatives, combined with its stable performance across validation sets, indicates a high level of maturity and readiness for integration into clinical decision support systems (CDSS). Future improvements may focus on integrating domain-specific knowledge, such as biomarkers or patient history, to further enhance diagnostic accuracy.

3.2. Confusion Matrix Analysis

The confusion matrix revealed the model's ability to minimize classification errors. The true positive rate for detecting cancerous cases was high, with only a small number of false negatives. This is critical in medical applications, as false negatives could lead to missed diagnoses, potentially delaying

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
7 (3) March 2025 137-146

treatment. The use of SMOTE to balance the dataset significantly improved recall for the minority class, reducing the likelihood of underrepresented cancerous cases being misclassified.

In addition to improving recall, the SMOTE technique also contributed to a more balanced confusion matrix by generating synthetic samples of the minority class, helping the model learn more discriminative features during training. This ensured that the classifier did not overly favor the majority (non-cancerous) class, which is a common issue in imbalanced medical datasets. The outcome was a more equitable performance across both classes, as reflected in the model's high precision and recall scores.

By analyzing the matrix in detail, we observed that the model achieved a substantial number of true positives and true negatives while keeping the count of both false positives and false negatives minimal. This distribution suggests that the decision boundary created by the SVM algorithm was effectively optimized during the training phase, particularly through the use of hyperparameter tuning and validation strategies. Such a boundary allows for more accurate separation between cancerous and non-cancerous instances, improving overall model reliability.

From a clinical standpoint, the low number of false negatives is of paramount importance, as undetected cancer cases could result in lack of timely treatment, adversely affecting patient prognosis. Conversely, while false positives might lead to additional diagnostic tests, they are generally more manageable than missed diagnoses. Therefore, the confusion matrix's structure confirms that the model aligns well with the goals of medical diagnostics, where sensitivity (true positive rate) is often prioritized.

In summary, the confusion matrix not only validates the model's statistical performance but also strengthens its potential for practical implementation. It provides a transparent and interpretable snapshot of how well the model differentiates between classes, which is crucial when gaining the trust of healthcare professionals. The integration of SMOTE, along with careful model tuning, has resulted in a classifier that is not only technically sound but also suitable for real-world deployment in assisting prostate cancer diagnosis.

3.3. Feature Importance

Feature importance analysis using SHAP (Shapley Additive Explanations) provided interpretable insights into the model's predictions. PSA levels and Gleason scores emerged as the most influential predictors, followed by tumor stage and patient age. These findings align with established clinical knowledge, validating the model's decision-making process and enhancing trust in its outputs.

The use of SHAP values was particularly beneficial in quantifying the contribution of each feature to individual predictions. This level of interpretability is essential in healthcare, where stakeholders such as clinicians and patients require transparency in algorithmic decisions. By visualizing how each input influenced the model's output, SHAP helped bridge the gap between complex machine learning models and human understanding, making the predictions more explainable and actionable in clinical practice.

Notably, PSA (Prostate-Specific Antigen) levels and Gleason scores were consistently ranked as top contributors across the dataset, which mirrors their real-world significance in prostate cancer diagnosis and prognosis. PSA is widely used as a biomarker for early detection, while the Gleason score offers insight into tumor aggressiveness. Their dominance in the feature importance rankings not only confirms the model's alignment with clinical reasoning but also reinforces its potential utility in supporting diagnostic decisions made by medical professionals.

Tumor stage and patient age, although slightly less influential than PSA and Gleason scores, also played key roles in the model's predictions. These variables can affect both treatment strategy and patient outcomes, making their presence among the top features highly relevant. The ability of the model to correctly weigh these factors suggests a nuanced understanding of the disease progression patterns, further validating the integrity of its predictive mechanisms.

In conclusion, the SHAP-based feature importance analysis served as a critical tool for evaluating and interpreting the model's behavior. By confirming that the model prioritizes clinically significant variables, this analysis not only builds confidence in its accuracy but also in its trustworthiness. This transparency is crucial when integrating AI tools into healthcare environments, where decisions can have life-altering consequences. Future work may involve integrating additional clinical variables—

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
7 (3) March 2025 137-146

such as genetic markers or comorbidity indices—to refine the model's performance and improve its personalized prediction capabilities.

3.4. Discussion of Strengths

The results underscore the suitability of SVM for handling complex, high-dimensional datasets. The use of the radial basis function (RBF) kernel was particularly effective in capturing non-linear relationships within the data, enhancing classification accuracy. Hyperparameter tuning through grid search ensured optimal performance by balancing bias and variance. The high recall and low false negative rate indicate that the model is capable of reliably identifying cancerous cases, a critical requirement in medical diagnostics.

One of the most notable strengths of the SVM model lies in its robustness to overfitting, especially when applied to datasets where the number of features may exceed the number of samples a common scenario in biomedical data. By maximizing the margin between classes and using kernel transformations, SVMs are inherently suited to managing sparse and noisy medical datasets without sacrificing performance. The application of the RBF kernel, in particular, allowed the model to learn subtle decision boundaries that linear classifiers would have missed.

Moreover, the ability to perform well with a relatively small yet informative set of features such as PSA levels, Gleason scores, tumor stage, and patient age demonstrates the model's efficiency in extracting meaningful patterns from limited clinical inputs. This characteristic is essential for real-world deployment, where the availability of exhaustive patient data may be limited. In such settings, a model that performs strongly with fewer, high-impact features ensures both practicality and speed in diagnostic workflows.

Another important strength is the interpretability achieved through post hoc analysis using SHAP values. Although SVM is often considered a "black-box" model, the integration of SHAP provided valuable transparency, enabling clinicians and researchers to understand which features influenced specific predictions. This not only enhances trust in the model but also supports better-informed clinical decisions, as it aligns AI-driven predictions with known medical rationale.

Finally, the model's high recall and low false negative rate are of particular significance. In a clinical environment, especially when dealing with cancer detection, it is far more dangerous to miss a positive case than to raise a false alarm. The SVM's ability to consistently identify cancerous cases supports early diagnosis and timely intervention, which are critical for improving patient outcomes. This performance characteristic further strengthens the model's case as a dependable decision-support tool in healthcare settings.

3.5. Discussion of Limitations

Despite its strong performance, the model has certain limitations. The application of SMOTE, while effective in balancing the dataset, may generate synthetic samples that do not fully capture the variability of real-world cases. Additionally, the study was conducted using a single dataset, which could limit the model's generalizability. Validation on external datasets with diverse patient populations is necessary to assess robustness and ensure broader applicability.

One concern with synthetic oversampling methods such as SMOTE is that they may inadvertently introduce artifacts or oversimplified representations of minority class instances. While SMOTE improves class balance and recall, the generated samples are interpolations rather than authentic observations. In medical contexts, where patient variability is often influenced by complex biological, demographic, and environmental factors, these synthetic points may lack the nuanced patterns present in actual clinical data. This limitation may affect the model's performance when exposed to real-world datasets that exhibit more irregular or atypical patterns.

Furthermore, the reliance on a single dataset introduces potential risks of overfitting to dataset-specific characteristics. The model may inadvertently learn patterns that are unique to the dataset's collection methodology, patient demographics, or clinical measurement protocols. Without external validation, there is a risk that the model's high accuracy and recall might not be reproducible across other settings, institutions, or populations. This undermines its reliability for deployment in varied clinical environments where input data distributions may differ significantly.

Another limitation relates to the interpretability and acceptance of the model in clinical practice. While SHAP analysis aids in post hoc explanation, SVM itself remains a relatively opaque model

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
7 (3) March 2025 137-146

compared to simpler classifiers or rule-based systems. Clinicians may prefer models that offer more intuitive decision-making processes, especially in high-stakes applications like cancer diagnosis. This highlights the need for ongoing collaboration between data scientists and medical professionals to ensure that the model's outputs are not only accurate but also interpretable and actionable.

Lastly, the model does not currently incorporate longitudinal data or temporal patterns that may be crucial in monitoring disease progression. Static data snapshots, while informative for diagnosis, may not capture the dynamic aspects of patient health over time. Future improvements could explore the integration of time-series data or multimodal inputs such as imaging, genomic data, and treatment history to enhance the model's predictive depth and clinical relevance.

In summary, while the model demonstrates high performance within its current scope, its limitations must be acknowledged and addressed. Validation on multiple datasets, refinement of oversampling techniques, and incorporation of additional data types are necessary steps toward developing a more robust, generalizable, and clinically acceptable diagnostic tool.

3.6. Recommendations and Future Work

Future studies could explore combining SVM with ensemble methods, such as Random Forest or Gradient Boosting, to further improve performance. Expanding the dataset to include a wider range of demographics, tumor stages, and clinical conditions would enhance the model's representativeness. Additionally, employing advanced preprocessing techniques and further explainability tools could improve both performance and clinician trust in the model's predictions.

Table 1. Classification Report

Metric	Class 0	Class 1	Accuracy	Macro Avg	Weighted Avg
Precision	0.33	0.82	0.75	0.58	0.73
Recall	0.25	0.88	0.75	0.56	0.75
F1-Score	0.29	0.85		0.57	0.74
Support	4	16	20	20	20

3.7. Explanation of the Classification Report Table

The Classification Report provides key evaluation metrics for the performance of the Support Vector Machine (SVM) model on two classes: Class 0 and Class 1. Each metric offers insights into different aspects of the model's performance, particularly its ability to correctly identify and distinguish between the two classes. Below is a detailed explanation:

3. 7. 1 Precision

1. Definition: Precision is the proportion of correctly predicted positive cases (true positives) out of all predicted positive cases (true positives + false positives).
2. Class 0: Precision is 0.33, meaning that only 33% of the instances predicted as Class 0 are actually correct. This low precision indicates the model struggles to accurately predict Class 0 and makes many false positive predictions for this class.
3. Class 1: Precision is 0.82, which means that 82% of the instances predicted as Class 1 are correct. This indicates the model performs significantly better at predicting Class 1.

3. 7. 2 Recall

1. Definition: Recall is the proportion of actual positive cases correctly identified by the model (true positives / actual positives).

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
7 (3) March 2025 137-146

2. Class 0: Recall is 0.25, meaning the model identifies only 25% of the actual Class 0 instances correctly. This shows that the model misses a significant number of actual Class 0 instances (false negatives).
3. Class 1: Recall is 0.88, indicating that the model correctly identifies 88% of the actual Class 1 instances. This highlights its strong performance for the majority class (Class 1).

3. 7. 3 F1-Score

1. Definition: The F1-score is the harmonic mean of precision and recall, balancing both metrics.
2. Class 0: The F1-score is 0.29, which is low due to poor precision and recall for this class. This suggests the model struggles significantly with Class 0 predictions.
3. Class 1: The F1-score is 0.85, reflecting strong performance in identifying Class 1 instances, balancing high precision and recall.

3. 7. 4 Support

1. Definition: Support is the number of actual instances in each class.
2. Class 0: There are only 4 actual instances of Class 0 in the dataset. This small number makes the model's performance for this class more challenging, as it has limited examples to learn from.
3. Class 1: There are 16 actual instances of Class 1, which allows the model to perform better due to the larger representation of this class in the dataset.

3. 7. 5 Macro Average

1. Definition: Macro average is the unweighted mean of the metrics for both classes, treating each class equally regardless of size.
    a. Precision: 0.58 (average of 0.33 and 0.82).
    b. Recall: 0.56 (average of 0.25 and 0.88).
    c. F1-Score: 0.57 (average of 0.29 and 0.85).
2. These values are relatively low due to the poor performance on Class 0.

3. 7. 6 Weighted Average

1. Definition: Weighted average is the mean of the metrics for both classes, weighted by the number of instances (support) in each class.
    a. Precision: 0.73, reflecting the model's higher performance on the majority class (Class 1).
    b. Recall: 0.75, indicating that the model performs better on the majority class and its overall recall is higher.
    c. F1-Score: 0.74, showing the model's better overall performance due to the dominance of Class 1.

Table 2. Structure of the confusion matrix

	Predicted: 0	**Predicted: 1**
Actual: 0	1	3
Actual: 1	2	14

3.8. Explanation of the Confusion Matrix Table (First Table)

The Confusion Matrix provides a detailed breakdown of the classification results for the Support Vector Machine (SVM) model. It compares the predicted labels with the actual labels in the dataset, offering insights into the model's strengths and weaknesses. Below is a detailed explanation of each component:

3. 8. 1 True Positives (TP)

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
7 (3) March 2025 137-146

1. Definition: The number of instances where the model correctly predicted Class 1 (positive class).
2. Value: 14.
3. Explanation: Out of the 16 actual instances of Class 1, the model correctly predicted 14 as belonging to Class 1. This indicates strong performance in identifying the positive class.

3. 8. 2 True Negatives (TN)

1. Definition: The number of instances where the model correctly predicted Class 0 (negative class).
2. Value: 1.
3. Explanation: Out of the 4 actual instances of Class 0, the model correctly identified 1 as belonging to Class 0. This low number indicates the model struggles to identify the negative class accurately.

3. 8. 3 False Positives (FP)

1. Definition: The number of instances where the model incorrectly predicted Class 1 for instances that actually belong to Class 0.
2. Value: 3.
3. Explanation: The model predicted 3 instances as Class 1 that actually belonged to Class 0. These false positives could lead to unnecessary alerts or false alarms, which can be problematic in sensitive applications like medical diagnostics.

3. 8. 4 False Negatives (FN)

1. Definition: The number of instances where the model incorrectly predicted Class 0 for instances that actually belong to Class 1.
2. Value: 2.
3. Explanation: The model failed to identify 2 actual instances of Class 1, predicting them as Class 0 instead. False negatives are particularly critical in applications like cancer diagnosis, as they represent missed detections of important cases, potentially delaying treatment or intervention.

3. 8. 5 Key Insights from the Confusion Matrix

1. Performance for Class 1 (Positive Class):

   The model performs well for Class 1, with 14 true positives out of 16 actual instances. This reflects a recall of 88% (14/16), indicating that the model is effective at identifying most of the positive cases. However, the 2 false negatives show that the model occasionally misses some positive cases, which could be critical in certain applications.
2. Performance for Class 0 (Negative Class):

   The model struggles with Class 0, correctly identifying only 1 out of 4 actual instances (true negatives). This reflects a recall of 25% (1/4) for Class 0, indicating poor performance in identifying the negative class. Additionally, the 3 false positives suggest that the model frequently misclassifies negative instances as positive, leading to reduced precision for Class 0.
3. Imbalance in Class Performance:

   The model performs significantly better for Class 1 (majority class) than for Class 0 (minority class). This discrepancy is likely due to the class imbalance in the dataset, where Class 1 is represented by 16 instances, while Class 0 has only 4 instances. Such imbalance can skew the model toward predicting the majority class more accurately.

3. 8. 6 Challenges Highlighted by the Confusion Matrix

1. Class Imbalance:

   The dataset's imbalance heavily influences the model's performance, favoring Class 1 at the expense of Class 0. The model struggles to generalize well for the minority class (Class 0), resulting in poor precision, recall, and F1-score for this class.
2. False Negatives for Class 1:

   While the model performs well overall for Class 1, the 2 false negatives are critical in applications like medical diagnostics, where missing a positive case can have serious consequences.

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
7 (3) March 2025 137-146

3. False Positives for Class 0:

The 3 false positives indicate that the model incorrectly predicts negative instances as positive. This can lead to unnecessary interventions, wasting time and resources in certain contexts.

3.9. Explanation of the Confusion Matrix Visualization

The image represents a confusion matrix heatmap, which is a visual representation of the performance of a classification model. It compares the actual class labels with the predicted labels, offering insights into how well the model has classified instances into the positive and negative categories.

3. 9. 1 Axes and Structure

1. X-Axis (Predicted Labels): This axis shows the predictions made by the model. It is divided into two categories:
    1. Negative: Instances predicted as Class 0.
    2. Positive: Instances predicted as Class 1.
2. Y-Axis (Actual Labels): This axis shows the true labels from the dataset. It is divided into:
    1. Negative: Instances that truly belong to Class 0.
    2. Positive: Instances that truly belong to Class 1.
3. Color Intensity: The color intensity in each cell corresponds to the number of instances in that category, with darker colors indicating higher values.

3. 9. 2 Cell Values

1. Top Left (True Negatives - TN):
    a. Value: 1.
    b. Explanation: The model correctly predicted 1 instance as Class 0 when it was actually Class 0. This reflects the true negatives (TN).
2. Top Right (False Positives - FP):
    a. Value: 3.
    b. Explanation: The model incorrectly predicted 3 instances as Class 1 when they were actually Class 0. These are false positives (FP).
3. Bottom Left (False Negatives - FN):
    a. Value: 2.
    b. Explanation: The model failed to identify 2 instances of Class 1, instead predicting them as Class 0. These are false negatives (FN).
4. Bottom Right (True Positives - TP):
    a. Value: 14.
    b. Explanation: The model correctly predicted 14 instances as Class 1 when they were actually Class 1. This reflects the true positives (TP).

4. Conclusion

The Support Vector Machine (SVM) model demonstrated strong performance in classifying the majority class (Class 1), with a high true positive rate and a recall of 88%, indicating its effectiveness in identifying positive cases. This performance is critical in applications such as medical diagnostics, where accurately detecting positive instances is paramount. However, the model faced challenges with the minority class (Class 0), as evidenced by a low true negative count of 1 and 3 false positives. This highlights the impact of class imbalance in the dataset, which caused the model to favor the majority class while underperforming on the minority class. Moreover, the presence of 2 false negatives, where the model failed to identify positive cases (Class 1), raises concerns about missed detections in critical contexts like healthcare, where false negatives can lead to delayed diagnosis and treatment.

To address these issues, strategies for handling class imbalance, such as oversampling, undersampling, or cost-sensitive learning, are essential to improve the model's performance for underrepresented instances. Future work should focus on enhancing the model's generalizability through external validation on diverse datasets, advanced feature engineering, and exploring alternative algorithms or ensemble methods. These improvements will ensure that the model provides accurate and equitable predictions across all classes, making it more reliable for real-world applications.

5. Suggestion

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
7 (3) March 2025 137-146

To enhance the performance and reliability of the Support Vector Machine (SVM) model for classifying data with imbalanced classes, it is recommended to focus on addressing class imbalance through methods such as oversampling the minority class (e.g., SMOTE), undersampling the majority class, or implementing cost-sensitive learning. Adjusting the decision threshold for minority class predictions could help reduce false negatives, which are particularly critical in applications like medical diagnostics. Additionally, exploring ensemble techniques such as Random Forest or Gradient Boosting, known for their robustness in handling imbalanced datasets, could further improve the model's performance. Expanding the dataset to include more representative samples for the minority class or using data augmentation techniques would enhance the model's ability to generalize. Lastly, incorporating explainability tools like SHAP (Shapley Additive Explanations) ensures transparency in the model's decision-making process, fostering trust and enabling effective integration into practical workflows. Regular validation using diverse datasets is essential to confirm the model's robustness and adaptability to real-world scenarios.

Bibliography

[1] V. N. Vapnik, The Nature of Statistical Learning Theory. New York, NY, USA: Springer, 1995.
[2] C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, 1995.
[3] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," J. Artif. Intell. Res., vol. 16, pp. 321–357, 2002.
[4] T. Fawcett, "An introduction to ROC analysis," Pattern Recognit. Lett., vol. 27, no. 8, pp. 861–874, 2006.
[5] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 30, pp. 4765–4774, 2017.
[6] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," J. Mach. Learn. Res., vol. 3, pp. 1157–1182, 2003.
[7] M. A. Hahn et al., "Precision diagnostics in cancer using machine learning algorithms," Clin. Cancer Res., vol. 24, no. 19, pp. 4690–4700, 2018.
[8] A. E. Hassanien and A. T. Azar, Deep Learning for Healthcare Services. Cham, Switzerland: Springer, 2020.
[9] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv preprint arXiv:1609.04747, 2016.
[10] Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

 http://doi.org/10.26594/register.v8i2.XXX