Journal of Computer Engineering, Electronics and Information Technology (COELITE)

Journal homepage: https://ejournal.upi.edu/index.php/COELITE

Hybrid Explainable AI Approach for DNA Sequence Classification: Feature Importance, Permutation Importance, and Local Explanations with LIME

Nur Alamsyah[1], Budiman[2], Reni Nursyanti[3], Elia Setiana[4], Venia Restreva Danestiara[5], Titan Parama Yoga[6]*

[1]Department of Technology and Informatics, Universitas Informatika Dan Bisnis Indonesia, Indonesia
Correspondence E-mail: nuralamsyah@unibi.ac.id

ABSTRACT

Understanding the contribution of features in DNA sequence classification is crucial for enhancing model interpretability and reliability. This study proposes a Hybrid Explainable AI (XAI) approach that integrates Feature Importance (FI), Permutation Importance (PI), and Local Interpretable Model-Agnostic Explanations (LIME) to analyse the most influential features in a Random Forest classifier. FI is utilized to determine the most significant features contributing to the model, while PI validates their impact by assessing performance changes when features are shuffled. Additionally, LIME provides local explanations, offering insight into how specific feature values affect classification decisions. Experimental results on a publicly available DNA sequence dataset reveal a strong correlation between FI and PI rankings, validating the stability of key features such as A84, A89, and A92. LIME further enhances interpretability by highlighting individual instance contributions, reinforcing the relevance of specific nucleotide positions in sequence classification. This hybrid approach provides a more comprehensive understanding of feature importance, improving trust and transparency in DNA sequence classification models.

ARTICLE INFO

Article History:
Submitted/Received 11 Mar 2025
First Revised 18 Mar 2025
Accepted 28 Mar 2025
First Available online 1 Apr 2025
Publication Date 1 Apr 2025

Keyword:
Explainable AI,
Feature Importance,
Permutation Importance,
LIME,
DNA Classification.

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149

1. INTRODUCTION

The classification of DNA sequences plays a crucial role in bioinformatics, particularly in identifying genetic variations, disease-associated mutations, and evolutionary relationships [1]. With the growing availability of genomic data, machine learning techniques have been increasingly applied to improve classification accuracy [2]. However, while machine learning models such as Random Forest (RF), Support Vector Machines (SVM), and Deep Learning architectures have shown remarkable performance, their black-box nature limits interpretability, making it difficult for researchers to understand the factors influencing predictions [3]. This lack of transparency poses challenges in critical applications such as genetic disease prediction and mutation impact analysis, where interpretability is essential for validating model decisions [4].

To address this issue, Explainable Artificial Intelligence (XAI) has emerged as a framework to interpret complex machine learning models [5]. Several XAI techniques have been introduced to provide insights into model behaviour, including Feature Importance (FI), Permutation Importance (PI), and Local Interpretable Model-Agnostic Explanations (LIME) [6]. FI is a widely used approach that measures the contribution of each feature based on decision trees' splits, providing a global explanation of feature relevance [7]. However, FI may be biased toward correlated features, necessitating Permutation Importance (PI) as a validation method. PI quantifies feature impact by measuring performance degradation when a feature's values are randomly shuffled, offering a more robust validation mechanism for identifying key features [8]. While both FI and PI provide global feature explanations, they do not provide instance-level interpretability.

To complement these global explanations, LIME (Local Interpretable Model-Agnostic Explanations) has been introduced to provide localized feature insights by perturbing input data and fitting a surrogate model to approximate local decision boundaries [9]. LIME has been effectively applied in biomedical research to understand mutation impact, disease classification, and sequence pattern contributions [10]. However, its application in DNA sequence classification remains limited, and no previous studies have integrated FI, PI, and LIME into a hybrid explainability framework for genetic data analysis.

Several studies have focused on applying individual explainability techniques to machine learning models for genomic analysis. Zhou *et al.* [11] and Zubair *et al.* [12] explored FI in genomic classification, demonstrating its effectiveness in ranking influential features. Cantor *et al.* [13] highlighted PI's ability to validate feature significance, making it more reliable for high-dimensional biological datasets. Meanwhile, Labory *et al.* [14] introduced LIME as a method for explaining predictions at an instance level, which was later adapted in genomic research for mutation-based classification Puiu *et al.*, [15] and genomic sequence interpretation Lee *et al.*[16]. However, the existing literature lacks a comprehensive hybrid approach that integrates FI, PI, and LIME into a unified framework.

This study proposes a Hybrid Explainable AI (XAI) approach for DNA sequence classification by integrating Feature Importance (FI), Permutation Importance (PI), and LIME. The proposed framework aims to:

1. Determine the most influential genomic features using FI,
2. Validate feature significance through PI, and
3. Provide local instance-based explanations with LIME.

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149

By combining these three explainability techniques, this research enhances interpretability, robustness, and transparency in machine learning-based DNA classification, enabling a more reliable understanding of feature contributions and their impact on model predictions. The results of this hybrid approach will not only improve model trustworthiness but also contribute to advancing interpretable AI solutions in bioinformatics and genomic research.

2. METHODS

This section describes the methodology used in this study, which follows a structured pipeline consisting of data collection, preprocessing, model training, and explainability analysis. The workflow begins with data collection, where a DNA sequence dataset is obtained and prepared for machine learning. The data preprocessing phase includes feature-target separation, normalization, and data splitting to ensure optimal model performance. The processed data is then utilized in the model training phase, where a machine learning classifier is trained to predict DNA sequence classifications.

Following model training, the explainability analysis is performed using a hybrid Explainable AI (XAI) framework comprising Feature Importance (FI), Permutation Importance (PI), and Local Interpretable Model-Agnostic Explanations (LIME). FI identifies globally important features by assessing their impact on model decisions, while PI validates the robustness of these feature contributions by measuring performance degradation when features are permuted. Lastly, LIME provides instance-level insights, explaining how individual features contribute to specific predictions.

The methodology is visualized in **Figure 1**, which outlines the main steps involved in this research. The subsequent subsections provide a detailed description of each stage, starting with data collection.

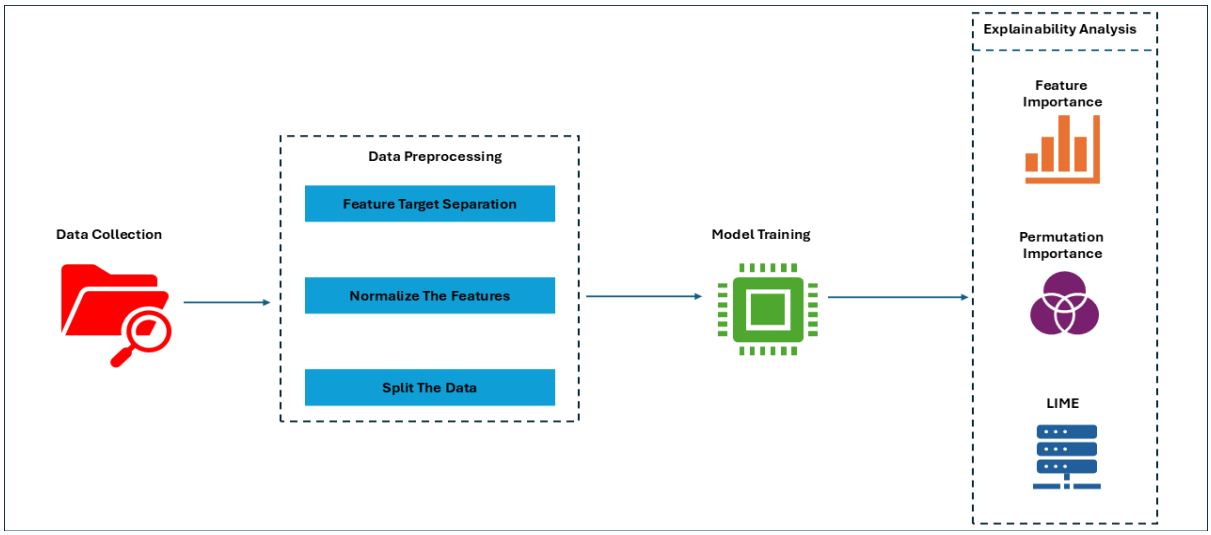

Figure 1. Propose Method.

2.1. Data Collection

The dataset used in this study was obtained from Kaggle, containing 3,187 DNA sequences. Each instance in the dataset represents a DNA segment, with features encoding nucleotide

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149

presence in specific positions. The dataset consists of 180 feature columns representing nucleotide encoding and one target column (class), which categorizes DNA sequences into distinct classes.

To facilitate the explainability analysis, we process the dataset by encoding the nucleotide sequences into numerical values. Each feature represents a specific nucleotide position in the sequence, encoded as a binary value. The target column (class) denotes the classification of the DNA sequence into different biological categories. A summary of the dataset features is provided in **Table 1.**

Table 1. Feature Description of the DNA Dataset

Feature	Description
A0 - A179	Encoded nucleotide presence (binary)
class	DNA sequence classification label

This dataset serves as the foundation for training a classification model and conducting explainability analysis using Feature Importance (FI), Permutation Importance (PI), and Local Interpretable Model-agnostic Explanations (LIME).

2.2. Data Preprocessing

Data preprocessing is a crucial step to ensure that the dataset is properly formatted and optimized for training machine learning models [17]. In this study, the preprocessing phase consists of three main steps: Feature Target Separation, Feature Normalization, and Data Splitting. The first step, Feature Target Separation, involves isolating the features from the target variable. All DNA sequence attributes serve as input features, while the classification label represents the target variable. This separation ensures a clear distinction between input and output data for the machine learning model.

Next, Feature Normalization is performed to standardize the scale of the features. Min-Max Scaling is applied to transform all feature values into a normalized range between 0 and 1, ensuring consistent feature representation. This normalization process enhances model stability and improves training convergence. The final step, Data Splitting, partitions the dataset into three subsets: Training Set, Validation Set, and Test Set. The training set is used to train the model, the validation set is used for hyperparameter tuning, and the test set evaluates model performance on unseen data. The data is split using a 70% training, 15% validation, and 15% testing ratio while maintaining class distribution using stratified sampling. Through this preprocessing pipeline, the dataset is structured and optimized for training, ensuring better model generalization and improved classification performance.

2.3. Model Training

The model training phase involves developing a classification model to predict DNA sequence categories based on extracted features. In this study, a Random Forest Classifier was chosen due to its robustness and capability to handle high-dimensional data effectively [18]. The training process follows these steps:

1. Model Initialization: A Random Forest classifier is initialized with optimized hyperparameters.
2. Model Training: The classifier is trained using the processed dataset.
3. Performance Evaluation: The trained model is assessed using standard classification metrics.

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149

A Random Forest Classifier consists of multiple decision trees, where each tree contributes to the final prediction through majority voting. The model's prediction function can be mathematically expressed as:

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} \quad f_i(X) \tag{1}$$

Where $(\hat{y})$ Is the final predicted class? (N) Is the number of decision trees and $(f_i(X))$ Is the prediction from the (i)-th decision tree. The training process aims to minimize the classification error by optimizing the splitting criteria in each tree, typically using the Gini impurity:

$$G = 1 - \sum_{i=1}^{C} \quad p_i^2 \tag{2}$$

Where (G) Is the Gini impurity, (C) Is the number of classes and (p_i) Is the probability of a sample belonging to a class(i).

2.4. Explainability Analysis

To enhance the interpretability of the classification model, we employ a comprehensive Explainability Analysis framework that integrates Feature Importance (FI), Permutation Importance (PI), and Local Interpretable Model-Agnostic Explanations (LIME). These techniques provide a deeper understanding of how features contribute to the model's decision-making process [19].

Feature Importance (FI)

Feature Importance quantifies the contribution of each feature to the model's predictive performance by measuring the reduction in impurity across decision trees [20]. The importance of a feature (f_i) A decision tree model is given by:

$$FI(f_i) = \sum_{t \in T} \quad I_t \cdot w_t \tag{3}$$

Where (T) represents all nodes in the decision trees where (f_i) is used, (I_t) Is the impurity reduction at the node (t) and (w_t) Is the weighted number of samples that reach the node (t).

Permutation Importance (PI)

Permutation Importance is a model-agnostic technique that validates feature relevance by randomly shuffling individual features and measuring the corresponding decrease in model accuracy. The PI score for a feature (f_i) Is computed as:

$$PI(f_i) = \frac{1}{N}\sum_{j=1}^{N} \quad \left[A - A_{f_i}^{(j)}\right] \tag{4}$$

Where (A) Is the baseline accuracy of the trained model? $(A_{f_i}^{(j)})$Is the accuracy after randomly permuting the feature? (f_i) and (N) Is the number of iterations. Higher PI scores indicate a greater impact on classification performance when the feature is shuffled, thus validating its importance.

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149

Local Interpretable Model-Agnostic Explanations (LIME)

LIME generates local surrogate models to approximate the behaviour of the classifier for individual predictions. It perturbs the input data and learns a linear approximation around a given instance.(x). The explanation model is given by:

$$\hat{f}(x) = arg\ arg\ L\ (f, g, \pi_x) + \Omega(g) \quad (5)$$

Where (g) is the interpretable model (e.g., a linear regression model), $(L(f, g, \pi_x))$Is the loss function ensuring approximations (f) locally, (π_x)represents the proximity measure defining locality and $(\Omega(g))$enforces complexity constraints on (g) For better interpretability.

By integrating FI, PI, and LIME, this explainability framework provides both global feature significance (FI, PI) and local interpretability (LIME), offering a transparent and accountable explanation for DNA sequence classification.

3. RESULTS AND DISCUSSION

3.1. Result

Feature Importance (FI) was employed to determine the most influential features in predicting DNA sequence classifications. The model utilized a Random Forest classifier to rank features based on their contribution to the classification task. **Table 2** presents the top 15 most important features, while **Figure 2** visualizes their relative importance scores. The results indicate that feature A89 holds the highest importance score of 0.091248, followed by A84 and A92, with scores of 0.080648 and 0.064267, respectively. Features A104 and A99 also show notable influence, suggesting their relevance in distinguishing DNA sequence classifications. The remaining features exhibit progressively lower importance scores but still contribute to the model's performance.

These findings provide a global understanding of how different features impact the classification model. However, to validate these results and ensure their robustness, we extend our analysis by incorporating Permutation Importance (PI) in the subsequent section.

Table 2. Top 15 most important features.

Rank	Feature	Importance Score
1	A89	0.091248
2	A84	0.080648
3	A92	0.064267
4	A104	0.061665
5	A99	0.037118
6	A82	0.033560
7	A88	0.016974
8	A83	0.016850
9	A93	0.015391
10	A87	0.015057
11	A95	0.014192
12	A81	0.013620
13	A85	0.013004
14	A96	0.011621

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149

15	A86	0.010789

Figure 2 illustrates the Feature Importance distribution, highlighting the relative significance of each feature in the model.

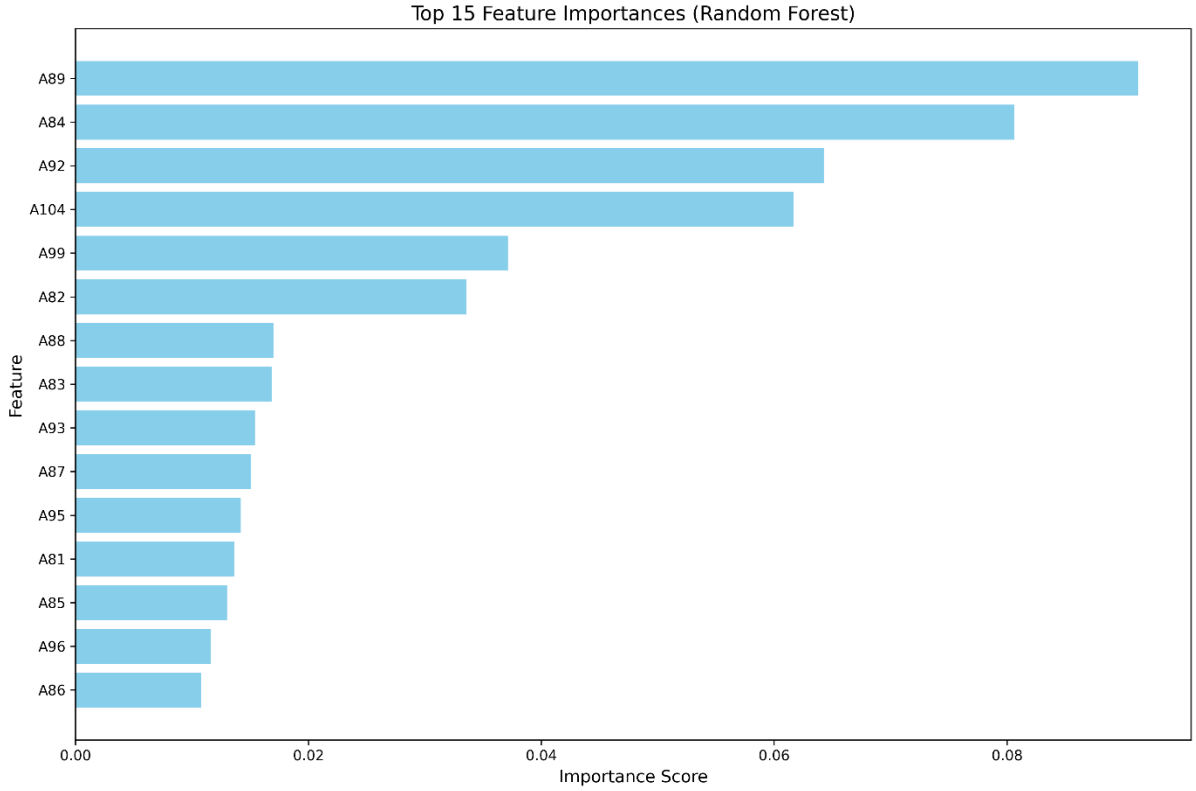

Figure 2. Importance Score

To validate the results obtained from Feature Importance (FI), we employed Permutation Importance (PI) as an additional validation method. PI measures the impact of each feature by shuffling its values and observing the drop in model performance. This method provides a robust verification of the relative importance of features. The results presented in **Table 3** and visualized in Figure 3 confirm the consistency between FI and PI rankings. The top three features remain unchanged, with A84, A89, and A92 maintaining their high importance scores. However, some variations can be observed in lower-ranked features, where A104 and A93 show slightly different rankings. The standard deviation values indicate the stability of feature contributions, with lower deviations implying more consistent impact.

Table 3. Top 15 Permutation Importance Scores.

Rank	Feature	Importance Score	Standard Deviation
1	A84	0.098536	0.009516
2	A89	0.092050	0.013065
3	A92	0.068410	0.007942
4	A104	0.048536	0.009013
5	A93	0.017573	0.004879
6	A99	0.015900	0.005056
7	A95	0.010460	0.005038

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149

Rank	Feature	Importance Score	Standard Deviation
8	A83	0.008368	0.004184
9	A94	0.007741	0.003110
10	A87	0.005439	0.002838
11	A96	0.004603	0.002440
12	A74	0.004184	0.001871
13	A103	0.003347	0.003268
14	A85	0.003347	0.002510
15	A88	0.003138	0.003771

Overall, the validation through PI strengthens the reliability of the feature importance findings, confirming that the identified key features significantly contribute to the classification of DNA sequences. **Figure 3** depicts the Permutation Importance analysis, highlighting the top 15 most influential features and their variations.

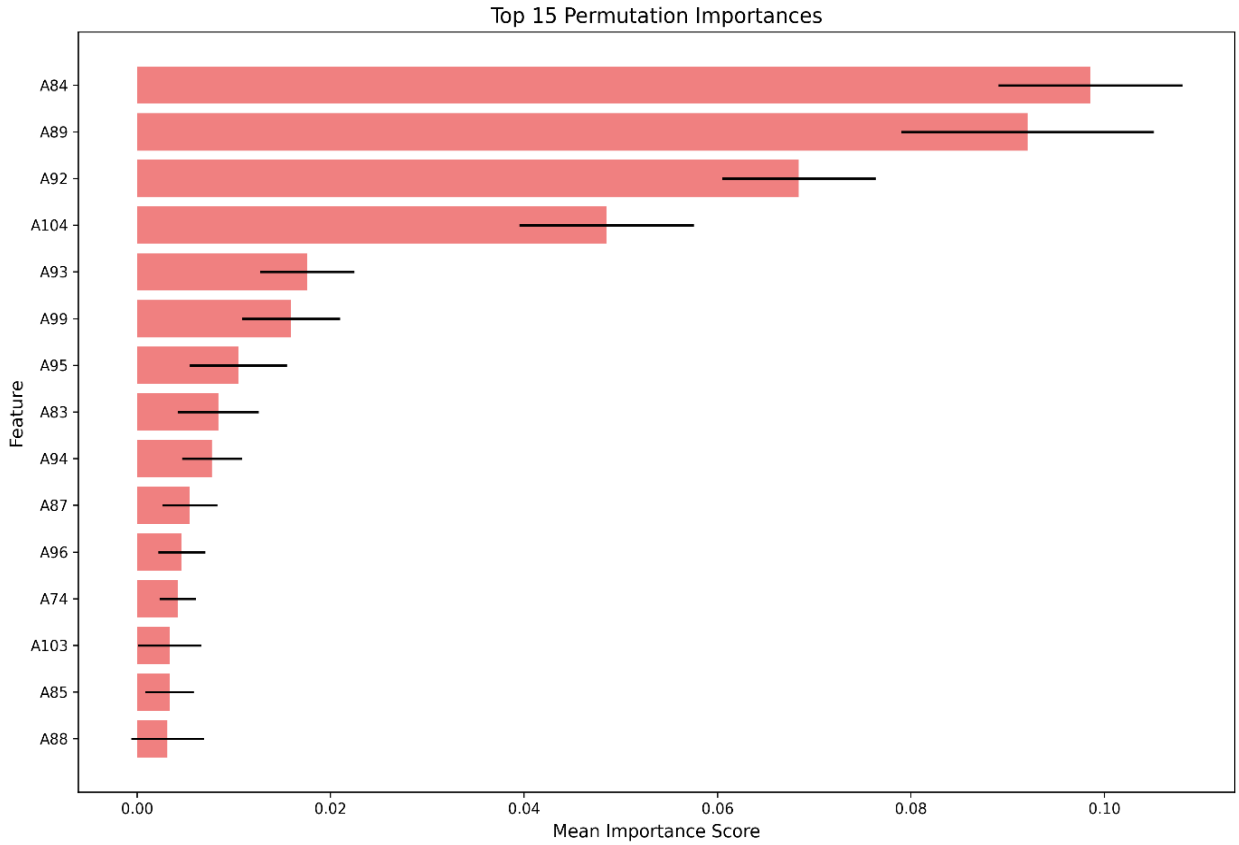

Figure 3. Permutation Importance Score

To complement the Feature Importance (FI) and Permutation Importance (PI) analyses, we conducted a Local Interpretable Model-Agnostic Explanations (LIME) analysis. LIME provides an interpretable, instance-based explanation by approximating the model's decision boundary with a simpler, interpretable model.

Figure 4 presents the local explanation for a single test sample. The horizontal bars represent the contribution of each feature to the classification decision. The red bars indicate

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149

negative contributions (reducing the probability of a specific class), while the green bars denote positive contributions (increasing the probability of the predicted class).

- A84 and A89 exhibit the highest negative influence on the classification outcome, indicating their strong contribution to shifting the prediction away from certain DNA classes.
- Conversely, A83 and A82 show significant positive contributions, reinforcing the model's classification decision.
- Other features such as A81, A88, and A86 play minor but noticeable roles in adjusting the classification outcome.

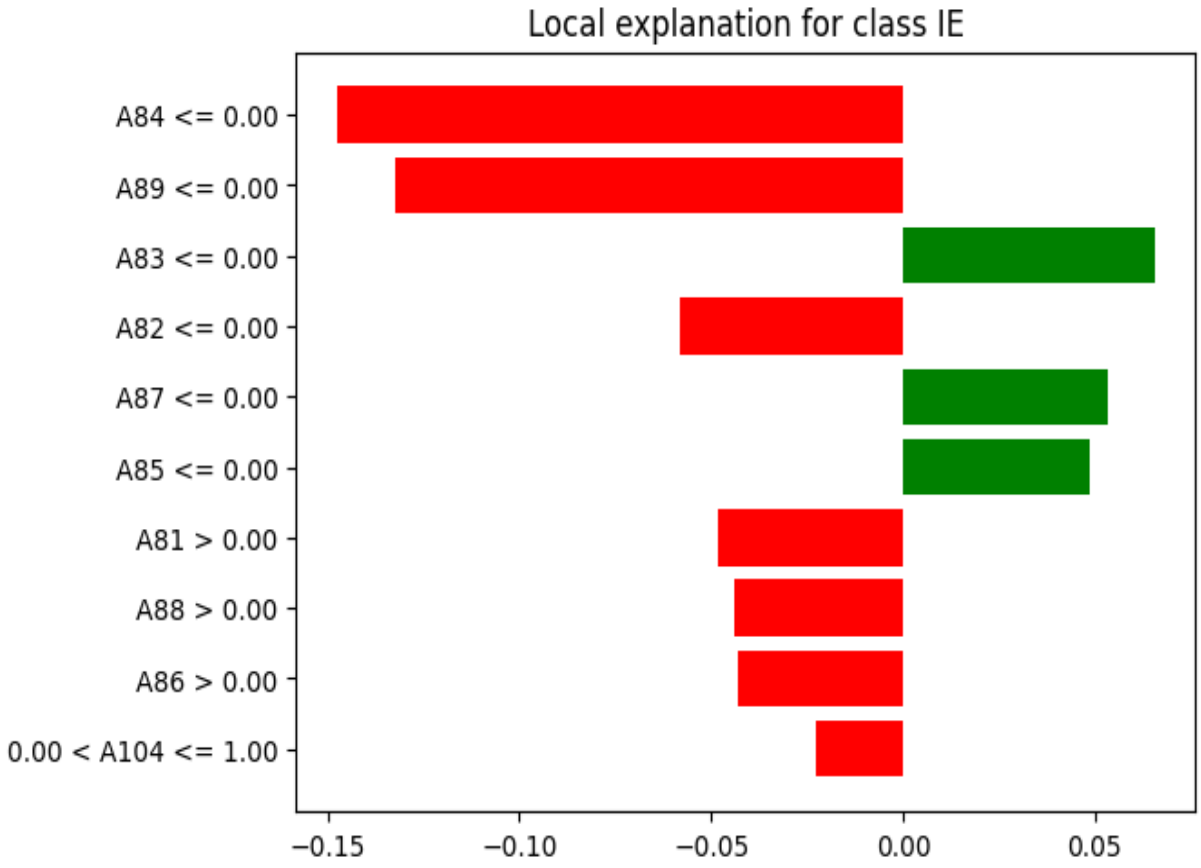

Figure 4. Local Explanation For Class IE

These results further validate the global Feature Importance (FI) and Permutation Importance (PI) findings by providing an instance-based justification for the significance of key features. This combination of global and local explainability techniques ensures a comprehensive understanding of the DNA sequence classification model.

4. CONCLUSION

This study presented a Hybrid Explainable AI Approach for DNA sequence classification, integrating Feature Importance (FI), Permutation Importance (PI), and Local Interpretable Model-Agnostic Explanations (LIME) to enhance the interpretability of machine learning predictions. The results demonstrated that features A84, A89, and A92 were consistently identified as the most significant across multiple explainability techniques. Feature

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149

Importance (FI) provided a global understanding of the dominant features influencing model decisions, while Permutation Importance (PI) validated their impact on classification performance. Furthermore, LIME offered localized explanations, allowing a more detailed analysis of individual predictions, which is crucial in understanding how specific DNA sequence characteristics contribute to classification outcomes. The combination of these explainability methods ensures a transparent and interpretable machine learning model, which is essential for applications in genomic research and biological analysis.

Although this approach offers valuable insights, there are several directions for future work. First, exploring more advanced models such as deep neural networks or Transformer-based architectures could improve classification performance while maintaining interpretability. Additionally, integrating domain-specific feature engineering may enhance feature representation and improve model accuracy. Further research can also investigate additional explainability techniques, such as SHAP (Shapley Additive Explanations) and counterfactual analysis, to gain deeper insights into feature contributions and interactions. Expanding the dataset with more diverse DNA sequences can improve the model's generalization ability, making it more robust for practical applications. Finally, applying this explainability framework in genomic medicine, disease prediction, and mutation analysis could extend its impact beyond classification tasks, contributing to advancements in computational biology and precision medicine.

5. ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the Faculty of Technology and Informatics and the Institute for Research and Community Service (LPPM), Universitas Informatika dan Bisnis Indonesia (UNIBI), for their continuous support and encouragement in conducting this research. Their invaluable assistance and resources have greatly contributed to the successful completion of this study.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism. Furthermore, all authors have significantly contributed to the research, including data collection, model development, explainability analysis, and manuscript writing. This research represents original work and has not been submitted or published elsewhere.

7. REFERENCES

[1] D. Bendigeri, L. Sakri, S. Mural, S. Hukkeri, and P. Tayannavar, "Human Genetic-based Disease Identification," in *2024 International Conference on Inventive Computation Technologies (ICICT)*, IEEE, 2024, pp. 486–491.

[2] H. Song *et al.*, "A novel approach utilizing domain adversarial neural networks for the detection and classification of selective sweeps," *Adv. Sci.*, vol. 11, no. 14, p. 2304842, 2024.

[3] F. M. Talaat, A. Aljadani, M. Badawy, and M. Elhosseini, "Toward interpretable credit scoring: integrating explainable artificial intelligence with deep learning for credit card default prediction," *Neural Comput. Appl.*, vol. 36, no. 9, pp. 4847–4865, 2024.

[4] A. Husnain, A. Shiwlani, M. N. Gondal, A. Ahmad, A. Saeed, and others, "Utilizing machine learning for proactive detection of cardiovascular risks: A data-driven

approach," *Int. J. Sci. Res. Arch.*, vol. 13, no. 1, pp. 1280–1290, 2024.
[5] E. Hikmawati and N. Alamsyah, "Supervised Learning for Emotional Prediction and Feature Importance Analysis Using SHAP on Social Media User Data," *Ingénierie Systèmes Inf.*, vol. 29, no. 6, 2024.
[6] V. Vimbi, N. Shaffi, and M. Mahmud, "Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection," *Brain Inform.*, vol. 11, no. 1, p. 10, 2024.
[7] U. Michelucci, "Feature Importance and Selection," in *Fundamental Mathematical Concepts for Machine Learning in Science*, Springer, 2024, pp. 229–242.
[8] N. Alamsyah, B. Budiman, T. P. Yoga, and R. Y. R. Alamsyah, "XGBOOST HYPERPARAMETER OPTIMIZATION USING RANDOMIZEDSEARCHCV FOR ACCURATE FOREST FIRE DROUGHT CONDITION PREDICTION," *J. Pilar Nusa Mandiri*, vol. 20, no. 2, pp. 103–110, 2024.
[9] J. Narkhede, "Comparative Evaluation of Post-Hoc Explainability Methods in AI: LIME, SHAP, and Grad-CAM," in *2024 4th International Conference on Sustainable Expert Systems (ICSES)*, IEEE, 2024, pp. 826–830.
[10] T. Khater *et al.*, "Explainable Machine Learning Model for Alzheimer Detection Using Genetic Data: A Genome-Wide Association Study Approach," *IEEE Access*, 2024.
[11] W. Zhou, Z. Yan, and L. Zhang, "A comparative study of 11 non-linear regression models highlighting autoencoder, DBN, and SVR, enhanced by SHAP importance analysis in soybean branching prediction," *Sci. Rep..*, vol. 14, no. 1, p. 5905, 2024.
[12] I. M. Zubair, Y.-S. Lee, and B. Kim, "A New Permutation-Based Method for Ranking and Selecting Group Features in Multiclass Classification," *Appl. Sci.*, vol. 14, no. 8, p. 3156, 2024.
[13] E. Cantor, S. Guauque-Olarte, R. León, S. Chabert, and R. Salas, "Knowledge-slanted random forest method for high-dimensional data and small sample size with a feature selection application for gene expression data," *BioData Min.*, vol. 17, no. 1, p. 34, 2024.
[14] J. Labory, E. Njomgue-Fotso, and S. Bottini, "Benchmarking feature selection and feature extraction methods to improve the performances of machine-learning algorithms for patient classification using metabolomics biomedical data," *Comput. Struct. Biotechnol. J.*, vol. 23, pp. 1274–1287, 2024.
[15] A. Puiu, C. Gómez Tapia, M. E. Weiss, V. Singh, A. Kamen, and M. Siebert, "Prediction uncertainty estimates elucidate the limitation of current NSCLC subtype classification in representing mutational heterogeneity," *Sci. Rep..*, vol. 14, no. 1, p. 6779, 2024.
[16] H. Lee, U. Ozbulak, H. Park, S. Depuydt, W. De Neve, and J. Vankerschaver, "Assessing the reliability of point mutation as data augmentation for deep learning with genomic data," *BMC Bioinformatics*, vol. 25, no. 1, p. 170, 2024.
[17] N. Alamsyah, A. P. Kurniati, and others, "A Novel Airfare Dataset To Predict Travel Agent Profits Based On Dynamic Pricing," in *2023 11th International Conference on Information and Communication Technology (ICoICT)*, IEEE, 2023, pp. 575–581.
[18] N. Alamsyah, B. Budiman, T. P. Yoga, and R. Y. R. Alamsyah, "Comparison of LINEAR REGRESSION AND RANDOM FOREST MODELS FOR PREDICTION OF UNDERGROUND DROUGHT LEVELS IN FOREST FIRES," *J. Techno Nusa Mandiri*, vol. 21, no. 2, pp. 81–86, 2024.
[19] N. Alamsyah, A. Hendra, E. Setiana, T. P. Yoga, V. R. Danestiara, and others, "Improved Prediction Of Global Temperature Via LSTM Using ReLU Activation And Hyperparameter Optimization," in *2024 International Conference on Information Technology Research and Innovation (ICITRI)*, IEEE, 2024, pp. 41–46.

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149

[20] A. G. Putrada, I. D. Oktaviani, M. N. Fauzan, and N. Alamsyah, "CNN Pruning for Edge Computing-Based Corn Disease Detection with a Novel NG-Mean Accuracy Loss Optimization," *Telematika*, vol. 17, no. 2, pp. 68–83, 2024.

DOI: https://doi.org/10.17509/coelite.v4i1.81356
p- ISSN 2829-4157 e- ISSN 2829-4149