

# Fast and Accurate COVID-19 Prediction with Optimized Machine Learning Approach

Abbas Jafar<sup>1</sup>, Myungho Lee<sup>2\*</sup>

<sup>1,2</sup> Department of Computer Engineering, Myongji University, Yongin, 17058, Republic of Korea

Corresponding author (Electronic mail: myunghol@mju.ac.kr)

The rapid transmission of COVID-19 has overwhelmed healthcare systems, creating an urgent demand for automated, accurate, and interpretable solutions to predict and control the virus. To address this, we propose an optimized machine learning framework that leverages clinical data with multiple features for precise COVID-19 detection. Our method first applies advanced feature selection techniques, including Mutual Information, Recursive Feature Elimination, and RidgeCV, to extract the most relevant features. These are then used to train machine learning models such as SVM, Gradient Boosting, Logistic Regression, and Adaboost. To further enhance model performance, we employ Genetic Algorithm-based hyperparameter optimization (GA-HPO) to fine-tune the model hyperparameters, achieving an accuracy of 96.30%. To ensure transparency and explainability, we use SHAP Explainable AI to identify key features, such as 'sex' and 'rate\_reducing\_mask,' that significantly influence the predictions. Our framework not only offers high accuracy but also ensures interpretability, making it a practical solution for early COVID-19 detection and resource allocation in healthcare.

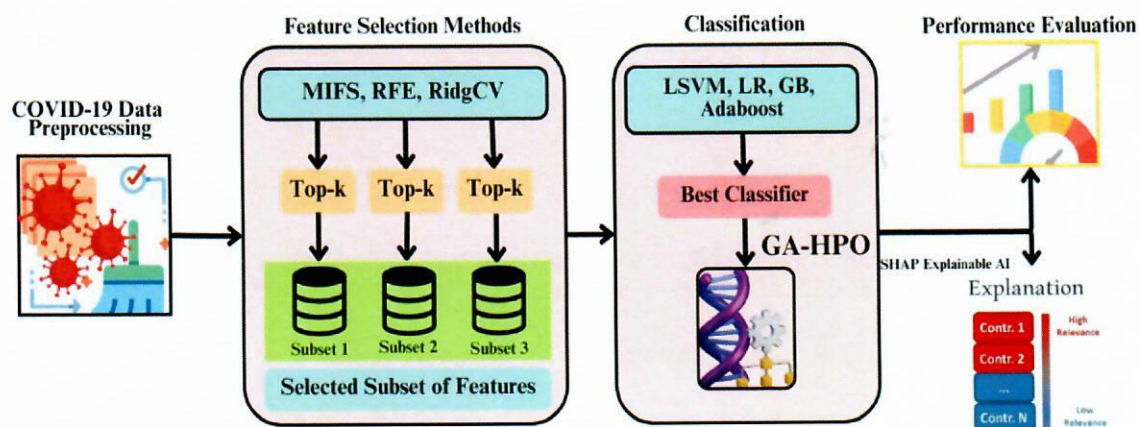


Figure 1 Overview of our proposed optimized ML approach for COVID-19 prediction

**Acknowledgments.** This work was supported by the National Research Foundation of Korea (NRF) grant funded by the government (MSIT) (RS-2023-00321688). ※ MSIT: Ministry of Science and ICT (Information and Communication Technology).

## References

- [1] G. Pascarella et al., "COVID-19 diagnosis and management: a comprehensive review," *J. Intern. Med.*, vol. 288, no. 2, pp. 192–206, Aug. 2020.
- [2] H. Lim, J. Lee, and D.-W. Kim, "Optimization approach for feature selection in multi-label classification," *Pattern Recognit. Lett.*, vol. 89, no. C, pp. 25–30, Apr. 2017.
- [3] K. Debjit et al., "An Improved Machine-Learning Approach for COVID-19 Prediction Using Harris Hawks Optimization and Feature Analysis Using SHAP," *Diagnostics*, vol. 12, no. 5, Art. no. 5, May 2022.
- [4] R. H. Ali and W. H. Abdulsalam, "The Prediction of COVID 19 Disease Using Feature Selection Techniques," *J. Phys. Conf. Ser.*, vol. 1879, no. 2, p. 022083, May 2021.