Original Article

Prediction of COVID-19 Vaccine Side Effects using SMOTE and Ensemble Machine Learning Models

R Vaishali[1], B Sarojini[2], D Sobya[3]

[1,2]*Department of Computer Science, Avinashilingam Institute for Home Science and Higher Education For Women, Coimbatore, Tamil Nadu, India.*
[3]*Department of Electronics and Communication Engineering, Rajarajeswari College of Engineering, Bangalore, Karnataka, India.*

[2]*Corresponding Author : sarojini_cs@avinuty.ac.in*

Abstract - While the rapid development of vaccines during the COVID-19 pandemic helped save billions of lives, a significant percentage of the population reported adverse reactions after vaccination. Post-vaccination surveys were conducted to understand the possible side effects. Analyzing and understanding the side effects enables the relevant stakeholders to gain more understanding of the vaccine and allows the individual to make informed decisions about whether to receive the vaccine. This work aims to identify a robust approach that can be used to predict the possibility of potential adverse symptoms. A dataset comprising 840 participants with 15 Post-Vaccination Symptoms was considered for the study. Synthetic Minority Oversampling TEchnique (SMOTE) was used to handle the class imbalance of the dataset. A combination of SMOTE and ensemble machine learning models was used to predict the adverse reactions to COVID-19 vaccines. The ensemble Machine Learning (ML) models that are considered for this study are Random Forest, Extreme Gradient Boosting Machine (XGBoost), Light Gradient Boosting Machine, and AdaBoost. The metrics accuracy, precision, recall and Receiver Operating Characteristic-Area Under the Curve (ROC-AUC) were used to measure the performance of the models. The dataset was pre-processed to handle missing values and one-hot encoding was applied to convert categorical variables into the numerical format. Insights into the data distribution and relationships between variables were gained through exploratory data and correlation analysis, respectively. Class imbalance in the target variables was addressed using SMOTE, resulting in a significantly improved F1-score and ROC-AUC score. Among the ensemble ML models, XGBoost delivered the best performance metrics. A combined performance score was calculated by averaging the F1-score and accuracy to identify the best model. XGBoost obtained the highest performance score among the ensemble ML models, and its performance is further enhanced by performing the threshold adjustment using the maximum F1-score strategy. The findings suggest that the combination of SMOTE and ensemble learning models with threshold adjustment provides a more efficient prediction of adverse effects after COVID-19 vaccination, aiding in healthcare decision-making.

Keywords - COVID-19, Ensemble Machine Learning, Exploratory Data Analysis, Imbalanced Dataset, SMOTE.

1. Introduction

The COVID-19 pandemic, triggered by the novel coronavirus SARS-CoV-2, has created an unprecedented global health emergency that has led to widespread illness and mortality [1, 2]. In response to this dire situation, the scientific community collaborated on an unprecedented scale to develop effective vaccines to combat the virus. The rapid development and deployment of COVID-19 vaccines marked a remarkable milestone in the history of medicine. Governments, healthcare organizations, and pharmaceutical companies worldwide joined forces to conduct rigorous clinical trials to ensure safety and efficacy [3–5]. The mass vaccination campaigns that followed sought to protect vulnerable populations and curb the spread of the virus. With millions of people receiving COVID-19 vaccines, the world witnessed an extraordinary effort to combat the pandemic and pave the way toward a more resilient and healthier future. However, despite these achievements, continuous monitoring, research, and evaluation of the vaccine's impact remain crucial to assess its effectiveness and address emerging challenges [6]. Post-vaccination surveys play a pivotal role in assessing the real-world impact and safety of COVID-19 vaccines.

These surveys involve collecting data from individuals who have received the vaccine, typically after a certain period following vaccination [7–10]. The surveys capture a wide range of information, including any adverse effects experienced, changes in health status, and overall vaccine satisfaction. By systematically gathering and analyzing this data, public health agencies and researchers can gain insights

into the prevalence and severity of post-vaccination side effects, identify rare or unexpected adverse events, and monitor the vaccine's effectiveness in preventing illness [7]. These surveys provide a valuable avenue for individuals to share their experiences and contribute to the collective understanding of vaccine safety and efficacy [11].

The data collected from post-vaccination surveys inform public health policies, guide communication strategies, and support ongoing efforts to optimize vaccination campaigns, ultimately contributing to the broader goal of controlling and mitigating the impact of the COVID-19 pandemic.

In 2021, a national referral hospital in Indonesia conducted a vaccination campaign among its staff and recorded post-vaccination symptoms. This study included 840 participants and 15 post-vaccination symptoms [5]. Statistical analysis was performed to study the impact of vaccination. A similar statistical study was conducted among healthcare professionals in Ethiopia [12]. Other studies include symptom-level impact analysis [13,14], vaccine-wise impact analysis [15], and long-term and short-term effects of the vaccines [16,17].

Various studies have used machine learning techniques to predict adverse effects following vaccination. These methodologies leverage the power of statistical data analysis and predictive modeling to enhance our understanding of potential post-vaccination outcomes [18, 19]. Researchers have harnessed machine learning algorithms to analyze large datasets, incorporating factors such as demographic information, medical history, and vaccine specifics to predict the likelihood and severity of adverse effects [19].

These predictive models aim to identify patterns and associations that can aid healthcare professionals and policymakers in making informed decisions. By utilizing machine learning in the context of post-vaccination adverse effects prediction, these studies contribute to the advancement of vaccine safety assessment and ultimately contribute to the overall success of vaccination campaigns and public health strategies.

A machine learning-based symptom prediction was attempted in research, where these post-vaccination symptoms were considered targets, and the participant information was taken as inputs to the model. The model attempted a multi-target classification model to predict multiple symptoms with one model. This model exhibited poor performance due to sparse patterns and class imbalance problems.

In the same work, another approach was attempted to predict the possibility of at least one symptom post-vaccination. After transforming the prediction target, the accuracy of prediction was 89% with the decision tree model.

A symptom-level model with class imbalance handling was suggested as an enhancement to work. Other machine learning works on adverse effects predictions were noticed on datasets extracted from the Vaccination Adverse Events Reporting System (VAERS) database [20-21].

2. Proposed Methodology
2.1. Framework
The framework of the proposed methodology is illustrated in Figure 1. The raw data was sourced from a survey conducted among hospital staff at a National Referral Hospital in Indonesia, specifically focusing on COVID-19 vaccination side effects.

The survey data collected from February 9th to February 13th, 2021, included responses from 840 participants, comprising 270 males and 570 females. To pre-process the data, missing values were replaced with default values, and categorical variables such as sex, profession, living area, and education were transformed into a numerical format using one-hot encoding.

Exploratory Data Analysis (EDA) provided insights into data distribution and helped identify patterns and outliers. Additionally, correlation analysis was conducted to understand the relationships between input features and target variables. To tackle the class imbalance in the target variables, SMOTE was employed and enhance predictive model capacity. The dataset was then split into training and testing sets, with 67% of the dataset allocated for training and the rest allocated to evaluate the performance of the models.

Four ensemble Machine Learning (ML) models are taken up for the study – Random Forest (RF), Extreme Gradient Boosting Machine (XGBoost), Light Gradient Boosting Machine (LGBM), and Adaptive Boosting (AdaBoost). The metrics accuracy, precision, recall, F1-score and Receiver Operating Characteristic-Area Under the Curve (ROC-AUC) were used to evaluate the performance of the models.

These metrics collectively offered a comprehensive evaluation of the model's predictive capabilities, with accuracy indicating overall correctness, F1-score providing a balance between precision and recall, precision assessing false positives, recall evaluating true positive rates, and ROC-AUC determining the models' ability to distinguish between different classes.

Following this rigorous methodology, the aim was to contribute valuable insights into predicting adverse effects of COVID-19 vaccination among hospital staff. The research findings hold potential significance in guiding healthcare decision-making and informing policy recommendations related to vaccination administration and side effect management.
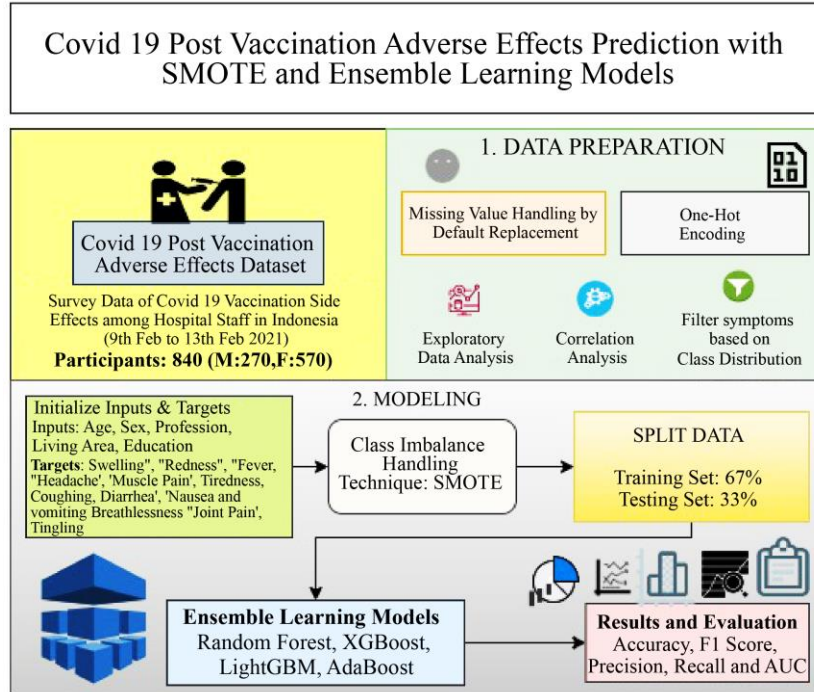
Fig. 1 Methodology of the end-to-end pipeline of the proposed modeling framework with SMOTE and Ensemble ML Models.

2.2. The SMOTE

SMOTE is a technique used to address the class imbalance in a dataset by generating synthetic samples for the minority class. This helps balance the class distribution and improves the performance of machine learning models. The SMOTE algorithm used in this work is obtained from the imblearn package of Sci-Kit Learn [22]. Choosing a suitable SMOTE model based on the data types is essential. SMOTE, in general, is designed to handle continuous data. While working with discrete or categorical data, it is essential to consider a suitable version of the SMOTE function [23-24].In SMOTE, synthetic samples for the minority class are generated to balance the classes in the dataset.

2.3. Ensemble ML Models

Ensemble ML models are used to enhance predictive accuracy, stability, and robustness by combining the outputs of multiple individual ML models. They mitigate the limitations of single models, capturing diverse patterns and reducing overfitting. Techniques like Bootstrap Aggregating (Bagging) and boosting ensemble models provide a more comprehensive understanding of complex data relationships, leading to improved generalization and more reliable predictions [25]. The ensemble ML models used in this work are:

2.3.1. RF

It is an ensemble learning model that constructs multiple decision trees during training and combines their predictions through voting or averaging. It mitigates overfitting by introducing randomness in the tree-building process.

Each decision tree is trained on a bootstrap sample of the data, and the final prediction is based on the majority vote or average of the individual tree predictions.

2.3.2. XGBoost

It is a boosting algorithm renowned for its efficiency and predictive power. It builds an ensemble of weak learners (typically decision trees) sequentially, with each new tree aimed at correcting the errors made by the previous ones.

XGBoost employs regularization techniques to prevent overfitting and provides insights into the importance of features.

2.3.3. LGBM

It is another gradient-boosting framework optimized for speed and efficiency. It employs a histogram-based approach for binning continuous feature values, leading to faster training times. LGBM uses a leaf-wise tree growth strategy and implements features like categorical variable handling and early stopping.

2.3.4. AdaBoost

AdaBoost is a pioneering ensemble model that iteratively merges the predictions of weak learners. It increases the weights of incorrectly classified instances in every iteration, thereby guiding subsequent learners to pay more attention to the more complex predictions. By adjusting the weights of the predictions from individual models according to their accuracy, AdaBoost effectively allocates greater significance to the more precise models.

2.4. Evaluation Metrics

The experiments were conducted in a cloud-based environment without General Processing Unit support, utilizing Python Programming. The performance of the ML models was evaluated using accuracy, recall, F1-score, precision, and the AUC.

2.4.1. Accuracy

Accuracy measures the ratio of correctly predicted instances to the total instances in the dataset. It provides an overall sense of how well the model's predictions match the outcomes.

2.4.2. Precision

Precision gauges are used to find the proportion of correctly predicted positive instances out of all instances predicted as positive. It indicates the model's ability to minimize false positives.

2.4.3. Recall

It is also known as Sensitivity or True Positive Rate, which measures the ratio of correctly predicted positive instances out of all actual positive instances. It quantifies the model's ability to capture all relevant positive instances.

2.4.4. F1-score

It is calculated by the harmonic mean of Precision and Recall. It provides a balanced evaluation metric, especially when there is an imbalance between classes or false positives and negatives must be minimized.

2.4.5. ROC-AUC

It assesses the model's ability to discriminate between positive and negative instances across different probability thresholds. A higher AUC indicates better overall performance in distinguishing between classes. Collectively, these metrics provide a thorough assessment of the model's predictive quality and ability to correctly classify instances and manage trade-offs between different aspects of performance.

3. Results and Discussions

This section discusses the performance of ensemble learning models. Table 1 shows the accuracy scores for symptom level prediction across four different ensemble ML models. Table 1 shows that the models performed better in detecting symptoms of redness, fever, diarrhea, nausea and vomiting, and breathlessness. The models predicted all the above-mentioned symptoms with accuracy scores ranging from 0.94 to 0.98. Conversely, the ensemble ML models predicted the symptoms of headache, muscle pain, tiredness, and coughing with comparatively lower accuracy scores, ranging from 0.54 to 0.69. XGB performed the best as it predicted all symptoms with the highest accuracy compared to the other models taken for the study.

Table 1. Accuracy scores of the ensemble ML models in predicting COVID-19 Symptoms

Accuracy Scores				
Symptoms	RF	XGB	LGBM	ADB
Swelling	0.83	0.84	0.82	0.77
Redness	0.96	0.95	0.95	0.94
Fever	0.97	0.96	0.96	0.97
Headache	0.67	0.69	0.65	0.65
Muscle Pain	0.6	0.61	0.6	0.58
Tiredness	0.58	0.59	0.56	0.54
Coughing	0.86	0.87	0.85	0.77
Diarrhea	0.95	0.96	0.94	0.93
Nausea and vomiting	0.96	0.96	0.95	0.94
Breathlessness	0.97	0.96	0.96	0.94
Joint Pain	0.98	0.98	0.98	0.96

Table 2. Precision scores of the ensemble ML models in predicting COVID-19 Symptoms

Precision scores				
Symptoms	RF	XGB	LGBM	ADB
Swelling	0.79	0.8	0.78	0.72
Redness	0.98	0.98	0.97	0.94
Fever	0.94	0.94	0.93	0.96
Headache	0.69	0.71	0.68	0.67
Muscle Pain	0.55	0.56	0.55	0.53
Tiredness	0.54	0.55	0.53	0.51
Coughing	0.85	0.86	0.83	0.77
Diarrhea	0.95	0.95	0.94	1.00
Nausea and vomiting	0.98	0.98	0.97	0.94
Breathlessness	0.94	0.93	0.92	0.95
Joint Pain	0.96	0.96	0.96	0.98

The precision scores of the ensemble learning models are tabulated in Table 2. High precision scores were observed for symptoms - redness, diarrhea, nausea and vomiting, breathlessness, and joint pain, with scores ranging from 0.92 to 0.98 and even reaching 1.00 for diarrhea predicted by AdaBoost. A high precision score for the above-mentioned symptoms indicates that the models make predictions with a low false positive rate. On the contrary, symptoms like muscle pain, tiredness, and headache exhibited lower precision scores, between 0.51 and 0.71, indicating a higher rate of false positives. These results reflect the models' varying levels of specificity across different symptoms, with some models demonstrating a high precision in their predictions.

In contrast, others suggest potential areas for reducing the incidence of false positives. Table 3 shows the recall scores of the ensemble ML models taken for the study/ The recall scores reveal the models' ability to identify positive cases for each symptom correctly. Notably, high recall scores were observed for symptoms - fever, breathlessness, and joint pain, with scores reaching up to 1.00 for some models, indicating an exceptional ability of these models to identify almost all positive cases for these symptoms.

Table 3. Recall scores of the ensemble ML models in predicting COVID-19 Symptoms

Recall scores				
Symptoms	**RF**	**XGB**	**LGBM**	**ADB**
Swelling	0.92	0.92	0.92	0.92
Redness	0.93	0.93	0.93	0.93
Fever	0.99	0.99	0.99	0.97
Headache	0.69	0.7	0.64	0.68
Muscle Pain	0.64	0.62	0.61	0.68
Tiredness	0.64	0.66	0.56	0.69
Coughing	0.88	0.90	0.88	0.8
Diarrhea	0.96	0.96	0.95	0.87
Nausea and vomiting	0.93	0.93	0.93	0.93
Breathlessness	1.00	1.00	1.00	0.91
Joint Pain	0.99	0.99	0.99	0.94

Table 5. ROC-AUC Scores of the ensemble ML models in predicting COVID-19 symptoms

ROC-AUC scores				
Symptoms	**RF**	**XGB**	**LGBM**	**ADB**
Swelling	0.93	0.93	0.91	0.88
Redness	0.99	0.99	0.98	0.99
Fever	0.99	0.99	0.99	0.99
Headache	0.74	0.76	0.71	0.68
Muscle Pain	0.64	0.65	0.64	0.65
Tiredness	0.62	0.61	0.59	0.55
Coughing	0.94	0.94	0.92	0.89
Diarrhea	0.99	0.99	0.98	0.97
Nausea and vomiting	0.99	0.99	0.99	0.99
Breathlessness	1.00	1.00	0.99	0.99
Joint Pain	1.00	1.00	1.00	1.00

The models predicted symptoms such as redness, swelling, coughing, nausea, and vomiting with high recall scores, ranging from 0.88 to 0.93, reflecting a strong capability in correctly identifying a high proportion of positive instances. Conversely, for symptoms like headache, muscle pain, and tiredness, the models exhibited lower recall scores, between 0.56 and 0.7, suggesting a relatively weaker performance in identifying positive cases for these symptoms.

These results demonstrate the effectiveness of the models in varying degrees across different symptoms, with some symptoms being identified with high accuracy. In contrast, others indicate areas where model sensitivity could be improved.

Table 4 depicts the evaluated F1-scores of the ensemble ML models, revealing insights into the models' precision and recall balance. It was found that the F1 scores of the ensemble ML models were notably high for symptoms - fever, diarrhea, nausea and vomiting, breathlessness, and joint pain, with values ranging between 0.93 and 0.98 across the models, indicating a strong balance between precision and recall for these symptoms.

Table 4. F1-scores of the ensemble ML models in predicting COVID-19 Symptoms

F1-scores				
Symptoms	**RF**	**XGB**	**LGBM**	**ADB**
Swelling	0.85	0.86	0.84	0.81
Redness	0.96	0.95	0.95	0.94
Fever	0.97	0.96	0.96	0.97
Headache	0.69	0.71	0.66	0.68
Muscle Pain	0.59	0.59	0.58	0.6
Tiredness	0.59	0.6	0.55	0.59
Coughing	0.86	0.88	0.86	0.78
Diarrhea	0.95	0.95	0.94	0.93
Nausea and vomiting	0.95	0.96	0.95	0.93
Breathlessness	0.97	0.96	0.96	0.93
Joint Pain	0.98	0.98	0.98	0.96

On the other hand, for symptoms such as muscle pain, tiredness, and headache, the model gave lower F1 scores, falling between 0.55 and 0.71, which suggests a lesser degree of balance between precision and recall for the prediction of these symptoms. The results reflect the models' varying effectiveness in accurately predicting different symptoms, with some symptoms achieving high precision and recall while others demonstrate areas for potential improvement in model performance.

The AUC associated with the prediction of various symptoms was analyzed to assess the models' ability to distinguish between the symptom's presence and absence. High AUC scores were observed for symptoms - redness, fever, diarrhea, nausea and vomiting, breathlessness, and joint pain, with scores reaching up to 1.00, indicating the models' perfect ability to differentiate between positive and negative cases for these symptoms. This suggests that the models are capable of ranking positive cases higher than negative ones with high confidence.

The AUC scores were also high for coughing and swelling, ranging from 0.9 to 0.94, reflecting a strong discriminatory ability between the symptom's presence and absence. However, for symptoms like headache, muscle pain, and tiredness, the AUC scores were comparatively lower, ranging from 0.59 to 0.76, indicating a lesser capability of these models to distinguish between positive and negative cases effectively. These lower scores suggest that there is a more significant overlap between the distribution of scores for positive and negative cases, leading to a higher chance of misclassification.

To decide the best ensemble, we compute the average performance score for each model, using Equation 1, Equation 2 and Equation 3. Acc_model $_{symptoms}$ is the accuracy score of the model for a particular symptom, $T_{Symptoms}$ is the total number of symptoms, Avg_acc$_{model}$ is the average accuracy of the model, F1_score_model $_{symptoms}$ is the F1-score of the

model for a particular symptom, Avg_F1_score$_{model}$ is the average F1-score of the model, Avg_score$_{model}$ is the average score of the model.

$$\text{Avg\_acc}_{model} = \left[\sum \text{Acc\_model}_{symptoms}\right]/\text{T}_{Symptoms} \qquad (1)$$

$$\text{Avg\_F1\_score}_{model} = \left[\sum \text{F1\_score\_model}_{symptoms}\right]/\text{T}_{Symptoms} \qquad (2)$$

$$\text{Avg\_score}_{model} = (\text{Avg\_acc}_{model} + \text{Avg\_F1score}_{model})\,/2 \qquad (3)$$

Figure 2 shows the average performance score of the four ensemble ML models. It is observed that the XGBoost outperforms other algorithms with an average performance score of 0.85. Random Forest and LGBM displayed a close performance.

To enhance the performance, we chose the best model, XGB, and tuned the thresholds for classification. The thresholds were varied between 0.4 to 0.7 with an increment of 0.01 in each step, and the performance metrics at each increment were computed. The threshold which yielded the maximum F1-score was decided to be the optimum threshold. The results after threshold adjustment were also compared with the existing results, as shown in Figure 3.

From Figure 3, it is observed that after threshold adjustment, there was a significant improvement in the recall rate and accuracy in most of the cases. This strategy could not improve the model's performance in some cases, like 'Tiredness' and 'Breathlessness'.

Hence, there is a need to check the data inconsistencies and efficiency of the oversampling techniques in future work. There is also a possibility of a statistical drift due to the oversampling strategy.

To understand the excellent performance exhibited by XGB, the average feature importance of the attributes across symptoms is visualized in Figures 4 and 5. The plots reveal that attributes such as age, profession, and education had more significance in the model decisions. The feature significance scores also indicate that doctors and medical professionals with master's degrees in the age group 36-40 years had more significance with the symptoms. From a real-world point of view, it is evident that medical professionals had high exposure to the infected population.

Also, the nature of work during the peak infection days could have contributed to the occurrence of adverse effects. The features such as age greater than or equal to 60 and education belonging to junior high school obtained lower feature importance scores, indicating the underrepresentation of these populations in the data sample taken for modeling. As the sample contains a hospital staff population, the chances of staff with education less than graduation or age above 60 years are less. This insight emphasizes the need for data variety in the dataset. Otherwise, the model for decision-making might not capture the underrepresented patterns.

In our work, the dataset was SMOTE to handle class imbalance. Oversampling tends to cause data to drift into statistical distributions due to its intrinsic nature of increasing the volume of minority samples. However, the feature importance insights show that the oversampling was able to retain the basic quality of the data sample.

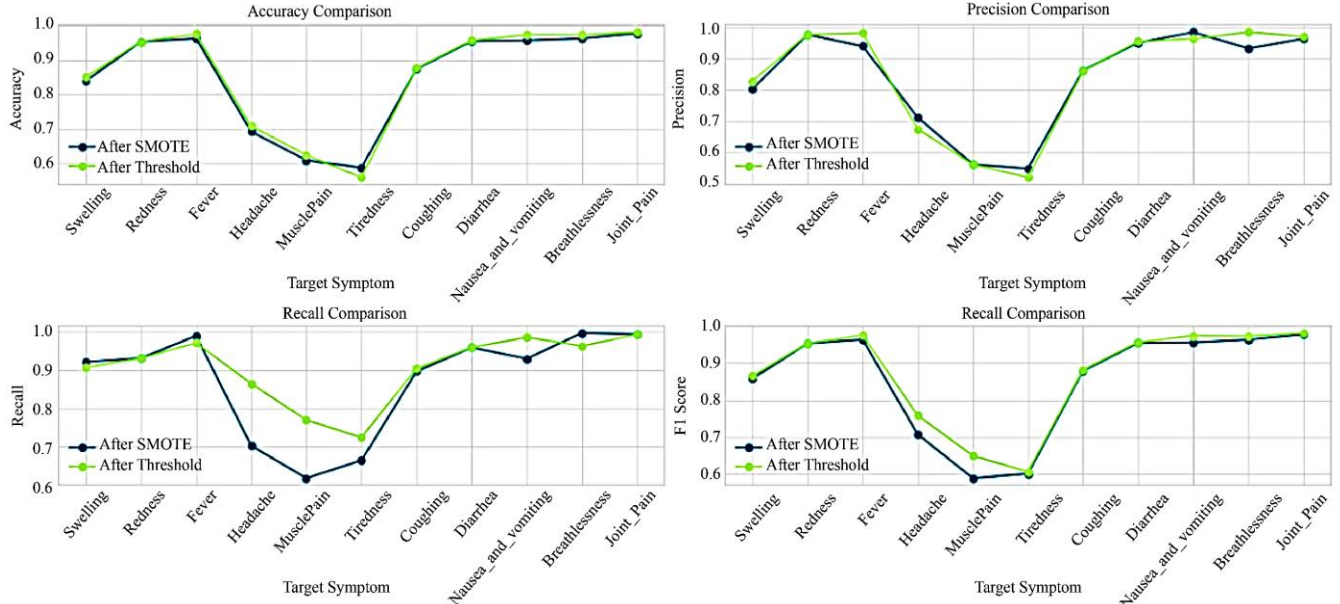**Fig. 2 Comparison of average performance of the ensemble ML models after SMOTE**

Fig. 3 Performance of the XGBoost model after Threshold adjustment

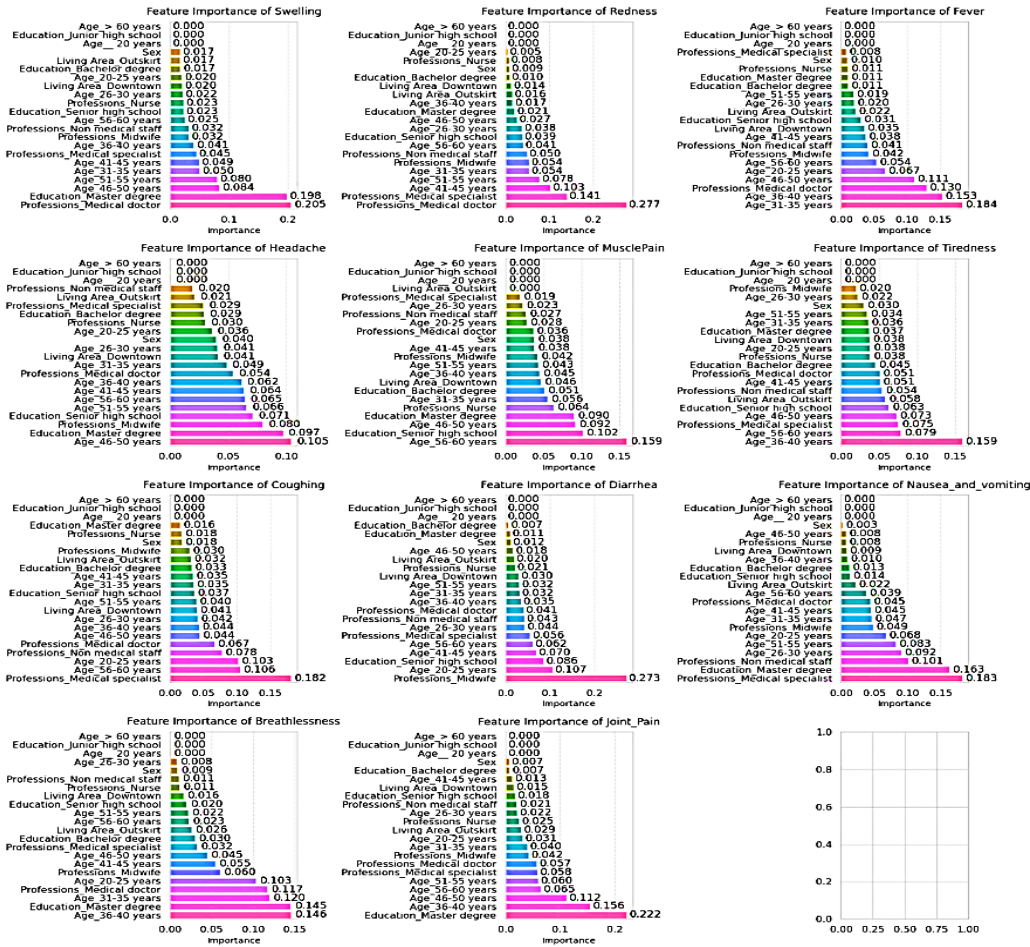**Fig. 4 Distribution of average of feature importance scored by attributes at each symptom level in the XGBoost model**
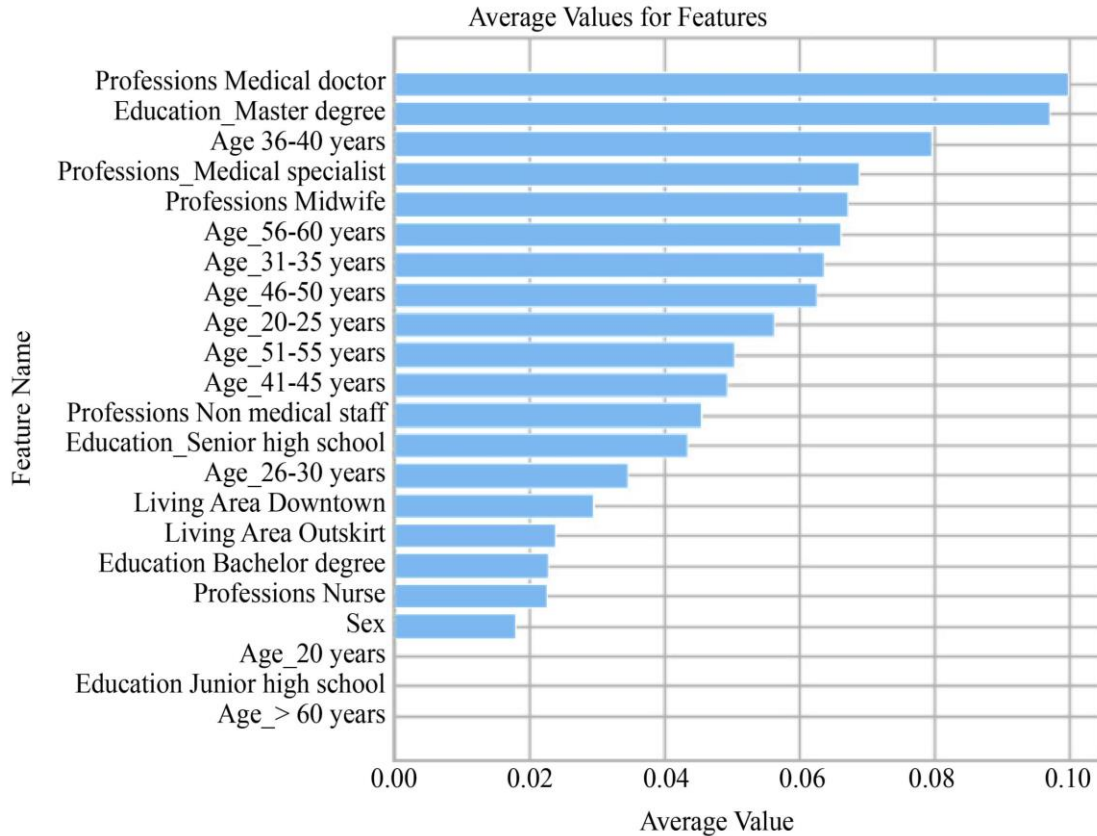
Average Values for Features

Fig. 5 Distribution of average of feature importance scored by attributes at each symptom level in the XGBoost model

4. Conclusion and Future Work

This work outlines a framework to predict the possible symptoms of COVID-19 vaccination better using the information on an individual background. The empirical study of ensemble ML models - RF, XGB, LGBM, and AdaBoost, for predicting post-vaccination symptoms has revealed substantial potential in leveraging these techniques for improved diagnostic accuracy and patient care. The analysis across various metrics - accuracy, precision, recall, F1-score, and ROC-AUC highlighted the models' effectiveness in predicting symptoms and showcased the critical role of machine learning in healthcare applications. However, the study also uncovered significant drawbacks, particularly in the models' performance on symptoms with lower predictive scores, underscoring a vital enhancement area. A key factor influencing these outcomes could be the implementation of addressing class imbalance in the dataset. While SMOTE helps by artificially generating new samples from the minority class, thereby balancing the distribution between classes, it can also introduce synthetic noise into the dataset. This may potentially affect the models' ability to learn from the training data to unseen data, impacting overall predictive performance. To mitigate these issues and further enhance predictive accuracy, future research will explore alternative oversampling techniques that might offer more nuanced approaches to handling class imbalance without introducing excessive noise. Additionally, advanced hyperparameter optimization techniques will be employed to fine-tune the models, and the exploration of deep learning models will be prioritized for their potential to capture complex data patterns more effectively.

Future work would also refine the understanding of vaccination outcomes and improve risk assessment with an enhanced method, offering new avenues for personalized medicine and public health initiatives.

References

[1] Neha Purohit et al., "COVID-19 Management: The Vaccination Drive in India," *Health Policy Technology*, vol. 11, no. 2, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Simran Preet Kaur, and Vandana Gupta, "COVID-19 Vaccine: A Comprehensive Status Report," *Virus Research*, vol. 288, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] Qutaiba A. Al Khames Aga et al., "Safety of COVID-19 Vaccines," *Journal of Medical Virology*, vol. 93, no. 12, pp. 6588-6594, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Mangalakumari Jeyanathan et al., "Immunological Considerations for COVID-19 Vaccine Strategies," *Nature Reviews Immunology*, vol. 20, pp. 615-632, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Dovy Djanas et al., "Survey Data of COVID-19 Vaccine Side Effects among Hospital Staff in a National Referral Hospital in Indonesia," *Data in Brief*, vol. 36, pp. 1-7, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] Nikolaos Tzenios, Mohamad Chahine, and Mary Tazanios, "Better Strategies For Coronavirus (COVID-19) Vaccination," *Special Journal of the Medical Academy and other Life Sciences*, vol. 1, no. 2, pp. 1-7, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Eng Eong Ooi et al., "Use of Analgesics/Antipyretics in the Management of Symptoms Associated with COVID-19 Vaccination," *NPJ Vaccines*, vol. 7, no. 1, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Georgi Bogdanov et al., "Cutaneous Adverse Effects of the Available COVID-19 Vaccines," *Clinics in Dermatology*, vol. 39, no. 3, pp. 523-531, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Jeffrey V. Lazarus et al., "A Survey of COVID-19 Vaccine Acceptance across 23 Countries in 2022," *Nature Medicine*, vol. 29, no. 2, pp. 366-375, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] Abanoub Riad et al., "COVID-19 Vaccines Safety Tracking (CoVaST): Protocol of a Multi-Center Prospective Cohort Study for Active Surveillance of COVID-19 Vaccines' Side Effects," *International Journal of Environmental Research and Public Health*, vol. 18, no. 15, pp. 1-10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Yung-Tsan Jou et al., "Factors Affecting Perceived Effectiveness of Government Response towards COVID-19 Vaccination in Occidental Mindoro, Philippines," *Healthcare*, vol. 10, no. 8, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Mesfin Tafa Segni et al., "Post COVID-19 Vaccination Side Effects and Associated Factors among Vaccinated Health Care Providers in Oromia Region, Ethiopia in 2021," *PLoS One*, vol. 17, no. 12, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Matteo Castaldo et al., "Headache Onset after Vaccination against SARS-CoV-2: A Systematic Literature Review and Meta-Analysis," *The Journal of Headache and Pain*, vol. 23, no. 1, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Kyung Ae Lee, Yu Ji Kim, and Heung Yong Jin, "Thyrotoxicosis after COVID-19 Vaccination: Seven Case Reports and a Literature Review," *Endocrine*, vol. 74, pp. 470-472, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Balsam Qubais Saeed et al., "Side Effects and Perceptions Following Sinopharm COVID-19 Vaccination," *International Journal of Infectious Diseases*, vol. 111, pp. 219-226, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Liane S. Canas et al., "Disentangling Post-Vaccination Symptoms from Early COVID-19," *EClinicalMedicine*, vol. 42, pp. 1-10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Kin Israel Notarte et al., "Impact of COVID-19 Vaccination on the Risk of Developing Long-COVID and on Existing Long-COVID Symptoms: A Systematic Review," *EClinicalMedicine*, vol. 53, pp. 1-19, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Martuza Ahamad et al., "Adverse Effects of COVID-19 Vaccination: Machine Learning and Statistical Approach to Identify and Classify Incidences of Morbidity and Postvaccination Reactogenicity," *Healthcare*, vol. 11, no. 1, pp. 1-22, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] V. Indira et al., "Prediction and Analysis of Covid-19 using the Deep Learning Models," *Research on Biomedical Engineering*, vol. 40, pp. 183-197, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] James Flora et al., "Usefulness of Vaccine Adverse Event Reporting System for Machine-Learning Based Vaccine Research: A Case Study for COVID-19 Vaccines," *International Journal of Molecular Sciences*, vol. 23, no. 15, pp. 1-19, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[21] Himanshu Gupta, and Om Prakash Verma, "Vaccine Hesitancy in the Post-Vaccination COVID-19 Era: A Machine Learning and Statistical Analysis Driven Study," *Evolutionary Intelligence*, vol. 16, pp. 739-757, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Nitesh V. Chawla et al., "SMOTE: Synthetic Minority Over-Sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[23] Mimi Mukherjee, and Matloob Khushi, "SMOTE-ENC: A Novel SMOTE-Based Method to Generate Synthetic Data for Nominal and Continuous Features," *Applied System Innovation*, vol. 4, no. 1, pp. 1-12, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] Elif Ceren Gök, and Mehmet Onur Olgun, "SMOTE-NC and Gradient Boosting Imputation Based Random Forest Classifier for Predicting Severity Level of COVID-19 Patients with Blood Samples," *Neural Computing and Applications*, vol. 33, pp. 15693-15707, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Jenni A. M. Sidey-Gibbons, and Chris J. Sidey-Gibbons, "Machine Learning in Medicine: A Practical Introduction," *BMC Medical Research Methodology*, vol. 19, pp. 1-18, 2019. [CrossRef] [Google Scholar] [Publisher Link]