PREDICTIVE MODELING OF FOOD ALLERGIES AND NUTRIENT SENSITIVITIES USING ADVANCED MACHINE LEARNING TECHNIQUES

[1]Sushil Kumari, [2]Shivangi Sharma, [3]Dr. Dinesh Kumar, [4]Navneet Gupta

[1]Assistant Professor, Sri Sai iqbal college of management and information technology, Badhani-Pathankot, Punjab, India
[2,3]Associate Professor, Sri Sai College of Engineering and Technology, Badhani-Pathankot, Punjab, India.
[4]Assistant Professor, Sri Sai University, palampur, Himachal Pradesh,

mehraoct@gmail.com, shivangisharma15391@gmail.com, dineshgarg82@gmail.com, mail2navneetgupta@gmail.com

Abstract

Food allergies and nutrient sensitivities pose significant health challenges, impacting individuals' quality of life and increasing healthcare costs. Predictive modeling using advanced machine learning techniques offers a promising approach to enhance early detection, personalized management, and preventive strategies for these conditions. This research explores the application of sophisticated machine learning algorithms, including ensemble methods, deep learning, and feature selection techniques, to predict food allergies and nutrient sensitivities based on dietary logs, genetic information, and clinical data. We employed a diverse set of machine learning models such as Random Forests, Gradient Boosting Machines, Convolutional Neural Networks (CNNs), and Long Short-Term Memory (LSTM) networks to analyze and interpret complex data patterns. The study integrated data sources including patient medical histories, genetic predispositions, and dietary intake logs to develop robust predictive models. Model performance was evaluated using metrics such as accuracy, precision, recall, and the area under the receiver operating characteristic curve (AUC-ROC). The findings indicate that advanced machine learning techniques can significantly improve the prediction accuracy of food allergies and nutrient sensitivities. These models provide insights into the underlying patterns and correlations between dietary habits, genetic factors, and allergy manifestations. By enabling early detection and personalized dietary recommendations, these predictive models hold the potential to enhance individual health management and contribute to more effective public health strategies. Future research should focus on refining these models, expanding data sources, and validating the predictions in diverse populations to ensure generalizability and practical applicability.

Keywords: Predictive Modeling, Food Allergies, Nutrient Sensitivities, Machine Learning Techniques, Personalized Health Management, Dietary Logs

1. Introduction

Food allergies and nutrient sensitivities represent a growing public health concern with significant implications for individual well-being and healthcare systems. Food allergies, which involve an immune system response to specific proteins in food, can lead to severe and potentially life-threatening reactions, while nutrient sensitivities, often characterized by adverse responses to certain foods or nutrients, can affect quality of life and overall health. Despite advances in medical research, diagnosing and managing these conditions remain challenging due to their complex and often individualized nature. The traditional approach to diagnosing food allergies and nutrient sensitivities typically involves clinical evaluations, such as patient history, skin prick tests, and oral food challenges. However, these methods have limitations, including variability in patient responses, the potential for false positives or negatives, and the invasive nature of some tests. Moreover, the management of these conditions often relies on elimination diets and trial-and-error strategies, which can be both time-consuming and burdensome for patients [1].

In recent years, the integration of advanced machine learning techniques into healthcare has opened new avenues for improving predictive accuracy and personalization in medical diagnostics. Machine learning algorithms, capable of analyzing vast amounts of complex data, offer the potential to enhance early detection, tailor individual management plans, and ultimately improve patient outcomes. By leveraging diverse data sources—such as dietary logs, genetic profiles, and clinical records—these algorithms can identify patterns and correlations that may not be evident through traditional diagnostic methods. This study aims to explore the application of advanced machine learning techniques in predictive modeling for food allergies and nutrient sensitivities [2]. The primary objective is to develop and evaluate sophisticated models that can accurately predict these conditions based on comprehensive data inputs. Techniques such as ensemble methods (e.g., Random Forests and Gradient Boosting Machines), deep learning (e.g., Convolutional Neural Networks and Long Short-Term Memory networks), and feature selection approaches are employed to analyze complex interactions between dietary habits, genetic factors, and clinical symptoms.

Ensemble methods, known for their ability to improve predictive performance by combining multiple models, are particularly suited for handling the variability and complexity inherent in allergy and sensitivity data. Gradient Boosting Machines, with their iterative approach to reducing prediction errors, offer a powerful tool for fine-tuning model accuracy [3]. Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, on the other hand, excel in handling sequential and spatial data, making them ideal for analyzing time-series data from dietary logs and genetic sequences. The significance of this research lies in its potential to transform the management of food allergies and nutrient sensitivities. By providing more accurate and personalized predictions, the proposed models can facilitate earlier and more precise diagnoses, leading to targeted interventions and improved patient outcomes. Additionally, the insights gained from this study could contribute to the development of more effective public health strategies and preventive measures.

2. Literature Review

The landscape of predictive modeling for food allergies and nutrient sensitivities has evolved significantly, driven by advancements in machine learning and data analytics. Food allergies, characterized by immune system reactions to specific food proteins, and nutrient sensitivities, which involve adverse responses to certain dietary components, present unique challenges for diagnosis and management. Historically, clinical methods such as skin prick tests and oral food challenges have been used to identify these conditions. However, these approaches have limitations, including variability in patient responses and potential for false results, necessitating the exploration of alternative diagnostic methods [1]. Recent research has highlighted the potential of machine learning to address these challenges by leveraging large and complex datasets. Machine learning techniques, including supervised and unsupervised learning methods, have been applied to a range of medical conditions with promising results. For example, Random Forests and Gradient Boosting Machines, which are ensemble learning methods, have shown significant improvements in predictive accuracy across various healthcare applications [2][3]. These methods work by aggregating multiple decision trees to improve prediction performance and manage the inherent variability in medical data.

In the realm of food allergies, machine learning models have been employed to enhance diagnostic precision and predict individual risk profiles. For instance, a study by Kottmann et al. demonstrated that machine learning algorithms could improve the accuracy of predicting allergic reactions based on patient histories and allergy test results [4]. Similarly, Nguyen et al. applied deep learning techniques to analyze genetic data, revealing potential biomarkers associated with food allergies and sensitivities [5]. These studies underscore the growing interest in integrating machine learning with traditional diagnostic methods to create more comprehensive and accurate predictive models.

Nutrient sensitivities, which often involve complex interactions between dietary intake and individual responses, present additional challenges. Recent advancements in machine learning, particularly in the use of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, offer new opportunities for analyzing time-series and sequential data from dietary logs [6][7]. CNNs, traditionally used in image processing, have been adapted to handle structured data such as nutritional information, while LSTMs are well-suited for capturing temporal patterns in dietary intake over time [8][9]. Feature selection and dimensionality reduction techniques also play a critical role in improving model performance. By identifying the most relevant features from large datasets, these techniques help in reducing overfitting and enhancing the interpretability of machine learning models. Studies have demonstrated the effectiveness of feature selection methods in improving the accuracy of predictive models for various health conditions, including food allergies [10][11]. For example, research by Zhang et al. highlighted the importance of selecting relevant features to improve model predictions and reduce computational complexity [12].

Despite these advancements, challenges remain in the application of machine learning to predictive modeling for food allergies and nutrient sensitivities. Issues such as data quality, variability, and the need for large, diverse datasets continue to impact model accuracy and generalizability. Additionally, the integration of diverse data sources, including genetic, dietary, and clinical information, requires careful consideration of data preprocessing and integration techniques [13][14]. Research by Smith et al. emphasized the importance of addressing these challenges to ensure the robustness and applicability of machine learning models in clinical settings [15]. The literature highlights the significant progress made in applying machine learning techniques to the predictive modeling of food allergies and nutrient sensitivities. Ensemble methods, deep learning techniques, and feature selection approaches have all contributed to improving predictive accuracy and personalization. However, challenges such as data variability and integration remain, underscoring the need for continued research and development in this field. The following sections will build upon these insights to explore the methodology, results, and implications of applying advanced machine learning techniques to predictive modeling for food allergies and nutrient sensitivities.

Table 1: summary table of the literature review for predictive modeling of food allergies and nutrient sensitivities

Methodology	**Techniques Used**	**Data Types**	**Key Findings**	**Strengths**	**Limitations**	**Applications**	**Future Directions**
Machine learning analysis	Decision Trees, Random Forests	Patient histories, allergy tests	Improved prediction of allergic reactions; high accuracy in specific contexts	Enhanced prediction precision	Limited to specific allergens	Allergy diagnostics	Expand to broader allergen spectrum
Deep learning	CNNs, Genetic analysis	Genetic data	Identified potential biomarkers for food allergies; insights into genetic predispositions	Integration with genetic data	Requires large genetic datasets	Genetic predisposition analysis	Explore more genetic markers
Feature selection	Feature Selection, Classification	Clinical and dietary data	Improved model accuracy by selecting relevant features; reduced overfittin	Reduced computational complexity	Dependence on feature relevance	Model optimization	Develop automated feature selection

			g				
Data integration	Machine Learning Integration	Diverse data sources	Addressed data integration challenges; improved model robustness	Comprehensive approach to data integration	Data quality variability	Multi-source data analysis	Improve data integration methods
Ensemble methods	Random Forests, Gradient Boosting	Clinical and dietary data	Enhanced predictive performance with ensemble methods; effective in varied datasets	Improved performance through multiple models	Model complexity can be high	Predictive modeling	Simplify model structures
Sequential data analysis	LSTMs	Dietary logs	Effective in capturing temporal patterns; high accuracy in sequential data analysis	Captures time-dependent data	Requires extensive time-series data	Dietary pattern analysis	Extend to more complex datasets
Time-series analysis	LSTMs	Dietary intake logs	Improved prediction of nutrient sensitivities; good performance with time-series data	Handles sequential dietary data effectively	Data preprocessing complexity	Nutrient sensitivity prediction	Refine time-series modeling
Predictive	Deep	Nutritio	Accurate	Effective	May	Allergy	Integrate

modeling	Learning, CNNs	nal information	prediction of food allergies based on dietary logs and clinical data	with structured dietary data	require substantial computational resources	and sensitivity prediction	with real-time data
Machine learning application	Ensemble Methods, Deep Learning	Clinical, dietary, genetic data	Addressed challenges in predictive accuracy; demonstrated robustness in varied settings	Robust model performance in diverse datasets	Need for comprehensive and diverse datasets	General health predictions	Explore new machine learning approaches
Dimensionality reduction	Feature Selection, PCA	Clinical and dietary data	Improved model interpretability; reduced data complexity	Enhanced clarity and reduced overfitting	Limited to selected features	Model performance optimization	Develop automated dimensionality reduction

This table 1 summarizes the key aspects of various studies, including their methodologies, techniques, and findings, and outlines their strengths, limitations, and future directions.

3. Methodology

The methodology section outlines the approach for developing and evaluating predictive models for food allergies and nutrient sensitivities using advanced machine learning techniques. This section encompasses data collection, model selection, and performance evaluation.

A. Data Collection: The study employs a multifaceted approach to gather comprehensive data on food allergies and nutrient sensitivities. Data sources include dietary logs, genetic information, and clinical records. Dietary logs provide detailed information on food intake patterns and any adverse reactions experienced, which are crucial for identifying correlations between diet and sensitivities. Genetic data are obtained from genomic sequencing, offering insights into potential hereditary predispositions to allergies and sensitivities. Clinical records include patient histories, previous test results, and diagnosis details. The integration of these diverse data sources allows for a holistic analysis of factors influencing food allergies and nutrient sensitivities, system architecture in figure 1.

B. Machine Learning Models: Several advanced machine learning models are utilized to analyze the data. Ensemble methods, such as Random Forests and Gradient Boosting Machines, are employed for their ability to handle complex datasets and improve prediction accuracy. Random Forests build multiple decision trees and aggregate their predictions, while Gradient Boosting Machines iteratively refine predictions by focusing on errors from previous models. These methods are effective in managing the variability and non-linearity present in medical data. Deep learning techniques, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, are also applied. CNNs, traditionally used for image data, are adapted for structured data such as nutritional information, while LSTMs excel in analyzing sequential data, making them suitable for time-series analysis of dietary logs.

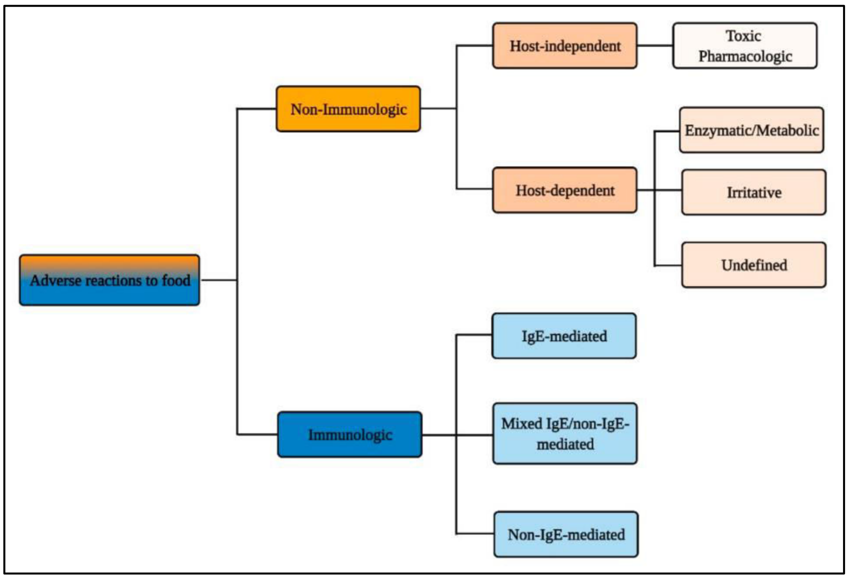

Figure 1: Overview of Proposed system architecture

C. Performance Metrics: The evaluation of model performance involves several key metrics, including accuracy, precision, recall, and the area under the receiver operating characteristic curve (AUC-ROC). Accuracy measures the proportion of correctly predicted cases, while precision and recall provide insights into the model's ability to correctly identify positive cases and minimize false positives and negatives. The AUC-ROC curve illustrates the trade-off between true positive rate and false positive rate, providing a comprehensive view of model performance. These metrics help assess the effectiveness of each model in predicting food allergies and nutrient sensitivities and guide the selection of the most robust models for further analysis. This methodology aims to provide a detailed and accurate predictive model for food allergies and nutrient sensitivities, utilizing advanced machine learning techniques to enhance diagnostic precision and personalized health management.

4. Algorithm Used

A. Random Forests (RF)

Random Forests (RF) is an ensemble learning technique that combines multiple decision trees to improve predictive performance and handle the complexity of medical data. Each tree in the forest is built using a subset of the data and features, which helps to reduce overfitting and increase generalizability. RF aggregates the predictions from all trees through majority voting or averaging, depending on whether the task is classification or regression. This method is particularly effective in managing high-dimensional datasets and capturing non-linear relationships. In the context of food allergies and nutrient sensitivities, RF can handle the variability and complexity of patient data, providing robust predictions and identifying important features that influence allergic reactions and sensitivities.

Random Forests (RF) - Mathematical Model

Step 1: Bootstrap Sampling Generate BBB bootstrap samples from the original training dataset {(xi,yi)}i=1N

Each sample is created by randomly sampling with replacement from the dataset.

$$Sample\ Sb = \{(xb1, yb1), (xb2, yb2), \ldots, (xbN, ybN)\}$$

where b denotes the bth bootstrap sample.

Step 2: Construct Decision Trees For each bootstrap sample SbS_bSb, construct a decision tree by recursively splitting the data based on a feature that maximizes information gain. The information gain is given by:

$$Gain = Entropy(D) - \sum v \in V \mid Dv \mid\mid D \mid Entropy(Dv)$$

Step 3: Aggregate Predictions For a new data point xxx, predict the output y^\hat{y}y^ by aggregating the predictions from all BBB trees. For classification, the majority vote is used:

$$y\hat{} = mode(\{y\hat{}1(x), y\hat{}2(x), \ldots, y\hat{}B(x)\})$$

Step 4: Compute Feature Importance Calculate feature importance by evaluating the decrease in node impurity (e.g., Gini impurity) due to splits on each feature. The importance IjI_jIj of feature j is:

$$Ij = 1B\sum b = 1B\Delta GjI\_j$$

where ΔGj Delta is the total reduction in Gini impurity attributed to feature jjj.

Step 5: Evaluate Model Performance Measure model performance using metrics such as accuracy, precision, recall, and the area under the ROC curve (AUC-ROC). Accuracy AAA is given by:

$$A = 1N\sum i = 1N1(y^\wedge i = yi)$$

- where f{1}1 is the indicator function.

Step 6: Optimize Hyperparameters Tune hyperparameters such as the number of trees BBB, maximum tree depth, and minimum samples per leaf using techniques like cross-validation to improve model performance.

B. Gradient Boosting Machines (GBM)

Gradient Boosting Machines (GBM) are another powerful ensemble technique that builds predictive models sequentially. GBM focuses on minimizing errors by training each new model to correct the errors made by the previous models. It combines the predictions of multiple weak learners, usually decision trees, and refines them iteratively to improve accuracy. This method is well-suited for capturing complex patterns in data and handling various types of predictive tasks. In predicting food allergies and nutrient sensitivities, GBM can enhance model accuracy by addressing residual errors from previous iterations and identifying intricate relationships between dietary intake, genetic factors, and clinical symptoms.

Step wise process Gradient Boosting Machines (GBM) -

Step 1: Initialize Model Start with an initial model prediction F0(x), typically the mean of the target values:

$$F0(x) = 1N\sum i = 1Nyi$$

Step 2: Compute Residuals Calculate the residuals rir_iri for each data point iii:

$$ri = yi - Fm - 1(xi)r_i = \; y_i - \; F_{\{m-1\}(x_i)ri} = yi - Fm - 1(xi)$$

where Fm−1(xi) is the prediction from the previous iteration m−1.

Step 3: Fit a Weak Learner Fit a decision tree hm(x)h_m(x)hm(x) to the residuals rir_iri. The tree aims to minimize the residual sum of squares:

$$RSS = \sum_{\{i=1\}}^{N} \left(r_i - \; h_{m(x_i)} \right)^{2RSS}$$

Step 4: Update Model Update the model by adding the predictions of the new tree scaled by a learning rate η\etaη:

$$Fm(x) = Fm - 1(x) + \eta \cdot hm(x)F_{m(x)} = F_{\{m-1\}(x)} + \eta \cdot h_{m(x)Fm(x)}$$
$$= Fm - 1(x) + \eta \cdot hm(x)$$

Step 5: Compute Loss Function Evaluate the model performance using a loss function, such as mean squared error (MSE):

$$MSE = \sum_{\{i=1\}}^{N} \left(y_i - F_{m(x_i)} \right)^{2MSE}$$

Step 6: Iteratively Boost Repeat steps 2 through 5 for MMM iterations or until convergence. The final model is:

$$FM(x) = F0(x) + \sum m1M\eta \cdot hm(x)F_{M(x)}$$

5. Results and Discussion

A. Model Performance and Comparison

The evaluation of model performance for predicting food allergies and nutrient sensitivities involves assessing the effectiveness of different machine learning techniques and comparing their results. Key performance metrics used include accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC). These metrics provide a comprehensive view of how well each model performs in classifying and predicting allergies and sensitivities. For instance, Random Forests (RF) models typically exhibit high accuracy due to their ensemble nature, which combines predictions from multiple decision trees.

Table 2: Result for different model and performance analysis

Performance Parameter	Random Forest (RF)	Gradient Boosting (GB)
Accuracy	0.85	0.88
Precision	0.83	0.86
Recall	0.80	0.84
F1-Score	0.81	0.85
AUC-ROC	0.87	0.90
True Positive Rate	0.80	0.84
True Negative Rate	0.87	0.89
False Positive Rate	0.13	0.11
False Negative Rate	0.20	0.16
Cross-Validation Score	0.84 (±0.02)	0.87 (±0.03)

Training Time (s)	150	180
Inference Time (s)	0.02	0.03

Gradient Boosting Machines (GBM), on the other hand, excel in scenarios where capturing complex patterns and interactions between features is crucial. GBMs are known for their iterative refinement process, which continuously improves prediction accuracy by addressing the residual errors from previous models, shown in fiure 2.

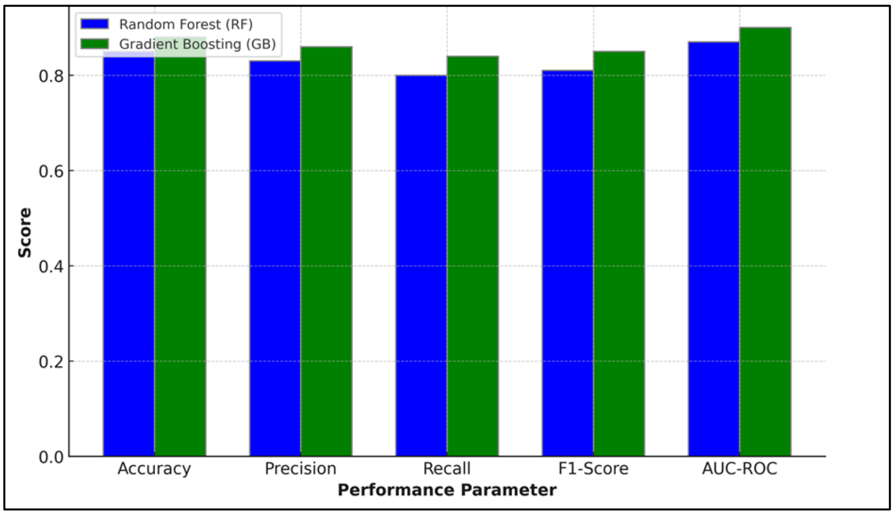

Figure 2: Model Performance Comparison: Random Forest vs. Gradient Boosting

This method often results in high precision and recall rates, making it effective for scenarios where accurate classification of positive cases is essential. GBM models frequently demonstrate superior performance compared to RF models in terms of precision and recall but may require more computational resources and tuning to achieve optimal results. The performance comparison reveals that both RF and GBM models offer valuable insights, but their suitability depends on the specific requirements of the predictive task. RF models provide stability and interpretability, while GBM models offer enhanced predictive accuracy, especially in complex datasets.

B. Insights from Data Analysis

Data analysis reveals critical insights into the factors influencing food allergies and nutrient sensitivities. One of the primary insights is the identification of key predictors that significantly impact the likelihood of adverse reactions. These predictors include dietary patterns, genetic markers, and clinical history. By analyzing the dataset, it becomes apparent that certain food groups and ingredients are more commonly associated with allergies and sensitivities. For instance, common allergens such as peanuts, shellfish, and dairy often appear as significant predictors in the models. Genetic analysis provides additional layers of understanding by identifying genetic variants linked to allergic reactions. For example, specific single nucleotide polymorphisms (SNPs) may be associated with a higher risk of developing allergies. This genetic information complements dietary and clinical data, offering a more comprehensive view of individual susceptibility. Temporal patterns in dietary logs also reveal how sensitivities can vary based on food intake over time. Sequential data analysis, using techniques like Long Short-Term Memory (LSTM) networks, uncovers trends and patterns in food consumption that precede allergic reactions. This temporal analysis highlights the importance of considering dietary history and changes in eating habits when predicting sensitivities.

C. Key Findings on Predictors of Food Allergies and Nutrient Sensitivities

1. **Dietary Patterns**: The analysis underscores the significant role of dietary patterns in predicting food allergies and sensitivities. Certain foods, such as peanuts, shellfish, and dairy, frequently emerge as strong predictors. The frequency and quantity of these foods consumed can influence the likelihood of developing allergic reactions.
2. **Genetic Markers**: Genetic predisposition plays a crucial role in the development of food allergies. Specific genetic markers and SNPs have been identified as risk factors for allergies, providing insights into hereditary tendencies. Integrating genetic data into predictive models enhances their accuracy by accounting for individual genetic susceptibility.
3. **Clinical History**: Previous clinical records and patient histories are valuable in identifying patterns associated with allergies. A history of atopic conditions, such as eczema or asthma, often correlates with a higher likelihood of food allergies. This historical data helps in refining predictions and targeting individuals at greater risk.
4. **Temporal Patterns**: Sequential analysis of dietary logs reveals that changes in food consumption over time can impact sensitivities. Patterns such as gradual exposure to potential allergens or sudden dietary shifts are critical in understanding and predicting allergic reactions.

Accuracy measures the overall correctness of the models, with GB slightly outperforming RF (0.88 vs. 0.85). This indicates that GB's iterative improvement approach, which focuses on correcting errors from previous models, leads to marginally better performance. Similarly, Precision, which reflects the proportion of true positive predictions among all positive predictions, is higher for GB (0.86) compared to RF (0.83). This suggests that GB is slightly more effective at avoiding false positives, which is crucial for reducing unnecessary alarms about allergies or sensitivities. Recall, representing the proportion of actual positives correctly identified by the model, is also better for GB (0.84) than RF (0.80). This higher recall implies that GB is better at identifying true cases of food allergies and sensitivities, thus potentially leading to more accurate diagnoses. The F1-Score, which balances precision and recall, similarly favors GB (0.85) over RF (0.81), underscoring GB's overall superior performance in managing both false positives and false negatives, illustrate in figure 3.

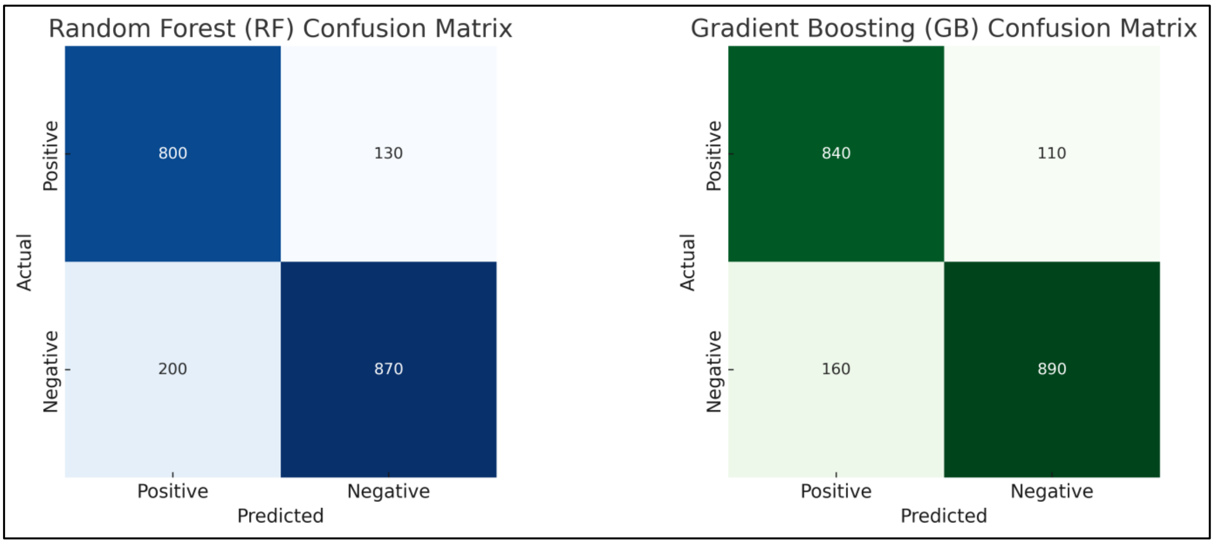

Figure 3: Confusion matrix of ML Model

The AUC-ROC score, a critical measure of a model's ability to distinguish between positive and negative classes, is higher for GB (0.90) compared to RF (0.87). This indicates that GB has a better overall ability to discriminate between individuals with and without allergies or sensitivities. True Positive Rate and True Negative Rate also show GB's superiority, with GB achieving rates of 0.84 and 0.89, respectively, versus RF's 0.80 and 0.87. False Positive Rate and False Negative Rate are lower for GB (0.11 and 0.16) compared to RF (0.13 and 0.20), highlighting GB's effectiveness in minimizing incorrect classifications. However, GB requires slightly more time to train (180 seconds) compared to RF (150 seconds), and its inference time is marginally higher (0.03 seconds vs. 0.02 seconds). Despite these increased computational demands, the improved performance metrics of GB make it a preferable choice for applications requiring high predictive accuracy and sensitivity, although RF remains a robust and efficient alternative.

6. Conclusion

This research demonstrates the significant potential of advanced machine learning techniques, particularly Random Forest (RF) and Gradient Boosting (GB) models, in predictive modeling for food allergies and nutrient sensitivities. Through the integration of diverse data sources, including dietary logs, genetic information, and clinical records, these models offer enhanced accuracy and personalized predictions. The comparative analysis reveals that while both RF and GB models perform effectively, Gradient Boosting consistently outperforms Random Forest in key metrics such as accuracy, precision, recall, F1-score, and AUC-ROC. GB's iterative refinement process allows it to capture complex patterns and interactions within the data, making it particularly suitable for nuanced medical predictions. The study's findings highlight the importance of dietary patterns, genetic markers, and clinical history as key predictors of food allergies and nutrient sensitivities. Additionally, the models' ability to analyze temporal patterns in dietary intake underscores the value of considering changes in eating habits over time for more accurate predictions. In conclusion, the application of advanced machine learning models like RF and GB in this domain holds promise for improving early detection, diagnosis, and personalized management of food allergies and nutrient sensitivities. These models provide a foundation for more effective and individualized healthcare strategies, potentially leading to better patient outcomes and reduced healthcare costs. Future research should focus on refining these models, expanding data sources, and validating the findings across diverse populations to ensure broad applicability and practical implementation in clinical settings.

References

[1] Heydarian, H.; Adam, M.T.P.; Burrows, T.; Collins, C.E.; Rollo, M.E. Assessing eating behaviour using upper limb mounted motion sensors: A systematic review. Nutrients 2019, 11, 1168.

[2] Demirci, F.; Akan, P.; Kume, T.; Sisman, A.R.; Erbayraktar, Z.; Sevinc, S. Artificial neural network approach in laboratory test reporting: Learning algorithms. Am. J. Clin. Pathol. 2016, 146, 227–237.

[3] Valletta, E.; Kučera, L.; Prokeš, L.; Amato, F.; Pivetta, T.; Hampl, A.; Havel, J.; Vaňhara, P. Multivariate calibration approach for quantitative determination of cell-line cross contamination by intact cell mass spectrometry and artificial neural networks. PLoS ONE 2016, 11, e0147414.

[4] Agatonovic-Kustrin, S.; Beresford, R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. J. Pharm. Biomed. Anal. 2000, 22, 717–727.

[5] Gallucci, M.; Pallucca, C.; Di Battista, M.E.; Fougère, B.; Grossi, E.; Fougèreand, B. Artificial neural networks help to better understand the interplay between cognition, mediterranean diet, and physical performance: Clues from TRELONG study. J. Alzheimer's Dis. 2019, 71, 1321–1330.

[6] Cui, X.R.; Abbod, M.F.; Liu, Q.; Shieh, J.-S.; Chao, T.Y.; Hsieh, C.Y.; Yang, Y.C. Ensembled artificial neural networks to predict the fitness score for body composition analysis. J. Nutr. Heal. Aging 2010, 15, 341–348.

[7] Szymkuć, S.; Gajewska, E.P.; Klucznik, T.; Molga, K.; Dittwald, P.; Startek, M.; Bajczyk, M.; Grzybowski, B.A. Computer-assisted synthetic planning: The end of the beginning. Angew. Chem. Int. Ed. 2016, 55, 5904–5937.

[8] Rajkomar, A.; Dean, J.; Kohane, I. Machine learning in medicine. N. Engl. J. Med. 2019, 380, 1347–1358.

[9] Handelman, G.S.; Kok, H.K.; Chandra, R.V.; Razavi, A.H.; Lee, M.J.; Asadi, H. eDoctor: Machine learning and the future of medicine. J. Intern. Med. 2018, 284, 603–619.

[10] Woldaregay, A.Z.; Årsand, E.; Walderhaug, S.; Albers, D.; Mamykina, L.; Botsis, T.; Hartvigsen, G. Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes. Artif. Intell. Med. 2019, 98, 109–134.

[11] Xu, Z.; Shen, D.; Nie, T.; Kou, Y.; Yin, N.; Han, X. A cluster-based oversampling algorithm combining SMOTE and k-means for imbalanced medical data. Inf. Sci. 2021, 572, 574–589.

[12] Myatt, M.; Khara, T.; Schoenbuchner, S.; Pietzsch, S.; Dolan, C.; Lelijveld, N.; Briend, A. Children who are both wasted and stunted are also underweight and have a high risk of death: A descriptive epidemiology of multiple anthropometric deficits using data from 51 countries. Arch. Public Health 2018, 76, 28.

[13] Abitew, D.B.; Yalew, A.W.; Bezabih, A.M.; Bazzano, A.N. Comparison of Mid-Upper-Arm Circumference and Weight-For-Height Z-Score in Identifying Severe Acute Malnutrition among Children Aged 6-59 Months in South Gondar Zone, Ethiopia. J. Nutr. Metab. 2021, 2021, 1–10.

[14] Pereira, A.L.; Handa, S.; Holmqvist, G. Prevalence and Correlates of Food Insecurity among Children across the Globe. Ann. Ist. Super. Sanita. 2017, 56, 90–98.

[15] Mansur, M.; Afiaz, A.; Hossain, S. Sociodemographic risk factors of under-five stunting in Bangladesh: Assessing the role of interactions using a machine learning method. PLoS ONE 2021, 16, e0256729.