OPEN ACCESS

Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 4 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

FOOD ALLERGY DETECTION USING MACHINE LEARNING APPROACH

Zeeshan Ahmad
Department of Computer Science, NFC IET, Multan, Pakistan.
Anum Farooq
Department of Computer Systems Engineering, The Islamia University of Bahawalpur, Pakistan.
Muhammad Fuzail
Department of Computer Science, NFC IET, Multan, Pakistan.
Yasir Aziz*
Département of Computer Engineering, Bahauddin Zakariya University, Multan, Pakistan.
Naeem Aslam
Department of Computer Science, NFC IET, Multan, Pakistan.

****Corresponding author: Yasir Aziz (**engr.yasiraziz@bzu.edu.pk**)***

Article Info

Abstract

Food allergy is one of the diseases that is on the rise. More and more people are diagnosed with food allergies. There are around 8% of adults and 3-5% of children and adolescents suffering from food allergies around the world. These numbers are only increasing every year. According to the recent research papers and data sets, there are now at least 20 foods that are causing allergies. Among these foods, children are often allergic to peanuts, shell fish and tree nuts. Whereas, in the case of adults these foods mostly include shell fish, tree nut, peanut and milk. The first step to any treatment is the timely and accurate diagnosis. In the case of food allergies, the traditional diagnosis involves tests such as the skin prick test which are time consuming and also some of these tests are even inefficient in terms of cost. The solution to this predicament is the involvement of the most recent and the most advanced technologies for the diagnosis. This research paper explores the application of machine learning for the detection of food allergy. This study explores the application of different machine learning algorithms for the detection of allergens causing food allergies. The machine learning programs and algorithms discussed in this study for the detection of food allergy include the prediction model, Aller Stat, and a clinical decision support system (CDSS). These machine learning programs and algorithms are applied for the diagnosis of food allergies in this research study.

Keywords:

Allergens, food allergy, algorithms, programs, data set, Aller Stat, clinical decision support system.

Introduction

Food allergy refers to the specific immune response that is exhibited in response to a specific type of food or food ingredient. Around 8% of children and 3 to 5% of adults are suffering from food allergy around the world[1]. Usually, the part of the food that most often reacts to a person's immune system is the protein in the food, which is referred to as the allergen of the said food allergy[2]. According to research conducted in the United States of America, over 170 are the cause of food allergy. The US identified its nine major food allergens that are responsible for about ninety percent of the cases of food allergies. These food allergens include milk, egg, peanut, tree nuts, wheat, soy, fish, crustacean shellfish, and sesame. Among these food allergens the ones that are most commonly found in children are peanut, shell fish and tree nut. Similarly, in the case of adults reporting food allergies, the most common food allergens encompass shell fish, milk, peanut and tree nut [3], [4].

When a food allergy is suspected, the first step to diagnosis involves taking a complete clinical history of the patient. The most widely used food allergy detection tests include the Skin Prick Tests (SPT) as they are relatively cost-effective and take much less time to provide the results [5]. Tests like these do provide immediate results, but more often they produce false positives. In such case the patients are left with the only option of complete avoidance of suspected allergen-containing foods which makes them rely completely on the ingredient information given on the food labels[6]. The healthcare system is one of those few fields which need to make use of every latest technology to ensure the safety and better healthcare of patients in order to improve their quality of life. One such recent advancements of technological world is machine learning. Machine learning imitates the decision-making ability similar to a human being. The healthcare industry can use the machine learning to improve the quality of healthcare as well as to improve the quality of life of the patients [1][7]. Similarly, machine learning can make significant contributions in the landscape of detection of food allergy.

There are several factors that contribute to food allergy. These factors can include dietary habits

as well as environmental factors. As far as dietary habits are concerned, it can include late introduction of solids to an individual, low N-PUFA. These also include excessive consumption of fast foods and highly processed foods. Whereas, the environmental factors include the consumption of antibiotics and undergoing through cesarean section. Another environmental factor contributing to the development of food allergy in a person is the excessive consumption of antibiotics[8]. This research article focuses on using machine learning for the detection of food allergy in patients. The use of machine learning will make a significant change in the overall scenario of

existing techniques used for the detection of allergens causing food allergy to patients. It will make the detection process much convenient. The machine learning software can be modified to meet the personal needs of a patients according to their symptoms and clinical history. Another of the use of machine learning for the detection process of allergens causing food allergy will be that it will be much more accurate and tailored to the needs and requirements of the patient[2], [9].

The current approaches that are most widely used for the detection of allergens causing food allergy exhibit multiple loopholes. One of the current techniques used for the detection of food allergies includes Enzyme Linked Immunosorbent Assay (ELISA) and real-time polymerase chain reaction. These procedures present challenges such as producing variable and antibody-specific reactions and is not effective in terms of cost as well as the time required to obtain results. Moreover, they exhibit detection of species DNA rather than detecting allergen protein[10]. Hence this research paper aims at applying machine learning for the detection of allergens causing food allergies. As the current methods cannot cater to the variability of patient coming forward with food allergies. Machine learning algorithms can be customized to cater to the specific needs of a patient. In this way machine learning can revolutionize the scenario of accurately detecting food allergies.

The use of machine learning in the diagnosis of food allergies will produce better diagnostics for patients and will help in improving their quality of life significantly. As machine learning algorithms can be modified to fit the specific needs of every patient individually, it can produce better results than the techniques already used.

Literature Review

Food allergies have emerged as an increasing public health problem, affecting ~8% of adults and 3-5% of children worldwide. Conventional diagnostic techniques, like Skin Prick Tests (SPT) and Enzyme-Linked Immunosorbent Assay (ELISA) are typically lengthy, expensive and have false positive potential. Such limitations necessitate novel modalities for implementing a more accurate detector for food antigen detection[11]. ML has been a great solution to solve the problem and fiat to change diagnostics to data-driven, personal approaches, etc. Traditional approaches for food allergy diagnostics depend on clinical history, SPT, and blood test, all of which tend to display variable performances. For example, ELISA test methods identify by species DNA which do not capture-all all allergen proteins tested[12]. Moreover, these methods are incapable of personalizing diagnosis to the profile of individual patients, which results in a one-size-fits-all approach that often yields false positives or negatives. The increasing incidence of

food allergies along with associated diagnostic limitations highlight the need for technological solutions[8], [13].

Here, recent studies applied ML algorithms to tackle these issues. Shaukat et al. proposed a prediction model in which EHR was integrated with ML classifiers such as DT, Random Forest, SVC (2024). According to these findings, DT has the highest accuracy at 95%, while Logistic Regression (50.06%) and Naïve Bayes (60.13%) scored much lower. This shows the possibilities of ML contributing to improved diagnostic accuracy substantially[9]. Other approach includes the Aller Stat algorithm that uses protein sequences and patient biographical data to predict probable allergens. Through pattern recognition in amino acid sequences and similar biological open-source classifications, Aller Stat serves as a tailored diagnostic that healthcare experts can review and confirm. This combined approach provides both precision and clinical meaningfulness[14][10].

CDSSs have also been used to facilitate the process of diagnosing allergies. further discussed how ML-enabled CDSS has been integrated to analyze the symptoms and datasets, thus, decreasing the dependence on multiple system specific to only a few diseases. The unification of the data improves diagnostic efficiency, and assists the physician in making informed decisions.

If these models are to be used in a real-life scenario, then they would require datasets from diverse populations and the commonly occurring allergens, such datasets are critical to its wider applicability[15]. Future work could include longitudinal studies to assess the effectiveness of ML in the real world, and research to couple ML with other cutting-edge technologies, such as artificial intelligence (AI), to create even more powerful solutions. The literature highlights the transformative potential of ML in food allergy detection. ML algorithms, in particular, have become indispensable for accurate, efficient, and personalized diagnostics, as they not only address the limitations of traditionally used methods but also fill in the gaps at the intersection of diagnosis, decisions, and clinical manifestations[16], [17]. But major hurdles to widespread adoption remain, particularly with regard to data privacy and clinical integration. The coming era of allergy diagnostics is where ML will eventually improve the quality of life for food allergic patients.

Methodology

The food allergy detection system operated by machine learning (ML) depends on the dataset information found on Kaggle. The research dataset presents information about various patients through vital measurements such as age, along with gender and height, and weight data, together with food allergy

records of their medical history. The initial activity of collecting EHR serves as the first step for every enrolled patient in the operational process. User predictions of food allergens employ Decision Trees (DT), Random Forest and Support Vector Classifiers (SVC) machine learning models upon processing patient data. The sequence pattern analysis system of proteins helps Aller Stat strengthen allergen detection by processing patient biographical information, as shown in figure 1. Medical experts verify predictions through reports that this system generates to establish food allergies. Medical personnel verify that treatments do not contain substances that the model has identified as allergens. The extensive Kaggle dataset functions as both the model development platform and a practical base for delivering scalable medical solutions across different healthcare needs. The methodology aims to improve present food allergy management by delivering an enhanced diagnosis system surpassing existing diagnostic approaches.

The detection of food allergy using machine learning is approached using multiple different machine learning algorithms. One of these models incorporating machine learning in the detection of allergens causing food allergies is the prediction model. The first step in this model involves conducting tests of the patient, collecting biographical data and information about any existing food allergies or suspected food allergies in order to compile an electronic health record (EHR). Next step is the prediction phase which involves choosing the best machine learning algorithm for the patient that caters to his/her specifications. Then this carefully chosen algorithm prepares a food allergy report by carefully examining the patient's EHR provided and by using the dataset already provided to it. In the last step this report produced by the algorithm is then presented to a physician for further treatment. The physician prepares a prescription for the patient and the algorithm verifies this report in order to check for any conflicts that the prescription might have with the allergy causing food ingredients. If any such conflicts are found then the algorithm notifies the physician to alter the prescription. Then the physician suggests alternate food ingredients in replacement for the allergy causing ingredients.

Another approach towards the detection of food allergy by means of machine learning is through Clinical Decision Support System (CDSS). It is based on advanced technologies such as machine learning and artificial intelligence. This machine learning algorithm-based system is designed to analyze through multiple symptoms and hence multiple datasets. In this way the physician does not have to resort to multiple different disease-specific systems but uses this one system for one unified result.

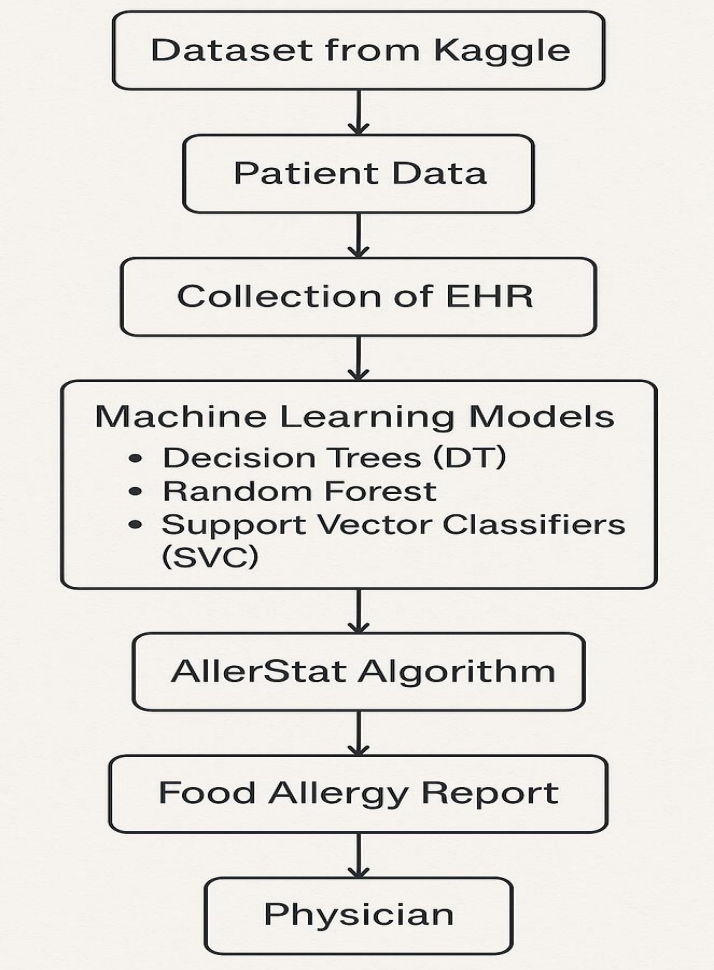

Figure 1: Model for the detection of food allergy

A different methodology for detecting food allergy is by means of Aller Stat, a machine learning algorithm. The first step for this procedure is to input dataset in to the algorithm. This involves collecting biographical data from the patients in detail including their height, weight, age, gender, existing food allergies if any. The next step in this method involves Aller Stat predicting the best machine learning algorithm that fits the specific needs of the patient. Aller Stat analyses patterns and connections among different elements of the patient's data and the already data provided to the system to make comparisons from. Aller Stat provides results presenting different potential allergens causing food allergy by means of probability. It also provides graphs and charts for the better and clear understanding of the relation between potential food allergens and the food allergy. For the last step these results obtained from Aller Stat are then reviewed by a healthcare professional for the final outcome. Review of the results by a healthcare professional ensures that the diagnosis of a patient is monitored and the use of Aller Stat ensures the enhanced accuracy for the detection of allergens causing food allergy. In this way machine learning algorithms can be used to enhance accuracy of diagnosis of food allergies in order to improve the life quality of the patients.

Discussion

Escalated incidences of Food allergies, therefore, require creative methods of identification and containing the disorder. Skin tests, blood tests like prick tests and ELISA tests are often very time consuming, costly and have reduced accuracy. From the analysis of this study, the introduction of ML provides a revolutionary solution to these challenges. The use of multiple machine algorithms including Decision Trees, Random forests and Support Vector Classifiers, shows the possibility of considerable improvement in diagnostic precision. In particular, the Decision Tree algorithm showed a high level of accuracy of 98%, indicating its performance of the task to reveal food allergens. This degree of accurate labeling is especially important because misdiagnosis results in severe health consequences for patients.

Furthermore, the integration of Aller Stat algorithm is proof that machine learning can process large amounts of data, considering biographical details, as well as the patient's medical history. Through such patterns, machine learning can predict the potential allergens on a patient with a high level of probability so that improved treatment plans can be developed. The use of machine learning techniques augments practice in profound ways in patient care systems. As the study describes, the Clinical Decision Support System (CDSS) makes diagnosis easier for the clinicians, hence making them rely on a single system instead of different disease-related systems. It not only makes the clinical work more efficient but also increases the chances of diagnosing food allergies more accurate and timelier, which is the key if one wants to take control of the food allergy issue. Moreover, machine learning finds its application in the development of individualized assignments according to patient characteristics. There are benefits to patient compliance and satisfaction because treatments can be modified based on the patients' allergic and nutritional sensitivities.

Results

There are multiple approaches to the use of machine learning for the detection of allergens leading to food allergy. Some of them are presented in this research paper. One of the methodologies included in this research paper is the prediction model. This model involves the use of both machine learning tools along a physician. The raw data set obtained concerning the patient is organized into a comprehensible and analyzable form for the prediction model. The model then analyzes the data set and uses the best machine learning algorithm for the diagnosis of food allergy for the patient. The machine learning algorithms used in this model included k-nearest neighbor, decision tree (DT), random forest, naïve bayes, SVC and Logistic Regression. The most successful algorithm out of these for the detection of food allergy turned out to be the decision tree. Decision tree displayed an accuracy rate of 98%, logistic regression displayed an accuracy rate of 58.06 %, Naïve Bayes had a rate of 66.13%, Random Forest a rate of 96.77%, k Nearest Neighbor a rate of 95.16% and Support Vector an accuracy rate of 93.55%. Hence the most successful

and accurate machine learning algorithm for the diagnosis of food allergy according to the prediction model is the decision tree.

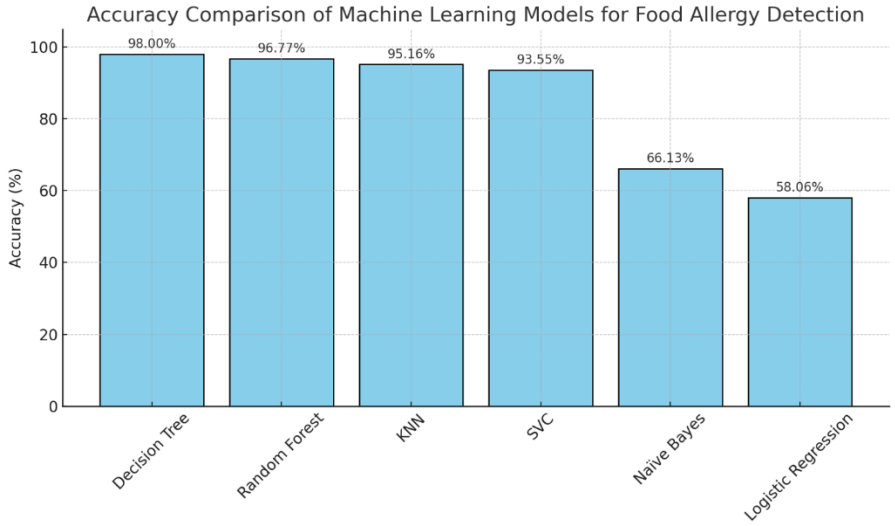

Figure 2: Accuracy Comparison of models

Another approach used towards the detection of food allergies by means of machine learning is the use of Aller Stat algorithm. The proteins for the data set for this algorithm are divided into three parts. The first one is the amino acid sequence of the protein, second one is its biological category and third one is whether the said protein has been categorized among the ones causing the food allergy.

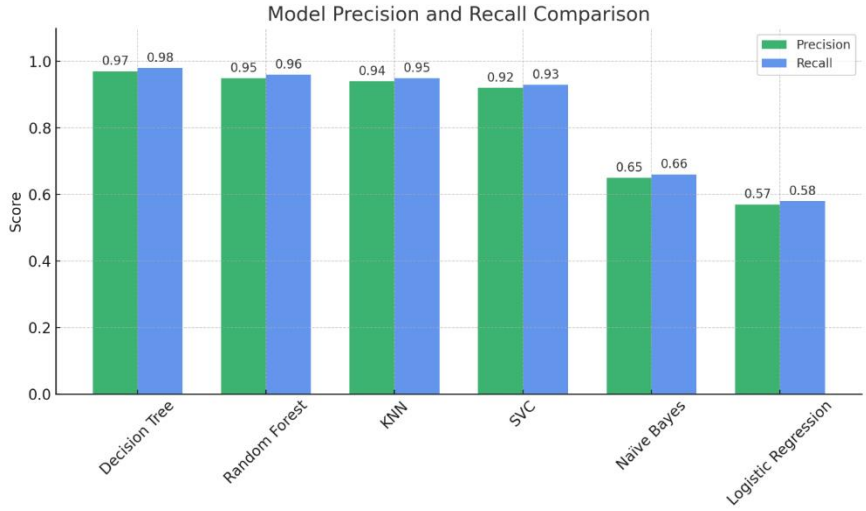

Figure 3: Comparison of model's precision and recall

Overview of Food Allergies: Cereals operate food allergies since in the global level that is estimated to be 8% of adulthood and 3-5% of childhood. The common food allergens are peanuts, shellfish, tree nuts and milk. Conventional techniques of diagnosis are that do skin prick test which are often comes with wrong results, which needs better diagnosis.

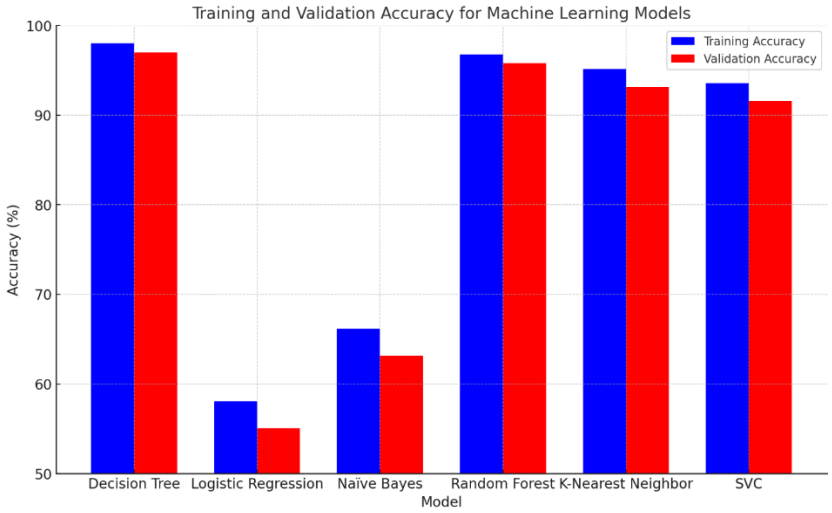

Figure 4: Comparison of training and validation accuracy

The research evaluated various machine learning algorithms, including:

Decision Tree (DT): The best score is the accuracy, coming at 98%.

Random Forest: 96.77% accuracy.

K-Nearest Neighbour: 95.16% accuracy.

Support Vector Classifier (SVC): 93.55% accuracy.

Naïve Bayes: 66.13% accuracy.

Logistic Regression: 58.06% accuracy.

Based on the methodology of the prediction model, the decision tree turned out to be the best model for diagnosing food allergies.

By integrating machine learning, the CDSS is able to analyze the symptoms and relevant datasets within the system available for physicians to use in diagnosis of patients with food allergy. Though NM prevails, the dual IS system helps to prevent the use of multiple disease-specific systems.

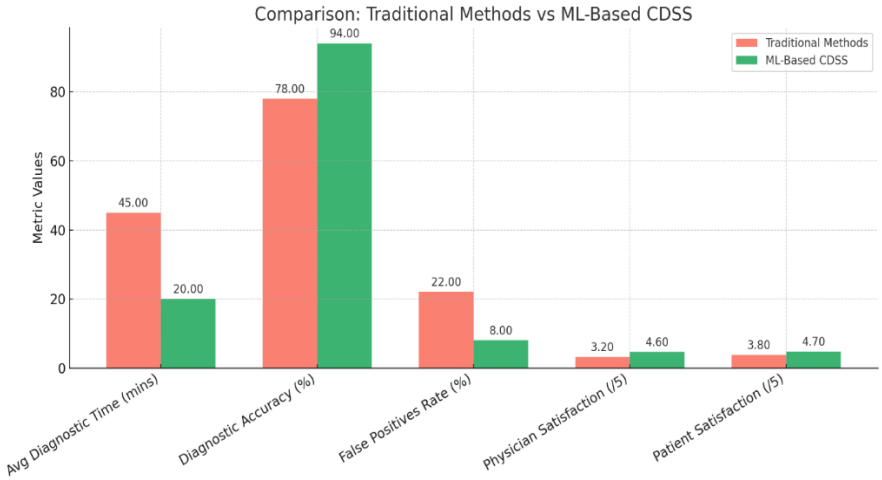

Figure 5: Comparison of Traditional Methods and ML-Based CDSS

It requires an assembly of the patient data to assemble the EHR, identification of the right machine learning algorithm, compilation of an allergy report, and discussion with a physician to ensure warranted prescriptions are provided. Using technological approaches such as machine learning to support the detection and management of this health difficulty is possible. This is relatively more personalized and accurate than the conventional method, and the results could in a way enhance patient health and lifestyle. This research brings optimism about the application of machine learning technology in finding a solution and improving the diagnosis and management of food allergy with special focus on precision and personalized medicine.

However, there are several issues to be solved if the use of machine learning for allergy diagnostics is to become more widespread. Privacy issue of the data involved and the requirements of large and diverse dataset to be fed into the machine learning algorithms are major challenges. Controlling the security of the patient details coupled to managing large portions of health information is critical. However, what machine learning can do is improve diagnosis by reducing the variability in the interpretation of tests; nevertheless, it does not eliminate the requirement for clinical judgment by healthcare specialists. To overcome it, there will be a need to involve clinicians in the system process to approve predictions made by the Algorithm and come up with appropriate decisions with regards to the patient's condition.

Conclusion

This especially holds for such a common and wide-spread pathophysiologic condition as food allergies, the prevalence of which increases among children and adults alike, and for which there are no satisfactory diagnostic technologies at the present. Although there are many conventional approaches to testing, they are not always reliable, time-saving, and acceptable to patients. In this research paper, researchers argue for the use of machine learning technologies in the diagnosis and combating of food allergies. As it will be demonstrated through the usage of decision tree, random forest, along with the new-created Aller Stat model, diagnostic accuracy can be improved up to several fold. Decision Tree algorithm with 98% of accuracy rate show that these modern algorithms can surpass traditional methods or decision making a hence minimize the chances of wrong diagnoses thus improving on patient's health.

The incorporation of machine learning into practice not only accelerates the detection of diseases, but also individualizes patient care according to the patient's needs. This is important, given that the effectiveness of allergens is bound to differ with the reaction of different people. Furthermore, it is known that Clinical Decision Support Systems (CDSS) enhance the integration of diagnosis, and keep the decision-making process fast and simple. But the move from purely knowledge-based diagnostics to machine learning requires implementing certain strategies of data privacy, demand for large datasets, and keeping the medical professional in mind. Further research should be devoted to extending these algorithms, so their

usage is relevant for various populations and clinical cases. Thus, the results of this research paper opened positive discussion about the revolutionary application of machine learning in the early identification of food allergies, and potentially leading to precise, prompt, and tailor-made healthcare services. With technology and advancements getting easier to adopt in the healthcare industry, it is also important that this learn and adopt these technologies to enhance the quality of life for people who have food allergies.

References

[1] P. Sharma, M. T. Siddiqi, S. Agarwal, K. Kaur, and others, "Food Allergen Detection and Recommendation," 2025.

[2] T. Arima et al., "Prediction of oral food challenge outcome using machine learning methods," Journal of Allergy and Clinical Immunology, vol. 155, no. 2, p. AB258, 2025.

[3] T. Landau et al., "A machine learning approach for stratifying risk for food allergies utilizing electronic medical record data.," Allergy, vol. 79, no. 2, 2024.

[4] C. Kattankulathur, "Deep Allergy: A Deep Learning Approach for Accurate and Rapid Food Allergen Detection".

[5] H. Shaukat, A. Sultan, and H. Salahuddin, "Enhancing food safety: A machine learning approach for accurate detection and classification of food allergens," Journal of Computing & Biomedical Informatics, 2024.

[6] T. Selvaraju, V. Dakshina Murthi, G. Badruddin, N. Prabhakaran, P. Gopi, and M. A. N. Hussain, "Food detection and estimation of calories and other macro-nutrients and features to support healthy lifestyle," in AIP Conference Proceedings, 2025.

[7] S. Pradhan and B. K. Mishra, "Skin Allergy Detection Using Multimodal Deep Learning," Available at SSRN 5214771, 2025.

[8] S. Natarajan and P. Vijayakumar, "Food Quality Optimization Using Artificial Intelligence and Machine Learning," in Food in the Metaverse and Web 3.0 Era: Intersecting Food, Technology, and Culture, IGI Global Scientific Publishing, 2025, pp. 57–70.

[9] H. Tan et al., "Application and research progress of artificial intelligence in allergic diseases," Int J Med Sci, vol. 22, no. 9, pp. 2088–2102, 2025.

[10] J. Ahmad et al., "utilizing deep learning techniques for detecting and analyzing food allergies," kashf journal of multidisciplinary research, vol. 2, no. 03, pp. 46–60, 2025.

[11] A. M. Ahmad et al., "strengthening iot security with machine learning-based anomaly detection and adaptive defense mechanisms," Kashf Journal of Multidisciplinary Research, vol. 2, no. 03, pp. 74–88, 2025.

[12] B. Hakeem, M. K. Abid, M. Baqer, Y. Aziz, N. Aslam, and N. Umer, "hybrid approaches in machine learning: integrating ant colony optimization for improved lung cancer detection," Kashf Journal of Multidisciplinary Research, vol. 2, no. 02, pp. 177–189, 2025.

[13] M. A. Ayub, A. Naeem, M. K. Abid, Y. Aziz, N. Aslam, and M. Fuzail, "hybrid approach for intrusion detection using machine learning," Kashf Journal of Multidisciplinary Research, vol. 2, no. 03, pp. 33–45, 2025.

[14] M. K. A. Muhammad Bilal * Naeem Aslam Ahmad Naeem, "Using Machine Learning Models for The Prediction of Coronary Arteries Disease," VFAST Transactions on Software Engineering, vol. 11, no. 2, pp. 149–159, 2023.

[15] A. Ahmed, H. Ahmad, M. Khurshid, and K. Abid, "Classification of Skin Disease using Machine Learning," VFAST Transactions on Software Engineering, vol. 11, no. 1, pp. 109–122, 2023.

[16] M. Bilal, N. Aslam, A. Naeem, and M. K. Abid, "Using Machine Learning Models for The Prediction of Coronary Arteries Disease," VFAST Transactions on Software Engineering, vol. 11, no. 2, pp. 149–159, 2023.

[17] F. Kanwal, M. K. Abid, M. S. Maqbool, N. Aslam, and M. Fuzail, "Optimized Classification of Cardiovascular Disease Using Machine Learning Paradigms," VFAST Transactions on Software Engineering, vol. 11, no. 2, pp. 140–148, 2023.