International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 430-435, November-2023

https://doi.org/10.47001/IRJIET/2023.711058

Anaphylaxis Prediction and Tailored Medicine Prescription System Using CNN-Based Tabular Data Analysis

[1]L.M. Dilanka Matheesha Rajapakse, [2]H.M. Samadhi Chathuranga Rathnayake, [3]Sanjeevi Chandrasiri

[1]Department of Information Technology, Faculty of Computing, Sri Lanka Institute of Information Technology, Sri Lanka
[2,3]Lecturer, Department of Information Technology, Faculty of Computing, Sri Lanka Institute of Information Technology, Sri Lanka

Authors E-mail: [1]it20019518@my.sliit.lk, [2]samadhi.r@sliit.lk, [3]sanji.c@sliit.lk

***Abstract* - Allergic diseases encompass a wide range of conditions in which the immune system reacts abnormally to harmless substances, leading to various symptoms and health complications. Anaphylaxis is a serious life-threatening generalized or systemic hypersensitivity reaction.[1] Mainly it is thought to be a serious systemic hypersensitivity reaction that is usually rapid in onset and may cause death.[2] It is triggered by exposure to specific allergens, such as certain foods, medications, insect stings, or latex.[3] The conventional approach to diagnosing anaphylaxis involves in-person consultations with healthcare professionals, including physicians, allergists, and immunologists.[4] However, this process can be time-consuming, costly, and dependent on the availability of specialized medical expertise. Identifying anaphylaxis as positive or negative is complex due to the similarity of its symptoms with common ailments, requiring the physician's expertise. However, manual identification can cause accuracy issues, leading to incorrect diagnoses and prescriptions. To address these challenges and provide a more efficient disease diagnosis system, this research aims to harness machine learning techniques. Specifically, a CNN-based analysis is employed to predict whether a patient has anaphylaxis. Moreover, the system's functionality extends beyond diagnosis. If anaphylaxis is positively identified, the system initiates the process of recommending the administration of adrenaline. In the case of patients aged less than 12, a specialized mathematical equation is applied to calculate the appropriate dosage of adrenaline based on the patient's age. Conversely, if the system determines a negative anaphylaxis diagnosis, it reevaluates the input symptoms and matches them with suitable specialists. The system maintains a repository of physicians categorized by their areas of expertise, allowing it to output both the specialist's field and the name of the corresponding physician for patient referral.**

Keywords: Machine Learning, CNN, Data preprocessing, allergic diseases, anaphylaxis, machine learning, disease diagnosis, treatment recommendations, healthcare, patient outcomes.

I. INTRODUCTION

Allergic diseases have become a significant global health concern, affecting millions of individuals worldwide. These conditions arise when the immune system overreacts to otherwise harmless substances, leading to a wide range of symptoms and complications. Among these allergic diseases, anaphylaxis stands out as a severe and potentially life-threatening allergic reaction characterized by rapid onset and systemic involvement.[5] The identification and management of allergic diseases, including anaphylaxis, have traditionally relied on in-person consultations with healthcare professionals.[6] However, this approach can be time-consuming, costly, and limited by the availability of specialized expertise.

To address these challenges and improve the diagnosis and treatment of allergic diseases, there is a growing interest in leveraging machine learning techniques. Machine learning offers the potential to develop automated and accurate disease diagnosis systems that can assist in identifying specific allergens and predicting patient outcomes. By analyzing large datasets, machine learning models can identify patterns, correlations, and predictive indicators that may not be immediately apparent to human observers.[7][8] This can enable more efficient and personalized healthcare interventions, leading to better patient outcomes.

This research paper is focused on the development of a disease diagnosis system for anaphylaxis using machine learning approaches. The primary objective is to create a system that can accurately identify anaphylaxis cases based on patient symptoms and provide appropriate treatment recommendations. A dataset obtained from the National

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 430-435, November-2023
https://doi.org/10.47001/IRJIET/2023.711058

Center for Primary Care and Allergy Research, University of Sri Jayewardenepura, which includes symptom profiles of patients diagnosed with anaphylaxis and those without anaphylaxis, will be used to train and evaluate machine learning models to achieve high accuracy in anaphylaxis prediction.

The significance of this research lies in the potential to improve the efficiency and accessibility of healthcare services for individuals with allergic diseases, particularly anaphylaxis. The aim is to reduce the burden on healthcare professionals by implementing an automated disease diagnosis system, and to provide patients with timely and accurate assessments, leading to more effective treatment strategies. Moreover, this research contributes to the broader field of machine learning applications in healthcare, showcasing the potential of data-driven approaches to revolutionize disease management and enhance patient care.

The subsequent sections of this research paper will delve into the methodology, dataset description, machine learning model development, evaluation metrics, and results. Through this comprehensive investigation, the effectiveness and practicality of machine learning-based approaches in diagnosing and managing anaphylaxis will be established, with the goal of improving patient outcomes and healthcare delivery.

II. LITERATURE REVIEW

Allergies and anaphylaxis pose significant health risks [9] and require accurate identification and appropriate treatment.[10] Traditional methods of diagnosing allergies and determining suitable treatments are often time-consuming and subjective. To address these challenges, there is a growing need for more efficient and effective methods of identifying allergens and recommending appropriate treatment options.[13] Advances in technology, such as machine learning algorithms and data analytics, offer promising approaches for developing personalized medicine recommendations for patients with allergies. [7]

Machine learning (ML) has emerged as a powerful tool for medical diagnosis and prediction.[11] [12] ML models can analyze large datasets, identify patterns, and make accurate predictions based on learned patterns. In the context of anaphylaxis prediction, ML models can utilize patient data, including demographics and symptoms, to classify individuals as either positive or negative for anaphylaxis. This classification can assist healthcare professionals in making timely and accurate diagnoses, leading to improved patient outcomes.

When choosing the ML model for anaphylaxis prediction, several factors need to be considered. Firstly, the model should be able to handle categorical and numerical data, as the dataset includes patient characteristics such as age, weight, and gender, along with symptom information. Secondly, since anaphylaxis prediction does not involve image analysis, models specializing in image recognition, such as convolutional neural networks (CNNs), are not suitable for this particular task. Instead, models that can handle tabular data and capture complex relationships between features should be prioritized.

Various ML model types can be considered for anaphylaxis prediction. Logistic regression, decision trees, random forests, support vector machines (SVMs), and neural networks are among the commonly used models. Logistic regression is a popular choice for binary classification tasks, but it may struggle to capture complex nonlinear relationships in the data. Decision trees and random forests excel at capturing nonlinear relationships but may overfit the data or struggle with high-dimensional feature spaces. SVMs are effective for high-dimensional data but can be computationally expensive. Neural networks, particularly deep learning models, have shown exceptional performance in various medical applications due to their ability to capture intricate patterns in the data.

In this research, the CNN model was chosen for anaphylaxis prediction. Although CNNs are primarily used for image analysis, they can also be adapted for tabular data analysis by treating each feature as an input channel. While it may seem unconventional to use CNN for non-image data, it offers several advantages for this research project. CNNs are capable of capturing complex patterns and relationships within the data, even in the absence of image inputs. The hierarchical structure of CNNs allows them to automatically learn relevant features and make accurate predictions. By leveraging the strengths of CNNs, this research aims to achieve high accuracy in anaphylaxis prediction without the need for image analysis.

III. METHODOLGY

A) Data Collection

For this research project, a dataset consisting of patients diagnosed with anaphylaxis was obtained from Dr. Thushali Ranasinghe, a physician working under Prof. Chandima Jeewandara. The dataset was provided in Excel format and contained five columns: patient number, age, weight, gender, and disease symptoms. Each row represented a unique patient. Additionally, a separate dataset containing patients identified as negative for anaphylaxis was obtained. This dataset

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 430-435, November-2023

https://doi.org/10.47001/IRJIET/2023.711058

followed the same structure as the positive dataset, with the additional symptoms column.

B) Choosing the ML Model

The selection of an appropriate ML model is crucial for accurate anaphylaxis prediction. Several ML model types were considered, including logistic regression, decision trees, random forests, SVMs, and neural networks. After evaluating their performance and considering the specific requirements of the research project, a CNN model was chosen.

CNNs, although commonly used for image analysis, can be adapted to process sequential data such as symptom patterns in patients. This adaptability makes CNNs suitable for analyzing the symptom patterns of anaphylaxis. While other ML models may also perform well on this task, CNNs were favored due to their ability to extract relevant features from the input data and capture complex relationships between symptoms and anaphylaxis.

C) Data Preprocessing and Cleansing

Prior to the development of the ML model, extensive data preprocessing and cleansing were performed to prepare the dataset for training. The symptoms present in both the positive and negative anaphylaxis datasets were identified and consolidated into a new Excel file. This new file included the columns: patient number, age, weight, gender, and individual symptom columns. Each symptom was represented as a binary value, with 1 indicating the presence of the symptom and 0 indicating its absence.

Furthermore, a new column named "Label" was added to indicate whether a patient was positive or negative for anaphylaxis. This column was populated with "Positive" for patients diagnosed with anaphylaxis and "Negative" for those identified as negative for anaphylaxis.

The data preprocessing and cleansing steps were performed using Colab, an online platform that provides a Python environment for data analysis and ML model development. This approach ensured efficient and automated processing of the dataset, avoiding manual data manipulation and potential errors.

D) ML Model Development Process for Disease Prediction

Once the dataset was prepared, the development process of the ML model for anaphylaxis prediction was initiated. The CNN model architecture was defined, taking into account the input dimensions, number of classes (positive or negative for anaphylaxis), and the desired complexity of the model.

The dataset was split into training and testing sets, with the majority of the data allocated for training the model and a smaller portion reserved for evaluating its performance. The training set was used to update the model's parameters through a process called back propagation, where the model learns from the input data and adjusts its internal parameters to minimize the prediction error.

During the training process, the model's performance was regularly evaluated on the testing set to monitor its accuracy and detect any signs of overfitting. Overfitting occurs when the model becomes too specialized in the training data and fails to generalize well to new, unseen data. Several techniques, such as regularization and early stopping, were applied to mitigate the risk of overfitting and ensure the model's generalization ability.

The ML model development process involved iterative refinement and parameter tuning to optimize the model's performance. Different hyperparameters, such as learning rate, batch size, and network architecture, were adjusted to achieve the best possible accuracy and robustness.

E) Information and Guidance Gathering Process and Self-Learning

Throughout the research project, valuable guidance and information were obtained from the external supervisors, Prof. Chandima Jeewandara and Dr. Thushali Ranasinghe. Their expertise in the field of anaphylaxis and medical diagnosis provided crucial insights and direction for the research. Regular meetings and discussions were held to address challenges, clarify concepts, and refine the research objectives.

In addition to guidance from the supervisors, self-learning played a significant role in acquiring knowledge and skills related to anaphylaxis, ML model development, and medical prescription processes. Online resources, scientific publications, and relevant literature were extensively studied to understand the underlying principles, techniques, and best practices in these areas.

F) Medicine Prescription Process Development

As part of the research project, the development of a medicine prescription process for anaphylaxis was undertaken. The process involved determining the appropriate medication based on the patient's age, weight, and symptoms.

If the patient's age was above 12 years, physicians prescribed the adult epinephrine auto-injector (EpiPen), which contains 0.5 mg of adrenaline for intramuscular injection. For patients under 12 years of age, the adrenaline dosage was

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 430-435, November-2023

https://doi.org/10.47001/IRJIET/2023.711058

calculated based on their weight. The weight of adrenaline in milligrams was determined by multiplying the patient's age by 0.01.[14] This approach ensured personalized dosage recommendations tailored to the patient's characteristics.

The developed system not only facilitated accurate anaphylaxis prediction but also addressed the scenario where the ML model predicted a negative result for anaphylaxis. In such cases, the system identified the patient's symptoms and provided a suitable physician recommendation based on their specialization. A database of physicians categorized according to their expertise was integrated into the system, allowing it to output the name of a relevant physician along with their specialty, based on the patient's symptoms.

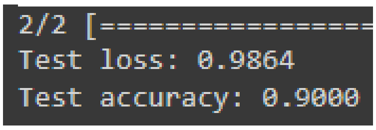

Figure 1: Accuracy Results

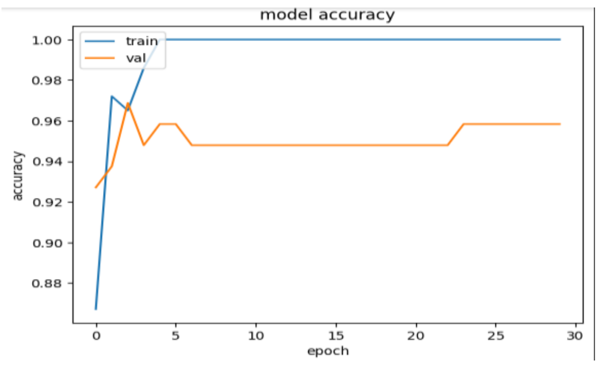

Figure 2: Model performance

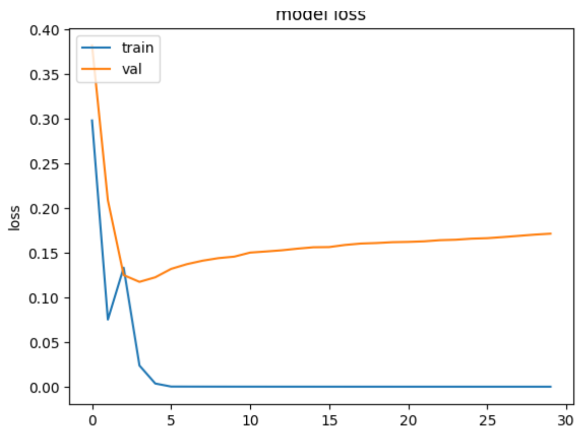

Figure 3: Model performance

G) Physician Recommendation System

The developed physician recommendations system plays a pivotal role in enhancing healthcare accessibility for patients who have been identified as negative for anaphylaxis. By leveraging the power of data analytics and machine learning, this system offers a personalized approach to connecting patients with specialized physicians in Sri Lanka. This approach ensures that patients receive appropriate medical attention tailored to their symptoms and medical needs.

The physician recommendations system relies on a comprehensive categorization of physicians based on their specialization areas. In this physician recommendation system, the output includes not only the physician's specialization but also their name. Based on the identified specialization of the physician, the system will provide a suitable name corresponding to that specialization to the user.

Each symptom is meticulously associated with the medical discipline that best addresses it. This categorization facilitates the precise matching of patients with the most suitable physicians for their condition. Information and availability of the specialized physicians are taken from the government hospitals in Sri Lanka.

The categorization of physicians based on symptoms is as follows:

Table 1: Categorization of physicians based on symptoms

Symptom	Type of the physician
Swelling of Eyes	Dermatologist
Hives	Dermatologist
Itching of Skin	Dermatologist
Swelling of Lips	Immunologist
Sneezing	Respiratory Physician (Pulmonologist)
Wheezing	Respiratory Physician (Pulmonologist)
Difficulty of Breathing	Cardiologist / Respiratory Physician (Pulmonologist)
Vomiting	Gastroenterologist
Fainting	Cardiologist
Abdominal Pain	Gastroenterologist
Diarrhea	Gastroenterologist
Loss of Consciousness	Neurologist
Dizziness	Neurologist
Nasal Congestion	Respiratory Physician (Pulmonologist)
Fever	General Medical Practitioner
Cough	Respiratory Physician (Pulmonologist)
Rhinorrhea	Respiratory Physician (Pulmonologist)
Dry Skin	Dermatologist
Fatigue	General Medical Practitioner
Acid Reflux	Gastroenterologist
Heartburn	Gastroenterologist

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 11, pp 430-435, November-2023

https://doi.org/10.47001/IRJIET/2023.711058

H) Enhancing Healthcare Access

The symptom-based physician recommendations system has a significant impact on healthcare access in Sri Lanka. By linking patients to specialized physicians, the system ensures that individuals receive prompt and accurate medical attention. This approach is especially important for patients who may not be familiar with the specific medical discipline required for their symptoms. Through this system, patients gain access to tailored medical expertise, contributing to improved diagnosis, treatment, and overall healthcare outcomes.

IV. CONCLUSION

In conclusion, this research project focused on developing an ML-based approach for anaphylaxis prediction and medicine prescription. The chosen CNN model demonstrated its effectiveness in accurately classifying patients as positive or negative for anaphylaxis, leveraging the input data comprising patient demographics and symptom patterns. By utilizing the strengths of CNNs in capturing complex relationships, the model achieved high accuracy without requiring image analysis.

The methodology encompassed various stages, including data collection, preprocessing, model selection, and development. The dataset obtained from external supervisors, Prof. Chandima Jeewandara and Dr. Thushali Ranasinghe, played a vital role in training the ML model. Their expertise and guidance contributed significantly to the research project's success.

Furthermore, the research project included the development of a medicine prescription process for anaphylaxis. By considering factors such as age, weight, and symptoms, personalized medication recommendations were generated. This process aimed to provide physicians and patients with accurate and tailored prescriptions, enhancing patient care and treatment outcomes.

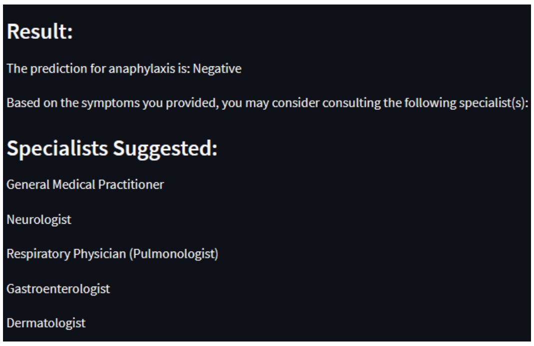

Figure 4: Final Output

Overall, the combination of ML-based anaphylaxis prediction, symptom-based physician recommendation, and personalized medicine prescription exhibited the potential to improve the diagnosis and treatment of anaphylaxis. Further research and refinement of the developed system could contribute to advancements in personalized medicine and healthcare decision-making.

REFERENCES

[1] F. E. R. Simons et al., "World allergy organization guidelines for the assessment and management of anaphylaxis," World Allergy Organization Journal, vol. 4, no. 2, pp. 13-37, 2011.

[2] P. J. Turner et al., "Time to revisit the definition and clinical criteria for anaphylaxis?," World Allergy Organization Journal, vol. 12, no. 10, 2019.

[3] R. Pawankar, G. Canonica, S. Holgate, and R. Lockey, "WAO, white book on allergy Milwaukee (WI)," World Allergy Organization, 2011.

[4] L. Harduar-Morano, M. R. Simon, S. Watkins, and C. Blackmore, "Algorithm for the diagnosis of anaphylaxis and its validation using population-based data on emergency department visits for anaphylaxis in Florida," Journal of allergy and clinical immunology, vol. 126, no. 1, pp. 98-104. e4, 2010.

[5] H. A. Sampson et al., "Second symposium on the definition and management of anaphylaxis: summary report—Second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network symposium," Journal of Allergy and Clinical Immunology, vol. 117, no. 2, pp. 391-397, 2006.

[6] R. L. Campbell et al., "Evaluation of national institute of allergy and infectious diseases/food allergy and anaphylaxis network criteria for the diagnosis of anaphylaxis in emergency department patients," Journal of Allergy and Clinical Immunology, vol. 129, no. 3, pp. 748-752, 2012.

[7] O. Isakov, I. Dotan, and S. Ben-Shachar, "Machine learning–based gene prioritization identifies novel candidate risk genes for inflammatory bowel disease," Inflammatory bowel diseases, vol. 23, no. 9, pp. 1516-1523, 2017.

[8] M. J. Matasar and A. I. Neugut, "Epidemiology of anaphylaxis in the United States," Current allergy and asthma reports, vol. 3, no. 1, pp. 30-35, 2003.

[9] D. A. Sclar and P. L. Lieberman, "Anaphylaxis: underdiagnosed, underreported, and undertreated," The American journal of medicine, vol. 127, no. 1, pp. S1-S5, 2014.

[10] O. Shahid et al., "Machine learning research towards combating COVID-19: Virus detection, spread prevention, and medical assistance," Journal of Biomedical Informatics, vol. 117, p. 103751, 2021.

ISSN (online): 2581-3048
Volume 7, Issue 11, pp 430-435, November-2023
https://doi.org/10.47001/IRJIET/2023.711058

[11] T. Latchoumi, J. Dayanika, and G. Archana, "A comparative study of machine learning algorithms using quick-witted diabetic prevention," Annals of the Romanian Society for Cell Biology, pp. 4249-4259, 2021.
[12] K. Goto et al., "Novel machine learning method allerStat identifies statistically significant allergen-specific patterns in protein sequences," Journal of Biological Chemistry, vol. 299, no. 6, 2023.
[13] A.Muraro et al., "EAACI guidelines: anaphylaxis (2021 update)," Allergy, vol. 77, no. 2, pp. 357-377, 2022.
[14] F. E. R. Simons and H. A. Sampson, "Anaphylaxis: unique aspects of clinical diagnosis and management in infants (birth to age 2 years)," Journal of Allergy and Clinical Immunology, vol. 135, no. 5, pp. 1125-1131, 2015.

AUTHORS BIOGRAPHY

L.M. Dilanka Matheesha Rajapakse,
Student,
Department of Information Technology,
Faculty of Computing,
Sri Lanka Institute of Information Technology,
Sri Lanka.
E-mail: it20019518@my.sliit.lk

Mr. H.M. Samadhi Chathuranga Rathnayake,
Lecturer,
Department of Information Technology,
Faculty of Computing,
Sri Lanka Institute of Information Technology,
Sri Lanka.
E-mail: samadhi.r@sliit.lk

Ms. Sanjeevi Chandrasiri,
Lecturer,
Department of Information Technology,
Faculty of Computing,
Sri Lanka Institute of Information Technology,
Sri Lanka.
E-mail: sanji.c@sliit.lk

Citation of this Article:

L.M. Dilanka Matheesha Rajapakse, H.M. Samadhi Chathuranga Rathnayake, Sanjeevi Chandrasiri, “Anaphylaxis Prediction and Tailored Medicine Prescription System Using CNN-Based Tabular Data Analysis” Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 11, pp 430-435, November 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.711058

****

