ABSTRACT

The research topic will revolve around assessing the accuracy, efficacy, speed, time, and storage of machine learning when utilized to identify allergies within individuals and how it compares to modern day allergy test methods. One of the methods to be conducted is training a Large Language Model(LLM) so that it will use data provided by researched documentation revolving around allergy likelihood factors such as environmental data, family medical history, or food patterns. The other methods to be conducted is once the LLM is trained, it will then be written into a web application, deployed, and then sent out through a link to a variety of students across Stetson University and University of Florida to receive user feedback on how accurate the AI was in the allergy prediction it made towards the user. The third method to be conducted is to implement a way to test the LLM in terms of time complexity when run on a regular machine, the computer science department server, and a regular laptop. The fourth method to be conducted would be to implement a way to test the LLM in terms of space complexity when on a regular machine, a regular laptop, and the computer science department server. The final method is to conduct research on the background information in regards to the modern day allergy test, then compare said information to the LLM details that were gathered from earlier.

Introduction

Allergy detection has been a method used as early as 1869 when the first skin test for allergies was invented. Since then, there have been many different ways that have been utilized in modern day times to aid in detecting potential allergies within an individual. Such ways include a skin prick test where an allergist may prick the skin of the patient with a small, sterile probe containing the food allergen. If a wheal develops, that would then indicate that the patient is positive for the test, but does not determine the severity of the reaction. Another popular method to determine the likelihood of allergies would be a blood test which, unlike the skin prick test, determines the amount of IgE, or immunoglobulin E, antibodies to whichever foods are tested at the moment. For context, an IgE antibody are antibodies that are produced by the immune system, and when an allergy is detected, the immune system will then react to the allergen by creating the IgE antibodies which then travel to cells, releasing chemicals, and thus creating an allergic reaction.

There are multiple types of IgE antibodies. Each one has a specific sensor of sorts for different varieties of allergens. For instance, some people may have multiple allergies due to having more IgE antibodies than their neighbors who may only have one allergy as they would have less IgE antibodies. When it boils down to it, this is exactly what the blood test is looking for when being used on patients. Another popular method used in determining allergies is also using an oral food test. This is where the patient is given gradually increasing amounts of the food that may be causing their reaction whilst being heavily monitored. This one proves to be the

most dangerous as there can be a higher risk of the evaluation going wrong, such as the patient having a severe reaction that will then need immense care quickly. The last common method found would be a food elimination diet placed on the patient by the allergist. With this method, the patient eliminates all food suspected of causing their outbreak for roughly two weeks to see how they will react towards the end of the period when they try the food again. If they flare up to the food item after having not eaten it for two weeks, it will determine the test positive for a potential allergic reaction.

Further along, with the advancement of technology, this thus creates a greater advancement in the healthcare industry as well. One popular software that has been released is known as CircleDna. CircleDna is a cross-compatible app that enables a user to take an at-home DNA test where they will swab their saliva with a kit they received, send it back to the facility, and get updated results on many different factors of their DNA. Such factors include diet, nutrition, common health risks, etc. that a user can check through results that are sent back on their user account. It allows the user to see the likelihood of how high or low risk they are in terms of if the user is receiving an adequate amount of vitamins in their body, if the user may have alcohol sensitivity, their likelihood of being lactose intolerant, and many other important pieces of information pertaining to the user's DNA.

Now with this in mind, it brings up another good question of how far can modern day technology really go in these advancements. One popular topic that has been talked about as of recent is that of artificial intelligence. Artificial intelligence(AI) is a way for machines to be both trained and built to handle certain tasks so that it may learn and adjust based on both input and expected outcome. AI, in a sense, acts almost human-like in nature and is a way to better aid in the advancement of modern day methods used in many different aspects. AI normally utilizes

what is known as a large language model(LLM). An LLM is a way to generate human-like text, whilst the model learns from patterns in data in a way that can create output that can be both adaptable and generalizable. Normally, LLM's are trained on large datasets of texts like articles, books, etc. so that they may adapt based on the anticipated outcome. During training, the model is able to predict patterns such as foreseeing the next word in a formulated sentence. Models can be fine tuned to operate on more specific datasets and basically reevaluate its understanding so that it can perform a task more accurately. In other words, fine tuning aids the model to become a specialist at a given task.

What this research aims to do with all of this information that has been gathered thus far is see how far can artificial intelligence go in this sector of the medical field. Can AI really aid in allergy predictability? How efficient can it become and will it compare to modern day methods? If so, what will that mean for the medical field in the unforeseeable future? These are all questions that are being aimed to be answered by the end of the research. AI has already proven to be so powerful with models such as GPT-4, BERT, Llama, etc. paving the way in society and the current workforce. There are curiosities revolving around just how far it can go in different aspects, especially on topics that are still being researched today by modern day professionals when it comes to specific allergy predictors (such as peanuts, walnuts, shellfish, etc.) outside of the typical tests we know.

This research is done to test that limit and see where current AI advancements are at and if there is a potential for non-invasive tests to be done on humans when determining allergy related factors. The research question to be asked is how accurately can an AI system predict the likelihood of an individual having specific allergies based on environment, genetic, and personal health data, and what is the efficacy of these predictions compared to traditional testing methods?

The hypothesis is that, when tested the AI will be roughly 60% accurate in its allergy predictions and will require the same amount of speed and storage when tested on different machines as well when tested with different amounts of queries being asked. The scope of the study revolves around the knowledge readily available on training the model and the vast amount of articles and prior research done on the topic of allergen detection.

Literature Review

Before providing further information on the research at hand, it is crucial to understand the similar research that has been conducted beforehand to gain a better understanding of the overall experimentation. While there has been limited research in relation to what factors primarily cause an individual to have a specific allergy outside of IgE antibodies, there has been more research conducted that revolves around the utilization of machine learning methods in allergy detection. One such study conducted back in 2014 focused on predicting phenotypes of asthma and eczema using machine learning. In an article written by Mattia CF Prosperi, Susana Marinho, Angela Simpson, Adnan Custovic, and Iain E Buchan, these individuals discuss the process of their research. The aim of their research was to recognize the extent that certain heterogeneous information can be merged to reveal particular clinical indications. By the end of their research, they had found that more complex modeling was necessary to better understand the mechanisms of such diseases and overall healthcare.

Another study conducted back in 2019 focused on utilizing machine learning methods to predict the development of asthma within adolescents. In the study conducted by Julie L. Harvey and Sathish A.P. Kumar, these individuals aimed to develop predictive models that analyzed the asthma health of a particular child. By the end of the research, they had found that allergies were amongst one of the variables that had the highest correlation with asthma. They intend and hope that their findings may be used within a clinical setting to further push to aid in asthma development predictability amongst medical professionals to execute asthma treatment early on.

Both this research, and the research mentioned prior have a pattern of using machine learning methods to aid in clinical predictability outcomes, primarily focused on asthma to see its correlation with any other factors that may play at hand.

Furthermore, with there being greater involvement in artificial intelligence and machine learning methods, there are now further recent studies that have been conducted which provide more up to date insight. One such study conducted back in 2023 aimed its focus at the current state and prospects of how artificial intelligence will be utilized in allergy detection. Within the study, they found that as of the time the study was conducted, there were no AI or machine learning based applications being used in the field of allergy. They theorize that with the rapid expansion of artificial intelligence and machine learning methods, there will be a potential for them to be used in the field soon. Throughout their tests, they found that machine learning based modeling of a component-resolved diagnostic multiplex array data displayed how the component-specific IgE responses to multiple allergenic proteins are functionally coordinated and co-regulated. This also showed how the interactions that occurred were associated with not only determining whether an individual has asthma, but also the severity.

Another study conducted back in 2020 focused on the development of a deep learning model for detecting allergic reactions by using the safety event reports found across certain hospitals. Their objective was to create a deep learning model that would identify allergic reactions in a free-text narrative of hospital safety reports. What they tested for were generalizability, efficiency, productivity, and interpretability. By the end of the conducted study, they had found a few interesting outcomes to note. For starters, the research conducted demonstrated that a deep learning model can indeed identify allergic reactions while using free-text narratives that were composed by differing health care professionals. What was

concluded was that the model can indeed aid in the improvement of allergy care, thus allowing guidance for errors that appear medically as well as the improvement of the system overall. With this in mind, it creates a better understanding of its relation to the study conducted within this paper as it provides further data on the allergy detection when being handled with AI and machine learning methods.

Furthermore, with the advancement of AI technologies progressing and being developed, there have been theoreticals appearing for what might come of the future. Some current theories are that there can be an anticipation for AI to further deepen our knowledge of disease mechanisms and contribute to precision medicine in allergy. Another theory states that it can be anticipated that AI and machine learning algorithms will be increasingly employed in allergy research and applications soon. Another theory states that as the field of AI research rapidly expands, there are many opportunities to implement AI to better understand and characterize allergic disease processes to deliver personalized care. This entails that, with the direction and pace that AI is moving, some present day theories would suggest that overall AI may soon be seen be utilized not just for allergy research, but also medical care towards the allergies themselves so that patients may soon see advancements in the treatment they receive as well once given a thorough assessment.

Moreover, what this research would contribute to the existing knowledge of today is further understanding behind the different methods that are conducted in allergy detection. Currently, there is limited knowledge on how AI can help in allergy detection overall. Most of the research that has been conducted is still somewhat far from reaching the goal of detecting specific allergies within individuals so that they receive the proper care and treatment for it. For the time being, there is still limited knowledge on what directly causes a specific allergy in and

of itself as most of the research done today focuses on what causes allergic reactions in general, but not what may cause a specific allergy. The research done in this paper will focus on figuring out if AI may find specific allergies a user may have through the use of simple questionnaires to see if there can at least be a further starting point in non-invasive allergy test methods. Hopefully, by the end of the research that can be proven true. With the knowledge gathered, there may be a future for further advancements for allergists to utilize and put into practice for patients to come.

References

https://acaai.org/allergies/testing-diagnosis/food-allergy-testing-and-diagnosis/

https://circledna.com/en

https://www.proquest.com/docview/922255910?parentSessionId=3P4dgAKCE4GEOa1wg72AeZL7Uqy0qUloajcpHb8oeec%3D&accountid=2193&sourcetype=Scholarly%20Journals(Dupont, Christophe. “Food Allergy: Recent Advances in Pathophysiology and Diagnosis.” *Annals of Nutrition & Metabolism*, vol. 59, 2011, pp. 8–18, https://doi.org/10.1159/000334145.)

https://jamaicahospital.org/newsletter/the-history-of-allergies/#:~:text=In%201869%20the%20first%20skin,allergies%2C%20was%20introduced%20in%201914.

https://www.aaaai.org/tools-for-the-public/allergy,-asthma-immunology-glossary/immunoglobulin-e-(ige)-defined

https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html

https://link.springer.com/article/10.1186/1755-8794-7-S1-S7#Abs1

https://ieeexplore.ieee.org/abstract/document/9002692/authors#authors

https://onlinelibrary.wiley.com/doi/10.1111/all.15849#:~:text=Currently%2C%20no%20registered%20AI%20and,of%20awareness%20of%20potential%20applications.

https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2772992

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10169188/#Sec16title

https://onlinelibrary.wiley.com/doi/full/10.1111/all.15849

- Add two yt vids watched from earlier