Allergy Prediction Using Artificial Intelligence

SDMAY24-48: Joseph Trembley, Noah Ross, Xerxes Tarman, Ella Godfrey, Alex Ong
ADVISOR/CLIENT: Ashraf Gaffar

Problem:

Testing for allergies can be expensive, time-consuming, and invasive, sometimes even requiring costly lab tests and blood work. When prescribing medication, a doctor needs to know what allergies their patient could have.

Solution:

AI can be leveraged to predict allergic reactions to medicines based on key genetic information such as age, gender, skin tone, and skin conditions. This process requires nothing more than the patient providing the information above and happens in seconds without skin pricks or lab time. The patient's name isn't attached to the data, so privacy is never a concern.

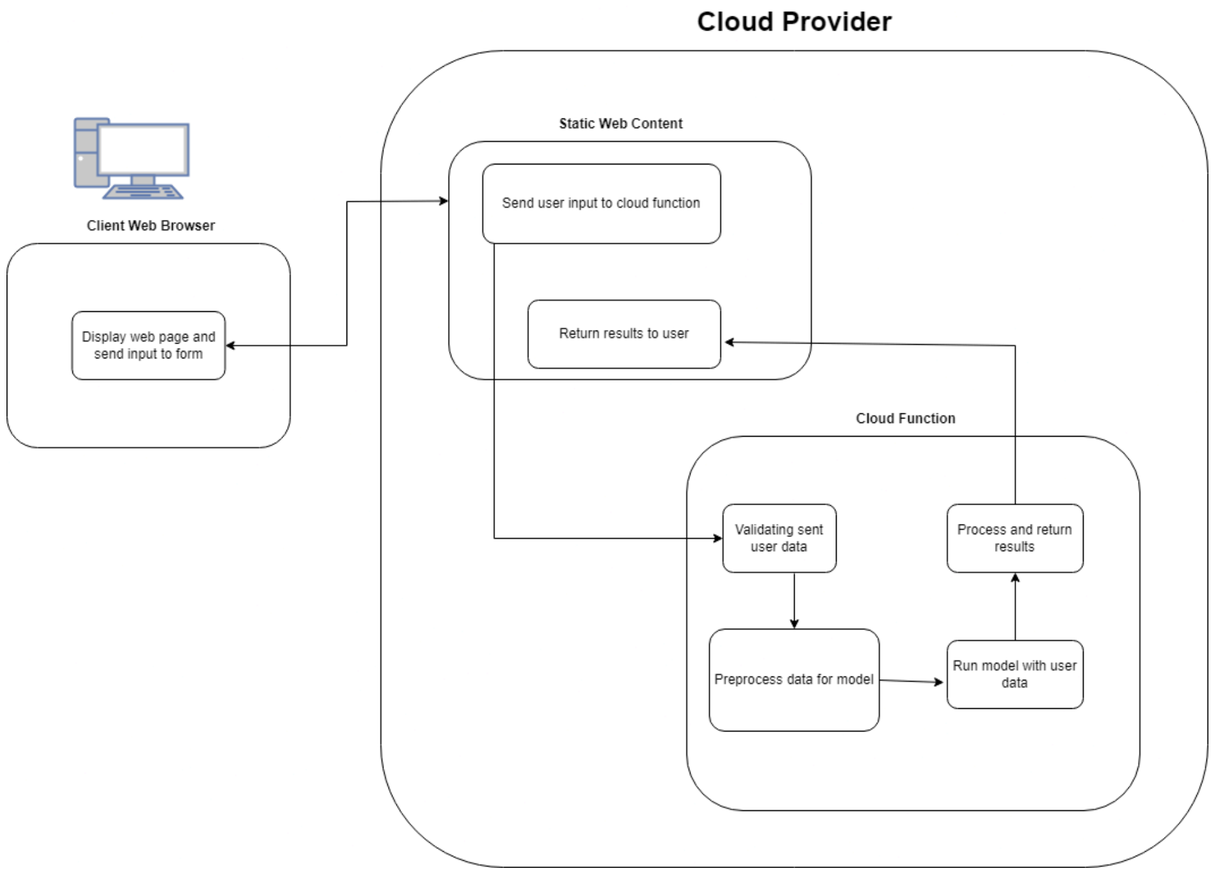
Design Requirements & Approach:

Functional Requirements:

- User-friendly interface that communicates with the model
- Neural network model that has a recall score of over 0.53 and a precision score of over 0.20

Non-functional Requirements:

- Keep data private and not violate any health information privacy laws
- Return a prediction to the patient in less than 5 minutes
- Must cost less than existing allergy detection methods



Final Design:

- For both the AWS and GCP Final designs, the final design included a cloud-based static bucket, reachable by a client's web browser, communicating with a neural network model uploaded and stored in a Docker image. This allows for lower operating costs because no model or backend is constantly running and is instead available on request. A potential downside to this is the cold start time which occurs when the model has not been used recently and needs to be initialized.

Technical Details and Descriptions

Model Type	Precision	Recall	F1-Score
SVM	0.25	0.13	0.17
Neural Network	0.35	0.86	0.65

- Neural Network: A neural network is a machine learning program that mimics the human brain by sending signals to small objects called neurons. They are all connected through complex layers, and this concept is the basis for how we created our model.
- F1 score: The F1 score is a metric used to evaluate the performance of a model when both precision and recall are important. It is a combination of precision and recall, providing a balanced measure of the model's accuracy. For each threshold value, the F1 score is computed based on the model's predictions compared to the ground truth labels.
- Recall: Recall, also known as sensitivity or the true positive rate, evaluates the model's ability to correctly identify all positive instances from the total number of actual positive instances in the dataset. A recall value of 1 indicates perfect recall, meaning all positive instances were correctly identified.
- Precision: Precision is a metric used to evaluate the accuracy of the model in minimizing false positives. It measures the proportion of true positive predictions out of all positive predictions made by the model. A precision value of 1 indicates perfect precision, where all positive predictions made by the model are correct.

Users & Use Cases:

Healthcare professionals can offer this initial screening as a fast and cheap alternative to existing allergy testing methods. With the help of a nurse or a doctor, the patient can fill out the questionnaire provided on our front end and receive a prediction of potential allergies to medications.

Potential Providers:

Part of the project was comparing Amazon Web Services (AWS) to Google Cloud Platform (GCP) to determine a better option for deployment. Cost to run, time to compute, and ease of use were the main metrics and considerations of quality. Overall, GCP was generally easier to use and provided greater transparency to their services and their costs. AWS, however, offers a free tier for any new account for a full year, which is more friendly for starting a project like this compared to the $300 credit GCP offered. Below is an analysis of the time and cost comparison for the two services.

Cloud Provider	Cost	Memory Usage	Cold Start Time	Warm Start Time
GCP	$0.02	> 512MB	11 seconds	.6 seconds
AWS	$0.00	700 MB	10 seconds	.4 seconds

Application:

Prediction with allergens found

Prediction with no allergens detected

The frontend client includes the ability to input necessary genetic conditions to the best of the user's ability to provide the necessary parameters for the model to use. The client then sends a POST request with a JSON payload to the model with the data, and waits for it to return a prediction, which it then displays. The prediction returned from the backend and model is a list of ingredient ids the patient may be allergic too. One of the limitations of the data set was that only ingredient ids were provided, with no correlation to product or ingredient labels, meaning only the id of the ingredient is possible to be displayed.