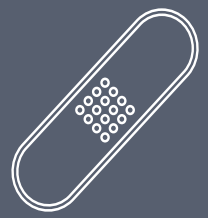


Flu Innovators

VACCINATIONS HAVE NEVER BEEN QUICKER AND SAFER BEFORE

About us

We are the
FluInnovators Team

Our personal goals

1. Provide clinics with excellent vaccine analysis software;
2. Be innovative and use the latest technology available;
3. Learn how to improve our abilities and skills to keep up with the needs of people all around the world.

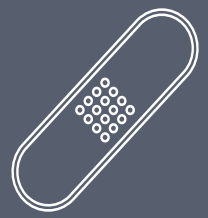
3 The three topics of today

I. BUSINESS & MARKET RESEARCH

II. TECHNICAL DETAILS & IMPLEMENTATION

III.AI MODEL DETAILS

(AND THEN, WE ANSWER QUESTIONS!)



Flu Innovators

MARKET RESEARCH

Market Research
> small wins from the past

The past has shown there is a growing interest in this kind of technology.

Through **statista**, we have found a few key remarks.

- Predictive analytics have a high rate of adoption in East Asia.
- Doctors are likely to use AI technology provided it works
- AI is much faster than traditional predictive models

Yet we still have challenges to overcome, mainly the skepticism in the technology.

Market Research > segmentation

- Performed by selecting the population according to three separate criteria.
- This will ensure the widest reach and will let us select where to deploy the product first to ensure its success.

Brief summary:

- Geographic: expected location
- Behavioral: expected behavior traits of users
- Demographic: expected population dynamics

Market Research > user persona

Meet **Jane Doe**, a medical professional at a clinic in **Cluj**.

- Needs a professional tool to streamline work.
- Would like to help as many people as possible without delays.
- Kind, analytical, and punctual.

Painful:

- Everything must be looked at and intuited by hand, with over 3000 variables to potentially track.

Challenge:

- Process large amounts of data, quickly and correctly

Market Research
> value proposition

VACCINATIONS HAVE NEVER BEEN QUICKER AND SAFER BEFORE

LET'S IMAGINE HOW THIS WOULD BE USEFUL IN A CLINIC

Market Research > business model

A B2B BUSINESS MODEL

- Our product is sold for use by clinics in a B2B transaction.
- **Option 1:** software as a service; you access our web application directly hosted by us
- **Option 2:** integration with existing platform; we integrate into the clinic's platform, and they subscribe so their backend & AI model servers continue operating.
- Each clinic chooses its option; the first one is cheaper but feels less tailored for their clinic.

DATA COLLECTION

- As you know, we must collect data to train our AI.
- The first package option will be stricter and require the clinics to anonymously provide details about each blood sample.
- The second package option will allow clinics to opt into optimizing the AI model. This allows for greater freedom of integration with their existing data collection terms and services.
- The blood sample data cannot be used to track back patients, as this data doesn't behave like fingerprints. We do not want to ruin patient privacy.

Market Research > business model

COST STRUCTURE/REVENUE STREAMS

- The following are money sinks:
  - Server hosting – backend, frontend, AI server
  - Employee salaries – programmers, technicians, support team, legal & financial team
- The following are our revenue streams:
  - Option 1: yearly subscription: €2400 / clinic
  - Option 2: yearly subscription: €5000 / clinic
- The prices can be adjusted as time goes or if they are a bit too cheap/expensive (we are open to discussion!)

INITIAL INVESTMENT

- The initial investment is small; we can test the waters by using option 1 only in the beginning. This is less risky.
- The total cost of server hosting initially should be at worst €200 / month, provided we get the most expensive hardware with the best uptime. We can easily cut costs here.
- The team can be small initially, as it will take time before many clinics choose to use our product. Investors don't have to risk investing in huge teams.
- Because of the relatively small investment price, even just 10-15 clinics would turn in a decent profit.
- Once clinics begin using our product, if they stick to it, we believe they will not be able to drop support for it. We may raise prices gradually.

Flu Innovators

TECHNICAL DETAILS & IMPLEMENTATION

Technical Details
> overall look

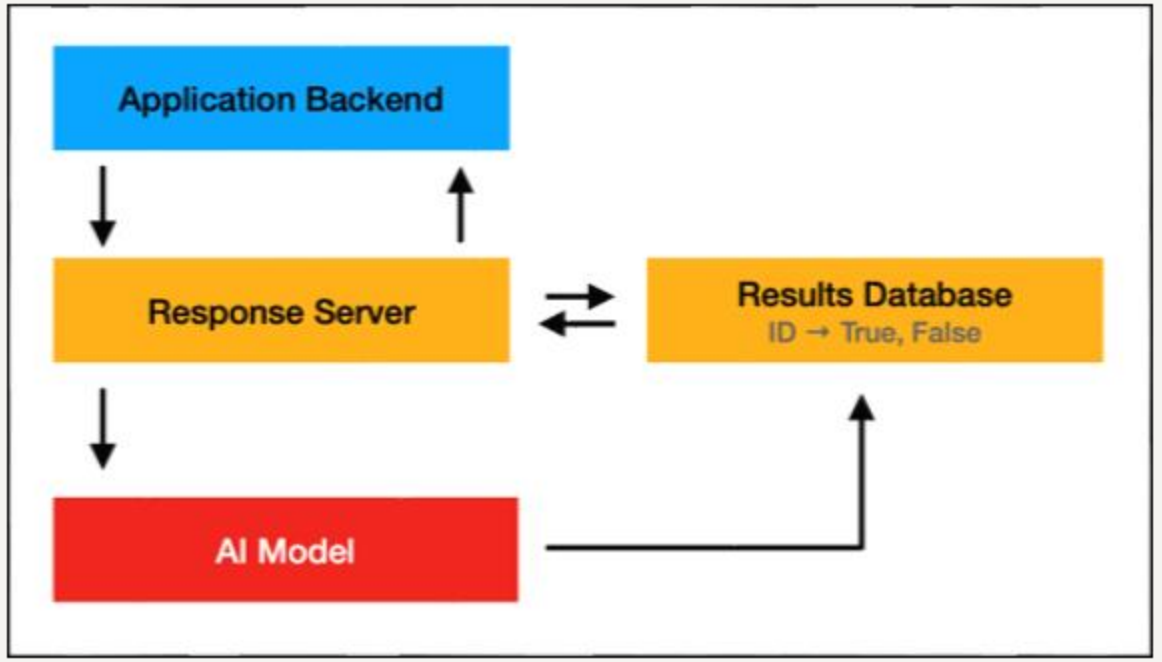
Our product is ultimately a web application, and because of its AI operations (AI-ops), we will have a few challenges:

1. High availability – the system should be resilient and remain often available
2. Efficiency – the backend-ops shouldn't slow AI-ops and vice-versa
3. Portability – the code should run on as many configurations as possible

Technical Details
> architecture

We'll use a true async/await parallel structure:

- The backend server communicates with another server, called "response server".
- The backend has no direct access to AI model calls. It may make requests which return lookups from a cached result database.
- The response server starts computations on multiple threads and then writes their result to the result database.
- The response server always includes in its result whether the computation is complete or whether you may need to ask it again later.



Technical Details
> availability, robustness

AVAILABILITY

- Computations are non-blocking, so the frontend and backend will never freeze waiting for computations.
- If the AI server encounters an error, the backend can still function (doctors can still manually check patient information)
- If the frontend encounters an error, the AI server continues computing so that by the time of a restart, results are available.

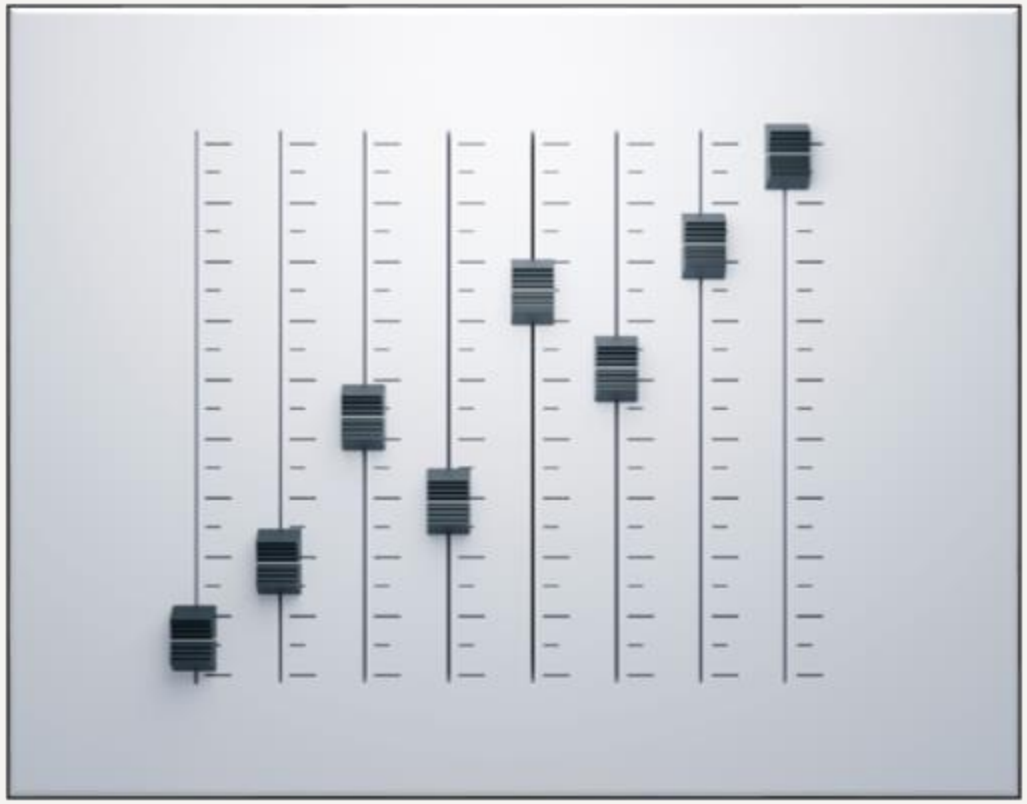
ROBUSTNESS

- The servers are loosely coupled via requests, so crashes do not propagate.
- The AI model is blind to complex details like patient information, clinic details, and geographic data which allows the AI team to focus on purely improving the model's performance without any clean-ups.
- The system resources can be dynamically allocated without stopping the entire application.

Technical Details
> innovation

- Unlike manual doctor checks and classic statistical models, our AI model supports more than 200 variables.
- The system is streamlined using a parallel approach to reduce waiting times for patients, but also to streamline medical staff experience.
- Patients could also view their results in real time by accessing Option 2 platforms, so everybody wins!

Technical Details
> deployment, portability



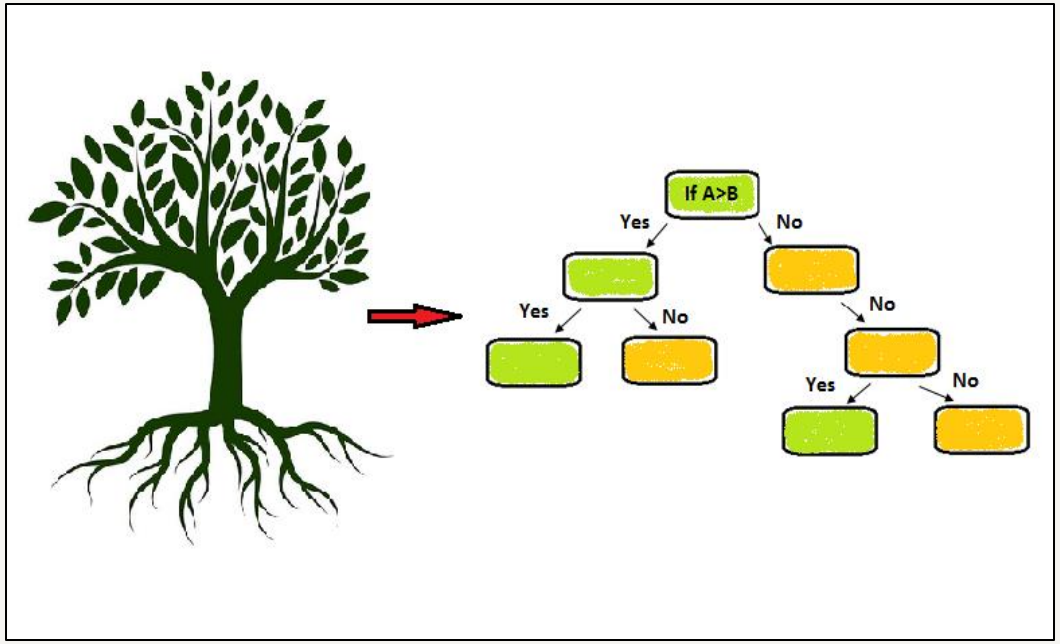
- The system is easy to deploy; in each of the 3 modules, configure the URL for the other modules, and just boot them up.
- The system is designed with horizontal scaling in mind, and very minor restructuring is required for this.
- The code is very portable, and anything from a personal computer to a VPS, to a dedicated server, to a cloud computer can run the overall app, including AI.

Flu Innovators

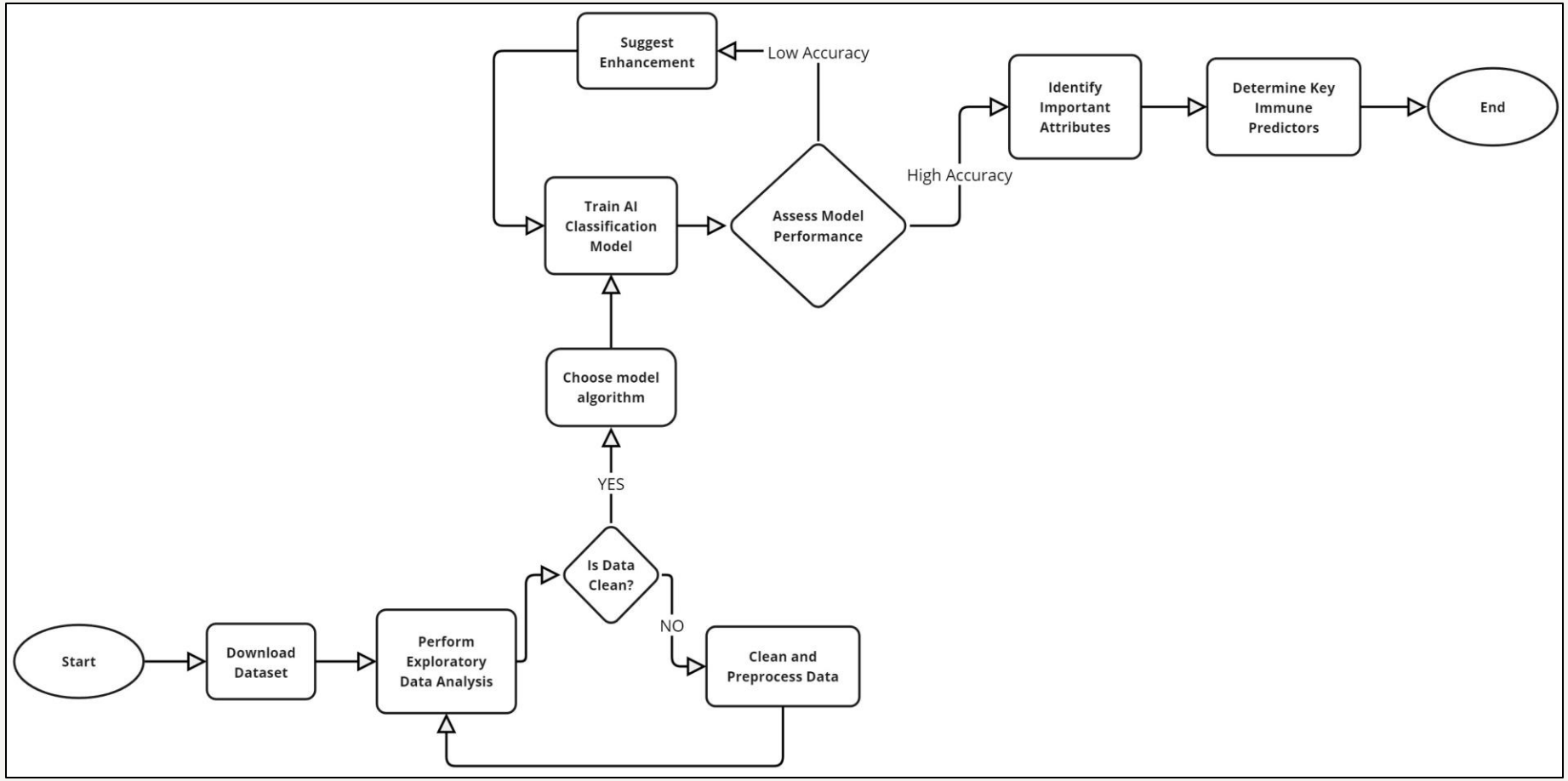
AI MODEL DETAILS

AI Model Details
> goal

- **Analyze medical data** containing individuals with diverse backgrounds to **predict vaccine response** and **determine critical features**
  - We decided to use the **FluPRINT dataset**, containing data about 740 patients in this study
- **Develop binary classifier models** to predict the risk of vaccine adverse reactions
  - We end up using:
    - **Random Forests**: combines multiple decision trees each trained on different parts of the data, reducing mistakes from individual trees
    - **Gradient Boosting (XGBoost)**: builds a series of simple decision trees, where each new tree learns from the mistakes of the ones before it



AI Model Details
> process



AI Model Details
> data preparation

CHALLENGES

- The data was **difficult to understand** because each patient's information was spread across multiple rows;
- There were **too many features**, making it harder to focus and find the important ones;
- Some patients had **missing data** for certain features;
- The data included a **mix of text and numbers**, which required processing to make it usable for AI models;
- Numeric values **varied widely**, needing normalization;
- **Imbalanced data**, meaning the database contained more low responders than high responders.

STEPS TAKEN

- **Transformed the text into numbers** and then brought all the numbers to a **common interval**, more specifically [0,1]
- **Transformed the structure of the dataset** so each patient had data spread across a single row by aggregating key features;
- Tried to **remove highly correlated features** using a Correlation Matrix;
- **Removed features** that had more than 70% missing values.

AI Model Details
> model training & results

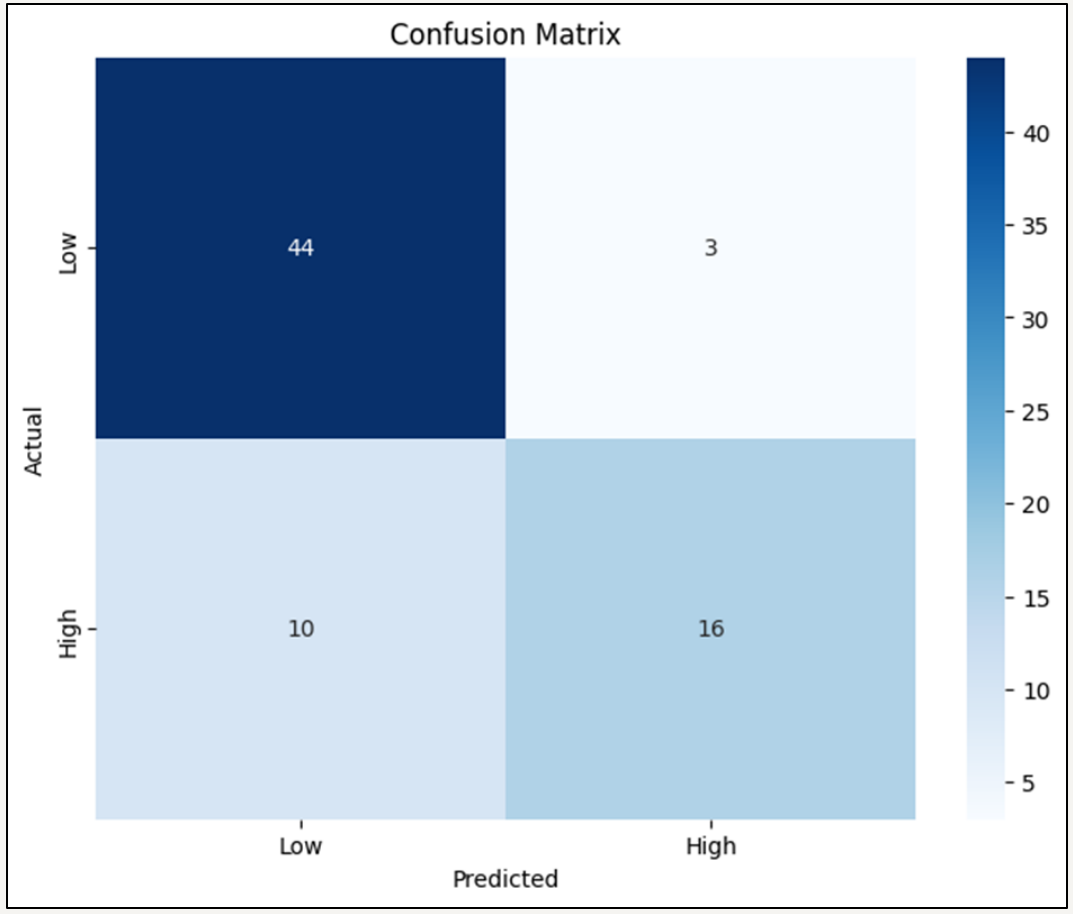
- We experimented with **different parameter values** to train and fine-tune the models to improve their performances.
- Both algorithms proved to have a **fast computation time**, reducing their CO2 footprint and electricity cost.

RANDOM FOREST PERFORMANCE

- **Accuracy**: 0.64
- **ROC AUC**: 0.54

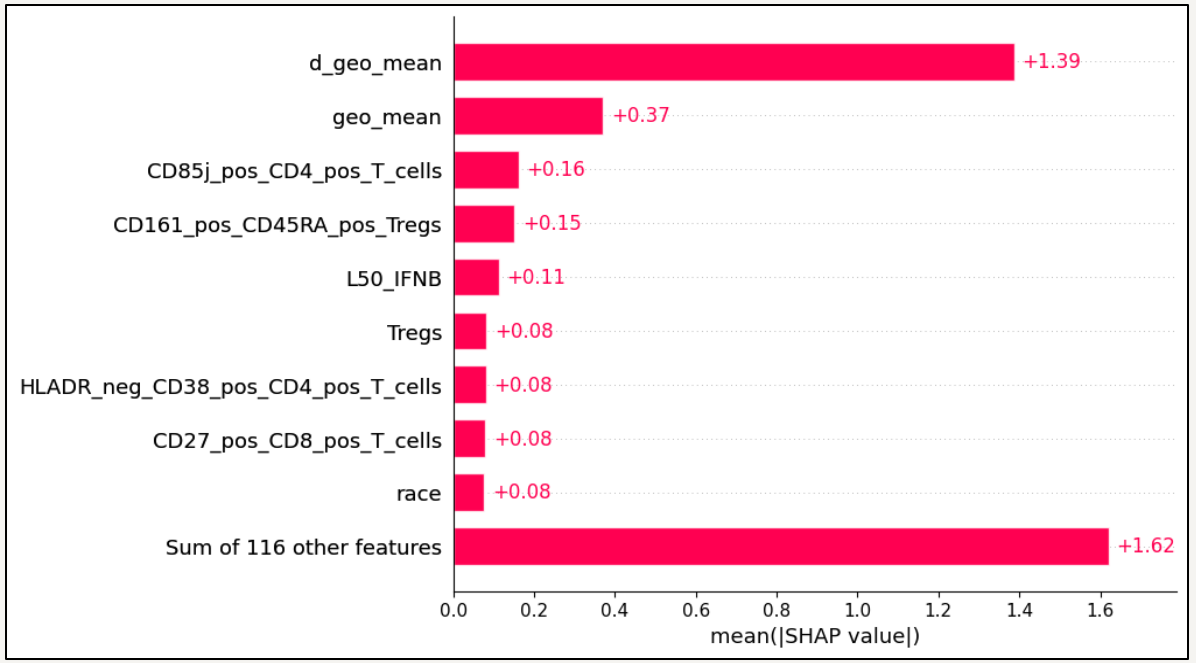
XGBOOST PERFORMANCE

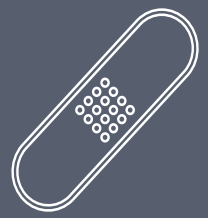
- **Accuracy**: 0.82
- **ROC AUC**: 0.90



AI Model Details
> feature importance

- For this step we used a framework called **SHAP**, which tries to explain how each feature impacts the model's prediction
- Behind the scenes, it tests how the model's prediction changes when it includes or excludes different features by trying out all possible combinations and averaging their impact.





Flu Innovators

DEMONSTRATION IN PROGRESS

Flu Innovators

THANK YOU

PLEASE FEEL FREE TO ASK
ANY QUESTIONS

DEMO VIDEO BACKUP