Term Project

Genetic Risk Assessment and Monitoring App

By Ritika Wadhwa
rwritikawadhwa@gmail.com
+1-857-4249618

Submission Date: 04/02/2023

EMGT5220: Engineering Project Management, Spring 2023
Northeastern University

Instructor: Prof. George Kontopidis

Table of Contents

Revision History

01/20/2023	Draft #1: Original proposal
01/29/2023	Draft #1a: Improvement of Scope/ Technical Approach & adding Organization
02/05/2023	Draft #1b: WBS, RACI, Mindmap, UML Sequence Diagram, UML Activity Diagram
02/10/2023	Draft #2: Preliminary Rough Estimate Financial Plan
02/26/2023	Draft #2a: Improvements in WBS, RACI, Org Chart. Adding SWOT Analysis, RPN, Risk Matrix
03/05/2023	Draft #2b: Critical Success Factors, Assumptions, Mitigation Plan
03/12/2023	Draft #3: Gantt Chart, PERT Chart, Financial Plan
04/01/2023	Final Report: Improved PERT Chart, Gantt Chart, Financial Plan, Org Chart. Added Resource Allocation, and Summary. Other improvements based on feedback

1. Purpose

Genetics plays a crucial part in maintaining our health. In addition to lifestyle and environment, understanding our genetic risks helps monitor and prevent various diseases. Usually, getting genetic insights from our family history and medical records involve finding a genetic counselor in your area and booking an appointment, followed by further testing. This project aims to develop an application that will help patients and the medical community in collecting and maintaining family history, and medical records efficiently and securely, from the comfort of their homes at a reduced cost. It will also use medical guidelines, and data to assess and predict risks for genetic conditions as well as provide guidance and recommendations on genetic counselors, doctors, testing and preventative measures according to the risk assessment. In addition, it will help clinicians save time and money by gathering the data in a secure global cloud-based platform to be easily accessible.

2. Objective

This project will develop an application which will have a platform for patients as well as healthcare providers/ clinicians. The app will be used by patients to store their family history and medical records which will be used to assess and predict genetic risks for patients by using the current guidelines and a predictive analytics machine learning model. The algorithm will provide analysis and insights across various inherited diseases from oncology to rare inherited diseases like Cardiomyopathy, Alzheimer's, pediatric diseases, mental health, and more, and accordingly connect them with clinicians and testing facilities. The data will be stored on a cloud-based platform with a secure environment, following the data protection law and regulations including HIPAA and GDPR compliance, which will allow healthcare providers to access the data easily and safely.

3. Scope

Scope of this project is to develop an application that will allow the users and patients to enter their family and medical history which will be stored in AWS, cloud-based platform.

We will use AWS Sagemaker to build a machine learning model like Regression, Decision Tree, etc. to predict the risks for various genetic diseases using the data from clinical guidelines (E.g., from the National Comprehensive Cancer Network), existing data, and current risk assessment tools like BOADICEA (The Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm), Bayes-Mendel, etc.

Cloud engineers will maintain the AWS servers where the data will be securely stored and use AWS Lambda to maintain code which will enable auto-updates for any new medical information for each patient. They will also maintain a server so the data can be accessed by healthcare providers.

The health tech advisor will be responsible for making sure correct guidelines and regulations are being followed and help create an app which is easy to understand by an average daily user by assisting in built-in Medical Ontology where users can search and get more information on any condition or terms.

We will partner with hospitals and testing facilities to enable users to directly book appointments for a physician or a test through the application based on the recommendations given in the results.

4. Funder/Investor

The funding for early-stage development will primarily be provided by an angel investor.

The funding will also come from genetic testing and counseling facilities like Ancestry, 23andme, etc. by partnering with them as the app can help save them time and money by collecting medical/family history and get new customers from the app who want to go for further testing as the app users can directly book testing and appointments at a discounted rate. Other partners will include healthcare providers for each area of medicine in different cities, so the users can directly book appointments through the app and as well as the healthcare providers can access the data through the app which will be continuously updated in real-time.

The final product will be sold on a subscription basis to users who want to find out their genetic risks with a full report, with data storage as well as healthcare provider companies for data collection and storage.

5. Critical Success Factors

Critical Success Factors for this project:
- Having a successful data collection from national guidelines: Data collection and modification to use it for our machine learning model is one of the most important factors. We need to be able to collect the data from legit sources and convert it for the model.
- Partnerships with testing facilities and hospitals: Building mutually beneficial partnerships with testing facilities like 23andme, and hospitals is extremely crucial for a profitable business and to gain new users and have shared access data.
- AWS Server settings: Integration of AWS Server with hospital data and the machine learning model is important to have real-time updates and share those updates with the users and healthcare providers.

6. Assumptions

The following assumptions have been made to ensure the success of this project:
- The Machine Learning model gives results with desired accuracy and precision that can be used to generate a report that is useful
- Healthcare providers will accept the use of this application and have a shared access database for real-time updates
- Storage of medical data is approved by the necessary government entities

7. Technical Approach

The application will have an interface for users to input data users which will be stored in AWS for a cloud-based secure environment.

AWS will store data and AWS SageMaker will be used to develop the Machine Learning model that will predict the genetic risks and display them in the app. A supervised machine learning model like Regression, Random Forest, etc. will be used to predict the risk based on the data provided by medical guidelines like the National Comprehensive Cancer Network.

The application will also have a portal for all healthcare partners where they can access data of all their respective patients and update any new information they wish to. The AWS server will also update any future medical reports/ genetic testing taken at our partner companies using AWS Lambda and SQL queries so it can be always accessed by both users and healthcare providers.

8. Organization

The organization structure is a weak type of matrix. There are functional managers for each department, Software Engineering, User Experience, Data Science, Marketing. The team reports to their functional managers but also works for the Project Manager. Head of Software Engineering, User Experience and Data Science reports to the CTO and the Head of Marketing and Sales reports to the CFO. The functional manager also reports to the Project Manager to ensure that all tasks are being performed according to the schedule and keep them informed. The project manager works for and informs CTO and CFO.
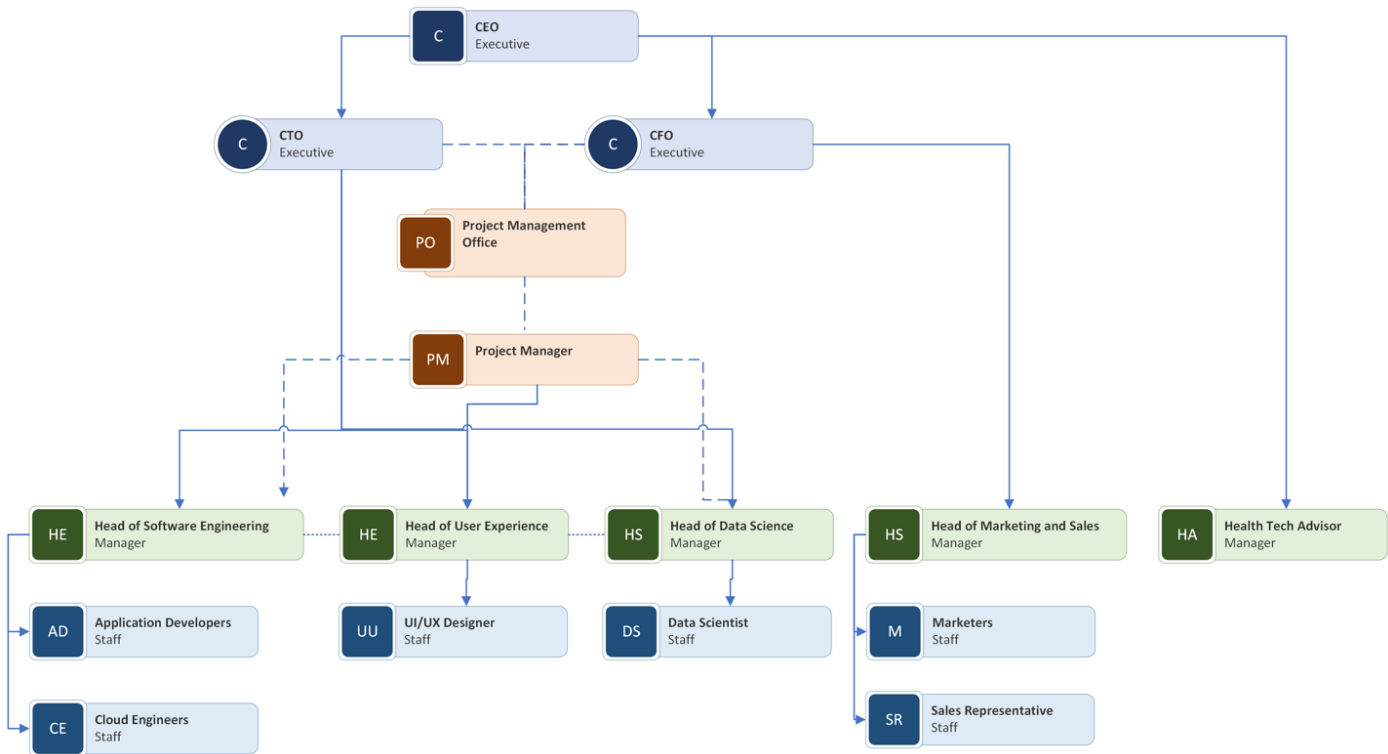
Figure 1: Organization Chart

9. Project Plan

a. Work Breakdown Structure (WBS)

The WBS is a detailed version of the mind map (Appendix 1). It shows how the project is divided into tasks and subtasks and the resources required for each task. There are 5 main phases with subtasks. The table shows how many people are required and how much time it will take to maximize resources and productivity. It also shows what are precursors of each subtask to determine when the task can be started.

TASK ID	PROJECT TASKS	RESPONSIBILITY	TIME (WEEKS)	PRECURSOR	PEOPLE REQUIRED	RESOURCES
1	USER EXPERIENCE	User Experience Manager				
1A	Application Market Research	UI/UX Designer	2		2	
1B	UI/UX Design Patient Interface	UI/UX Designer	2	1A	2	Figma
1C	UI/UX Design Hospital Interface	UI/UX Designer	2	1A	2	Figma
1D	UI/UX Testing	UI/UX Designer	2	1B, IC	2	Figma
1E	Medical Ontology	Health Tech Advisor	2	1A	1	
2	APPLICATION DEVELOPMENT	Software Engineering Manager				
2A	Front-End Development	Front-end developer	2	1D, 1E	2	
2B	Back-End Development	Back-end developer	4	2A	3	
2C	Connection with AWS Server	Cloud Engineer	2	2B, 3B	1	AWS Account
3	AWS SERVER SETTING	Software Engineering Manager				
3A	Create a server	Cloud Engineer	3		2	AWS Account
3B	Server Settings Query	Cloud Engineer	5	3A	3	AWS Account
3C	Machine Learning Model connection with AWS	Cloud Engineer	2	3B, 4D	3	AWS Account
4	MACHINE LEARNING MODEL	Data Science Manager				
4A	Data Collection from guidelines	Data Scientist	1		2	Guidelines & Health Tech Advisor
4B	Data Analysis	Data Scientist	2	4A	2	Python
4C	Model Development	Data Scientist	3	4B	4	Python, AWS Account
4D	Model Testing & Validation	Data Scientist	2	4C	2	Python
4E	Report Generation	Data Scientist	1	4D	2	Health tech Advisor
5	MARKETING & PARTNERSHIPS	Marketing Manager				
5A	Partnerships with healthcare provided	Sales Representative	4		5	
5B	Contracts & Negotiations	Project Manager	2	5B	1	

Table 1: Work Breakdown Structure

b. Resource Plan and Responsibilities (RACI)

The RACI matrix is used to show who is responsible and accountable for each task. It also shows which are the teams that should be consulted for their expertise and other teams who are providing support should be kept informed. This ensures that only one person/ team is responsible for one task, and only one person is accountable for a task so there are no overlaps and there is a clear communication, and everyone can keep track of the progress.
Key: R– Responsible, A- Accountable, C– Consulted, I- Informed

TASK ID	PROJECT TASKS	DEPT of USER EXPERIENCE		DEPT of SOFTWARE Engineering				DEPT of DATA SCIENCE		DEPT of MARKETING			PROJECT MANAGER	HEALTH TECH ADVISOR
		Manager	UI/UX Designer	Manager	Front-end developer	Back-end developer	Cloud Engineer	Manager	Data Scientist	Manager	Marketer	Sales Representative		
1	**USER EXPERIENCE**													
1A	Application Market Research	A	R										I	I
1B	UI/UX Design Patient Interface	A	R		I								I	C
1C	UI/UX Design Hospital Interface	A	R		I								I	C
1D	UI/UX Testing	A	R		I			I					I	
1E	Medical Ontology	R	I					I						A
2	**APPLICATION DEVELOPMENT**													
2A	Back-End Development			A	I	R							I	
2B	Front-End Development	C		A	R	I							I	
2C	Connection with AWS Server			A	I	I	R						I	
3	**AWS SERVER SETTING**													
3A	Create a server			A			R						I	
3B	Server Setting Maintenance			A			R						I	
3C	Machine Learning Model connection with AWS			A			R	C					I	
4	**MACHINE LEARNING MODEL**													
4A	Data Collection from guidelines							A	R				I	C
4B	Data Analysis							A	R				I	C
4C	Model Development							A	R				I	
4D	Model Testing & Validation							A	R				I	C
4E	Report Generation							A	R	I	I	I	I	C
5	**MARKETING & PARTNERSHIPS**													
5A	Partnerships with healthcare provided				I					A		R	I	C
5B	Contracts & Negotiations									A			R	C

Table 2: RACI Matrix

c. Financial Plan

The financial plan shows the monthly budget based on each task. The budget is calculated on a weekly basis based on the timeline from the WBS and Gantt Chart. The distribution of the funds is based on the work done each month on a particular task, and the total monthly and the project budget is also computed.

TASK ID	PROJECT TASKS	TOTAL	Monthly Budget			
			1	2	3	4
1	USER EXPERIENCE					
1A	Market Research	7500	7500			
1B	UI/UX Design Patient Interface	7500	7500			
1C	UI/UX Design Hospital Interface	7500	7500			
1C	UI/UX Testing	7500		7500		
1D	Medical Ontology	6000	6000			
2	APPLICATION DEVELOPMENT					
2A	Front-End Development	12000		12000		
2B	Back-End Development	36000		9000	27000	
2C	Connection with AWS Server	6000			3000	3000
3	AWS SERVER SETTING					
3A	Create a server	18000	18000			
3B	Server Settings Query	45000	18000	27000		
3C	Machine Learning Model connection with AWS	18000		9000	9000	
4	MACHINE LEARNING MODEL					
4A	Data Collection from guidelines	7000	7000			
4B	Data Analysis	14000	14000			
4C	Model Development	42000	14000	28000		
4D	Model Testing & Validation	14000		14000		
4E	Report Generation	7000			7000	
5	MARKETING & PARTNERSHIPS					
5A	Partnerships with healthcare provided	32000	32000			
5B	Contracts & Negotiations	8000		8000		
	TOTAL		131500	114500	46000	3000
					TOTAL BUDGET	295000

Table 3: Financial Plan: Monthly Estimate

d. PERT Chart

The PERT gives a representation of the tasks that are part of the project. It shows the chronological sequence of the tasks with the duration allocated for each task. It is a useful tool for identifying the order that the tasks need be to be done in and to identify the critical path to stay on the schedule and avoid any delays.
By solving the PERT Chart, we can identify the Critical Path which is shown by red arrows.

Critical Path: Task 1A → Task 1B & Task 1C → Task 1D → Task 2A → Task 2B → Task 2C
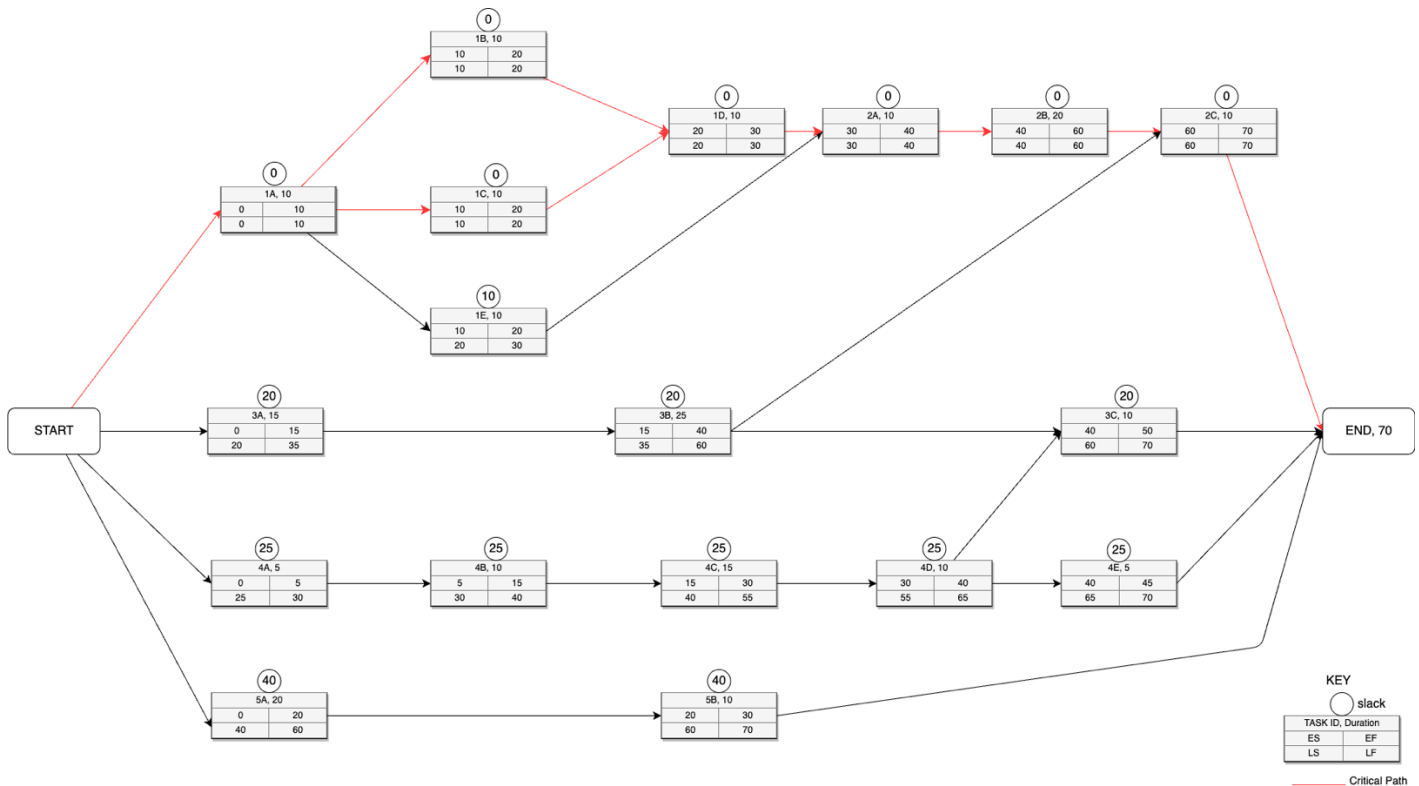The total duration will be 70 days.

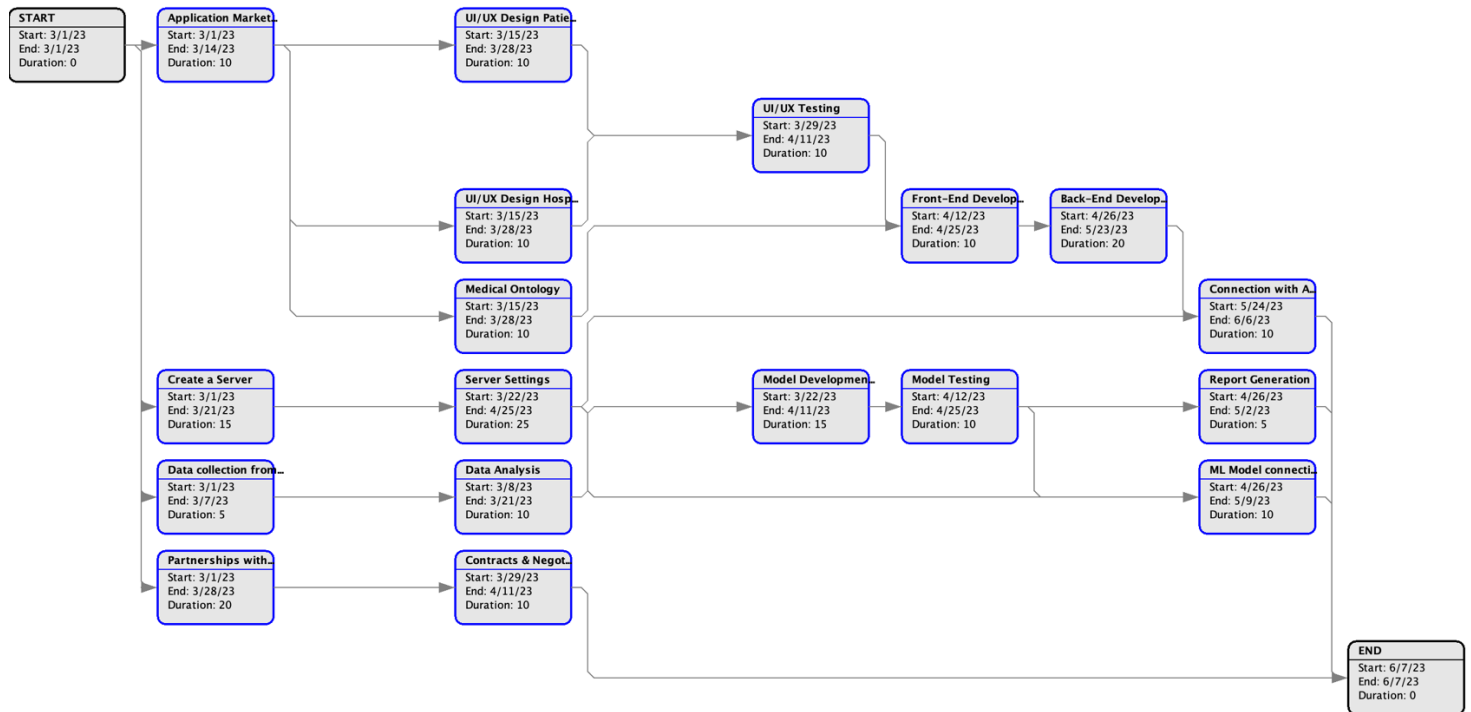
Figure 2: PERT Chart (Solved)

Figure 3: PERT Chart (Based on Gantt Chart)

e. GANTT Chart

The GANTT Chart provides a breakdown of the schedule and tasks. It shows all tasks from the WBS where each task is represented by a bar, and the duration is shown by the length of the bar. Each task's perquisite tasks are also mentioned to show the order and requirements needed to start a new task. The resources allocated are derived from the WBS and the RACCI table. It helps us in estimating the cost required and creating the financial budget.
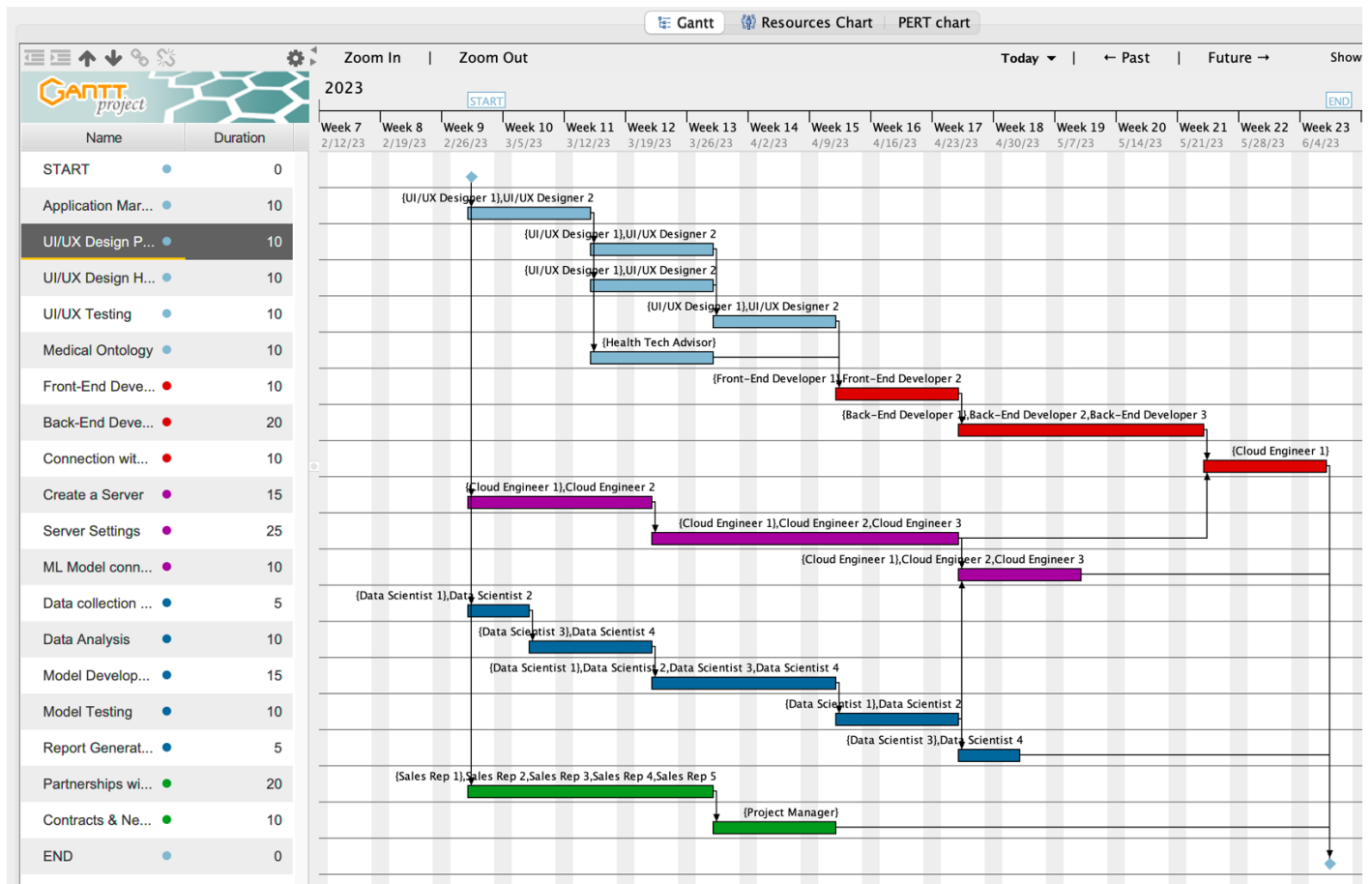
Figure 4: Gantt Chart

10. Risk Assessment

a. SWOT Analysis

Swot Analysis is done to identify Strengths, Weaknesses, Opportunities and Threats. It will help is developing a plan to determine priorities, maximize opportunities and plan for any risk and roadblocks we may encounter.
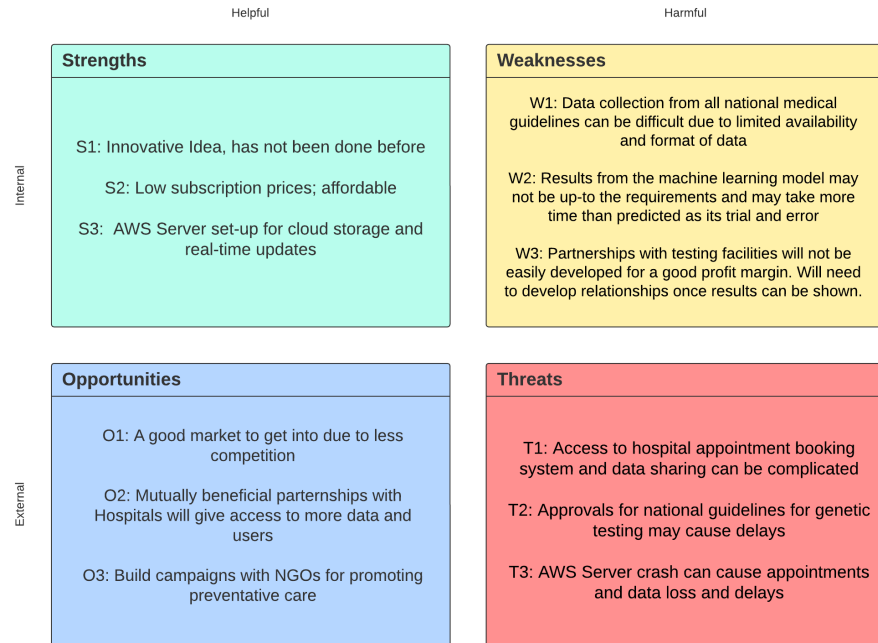
Strengths	Weaknesses
S1: Innovative Idea, has not been done before S2: Low subscription prices; affordable S3: AWS Server set-up for cloud storage and real-time updates	W1: Data collection from all national medical guidelines can be difficult due to limited availability and format of data W2: Results from the machine learning model may not be up-to the requirements and may take more time than predicted as its trial and error W3: Partnerships with testing facilities will not be easily developed for a good profit margin. Will need to develop relationships once results can be shown.

Opportunities	Threats
O1: A good market to get into due to less competition O2: Mutually beneficial parternships with Hospitals will give access to more data and users O3: Build campaigns with NGOs for promoting preventative care	T1: Access to hospital appointment booking system and data sharing can be complicated T2: Approvals for national guidelines for genetic testing may cause delays T3: AWS Server crash can cause appointments and data loss and delays

Internal / External

Figure 5: SWOT Analysis

b. Risk Priority Table

The RPN Table helps us detect severity, likelihood to occur, and the inability to detect the risks.
The product of these factors is used to compute the Risk Priority Number.
The risks are given on a scale of 1-5, 1 being the lowest and 5 being the highest.

THREATS	SEVERITY (S)	LIKELIHOOD (L)	INABILITY TO DETECT (D)	RPN
W1: Data collection from all national medical guidelines can be difficult due to limited availability and format of data	3	3	1	9
W2: Results from the machine learning model may not be up-to the requirements and may take more time than predicted as its trial and error	4	3	4	48
W3: Partnerships with testing facilities will not be easily developed for a good profit margin. Will need to develop relationships once results can be shown.	2	2	2	8
T1: Access to hospital appointment booking system and data sharing can be complicated	5	3	3	45
T2: Approvals for national guidelines for genetic testing may cause delays	4	2	2	16
T3: AWS Server crash can cause appointments and data loss and delays	3	5	3	45

Table 4: Risk Priority Table

c. Risk Matrix

The Risk Matrix helps us visualize the risks that have the highest likelihood and severity. It is useful in creating a mitigation plan to minimize roadblocks and plan for the threats and weaknesses with highest impact. The Matrix shows severity on the X-axis and Likelihood on the Y-Axis.
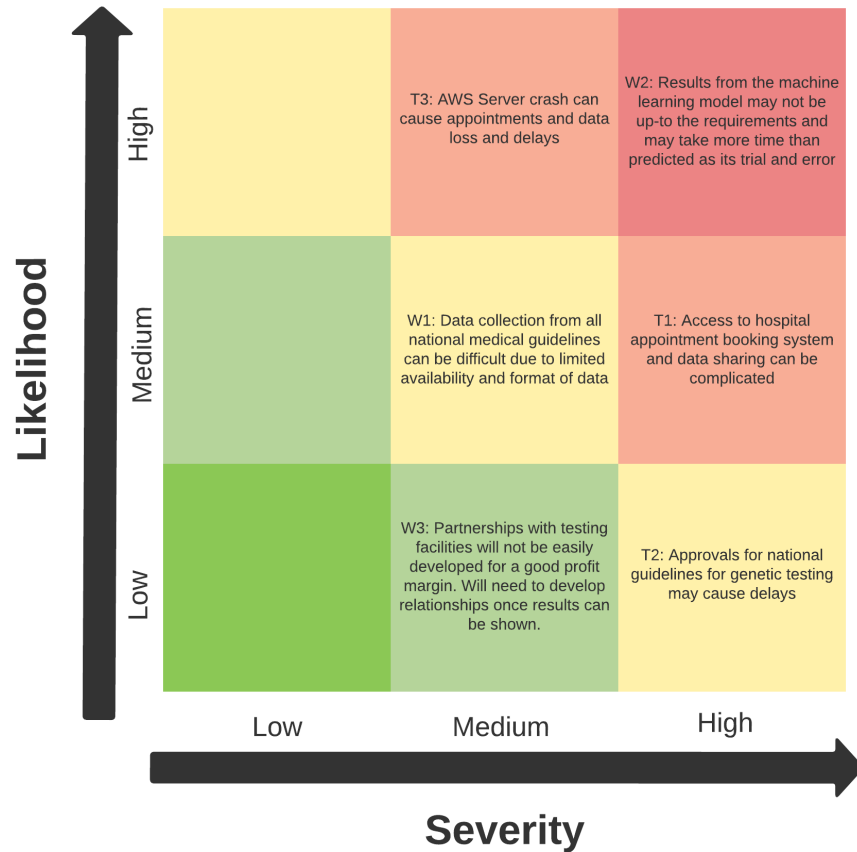

Figure 6: Risk Matrix

d. Risk Mitigation

- W1, T2: Correct authorities will be contacted, and extra time is allocated for all formalities to be done, and to get all approvals to avoid delays in the rest of the schedule
- W2: Ample amount of time allocated to the model development and validation to conduct proper testing and trial-error method to achieve the desired results
- T1: Build partnerships with testing facilities and hospitals to build a shared access database that benefits both parties.
- T3: Build an AWS server with maintenance team to make sure the server crashes are minimum and the real-time data updates are not affected.

11. Resource Allocation

The resource chart has been derived from WBS and Gantt Chart. The resource chart helps in efficiently distributing the workload among the team members. The visualization helps in understanding how the resources are allocated and if any resources are overloaded, so they can be levelled to reduce the workload.

From the chart below, we can see that all tasks are efficiently distributed. There is overloading of resources for UI/UX Designer which is then balanced by increasing the duration to reduce the workload and can be seen in the next chart.
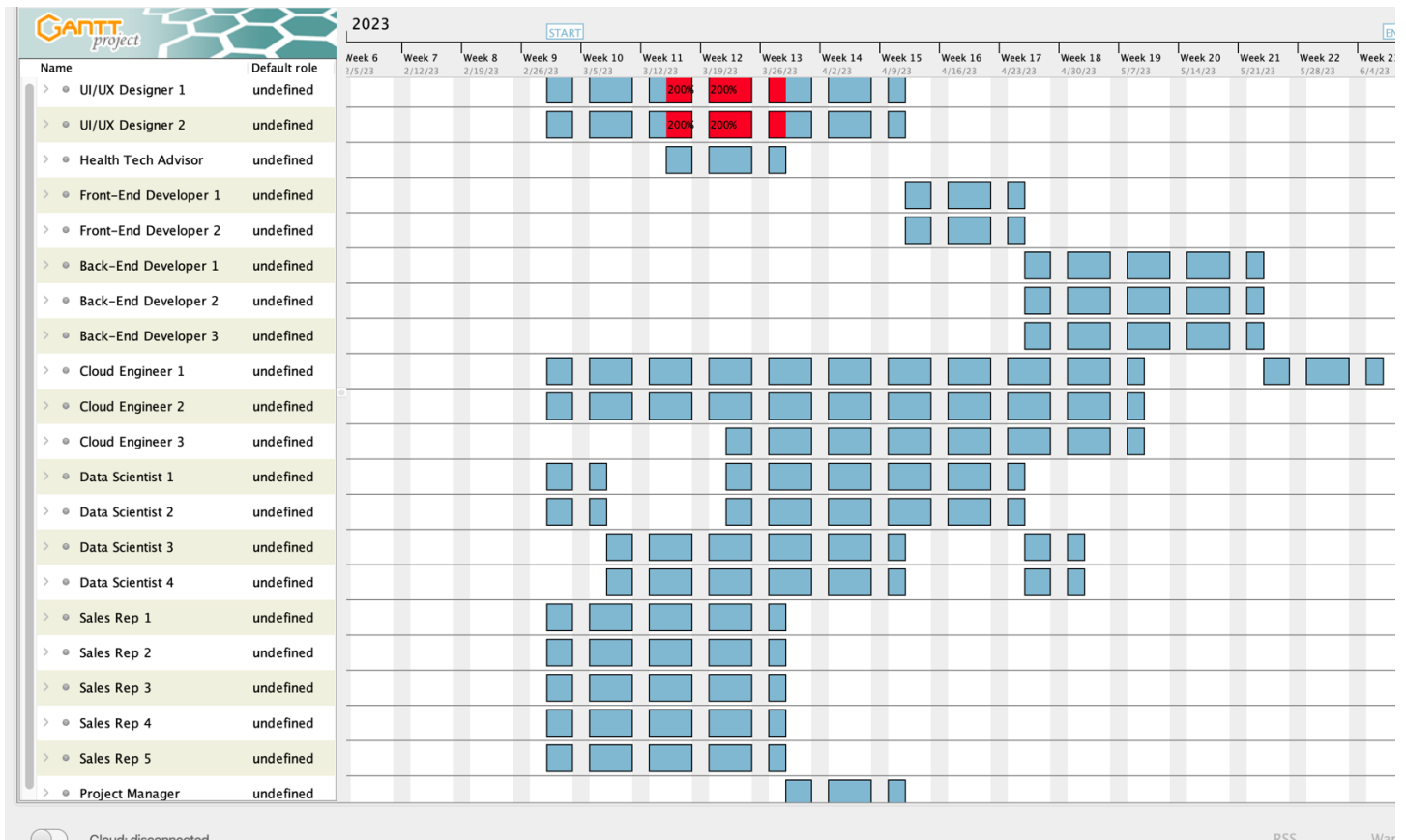
Figure 7: Resource Chart – Unbalanced
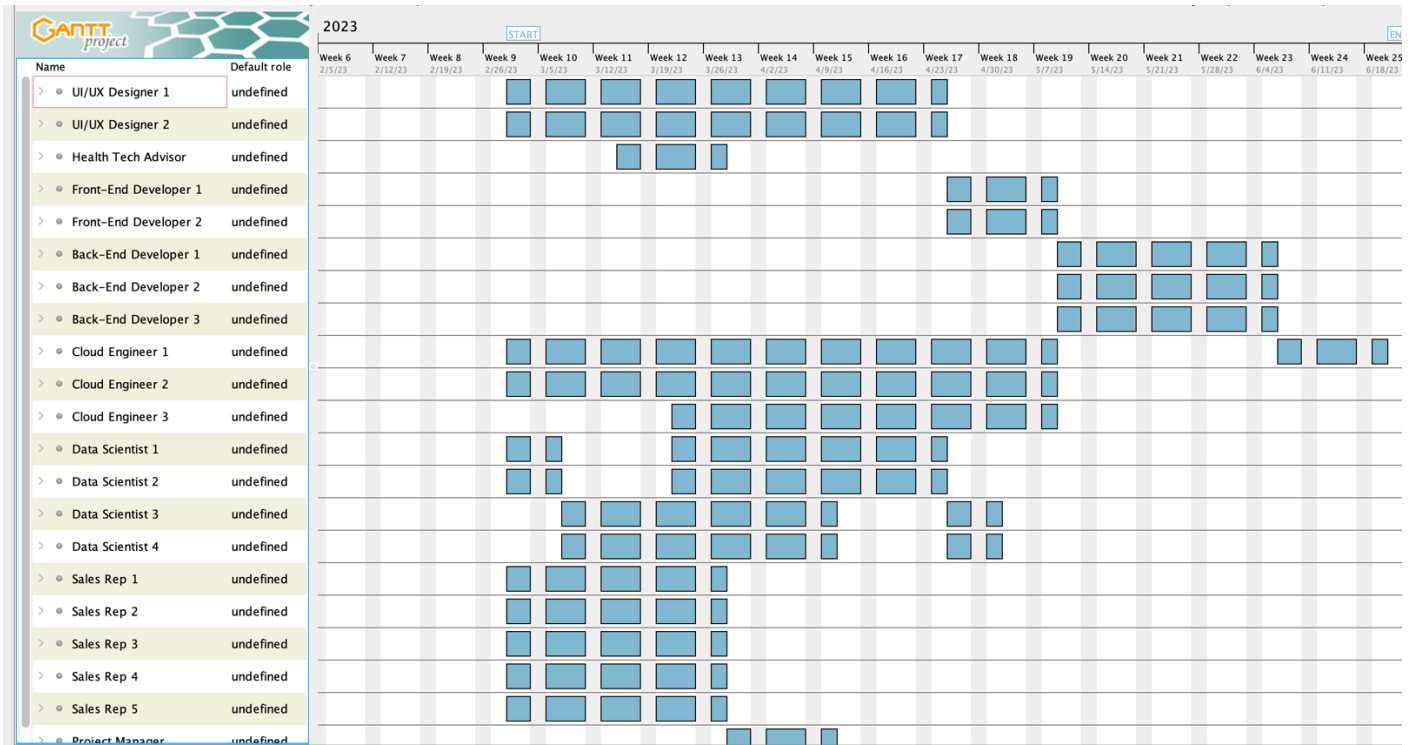
Figure 8: Resource Chart – Balanced

12. Summary

Personalized medicine has the potential to transform healthcare and this project aims to contribute to that. The project goal is to build an application that will help users get a quick genetic risk assessment report from the comfort of their home and they can use that report to book appointments with a physician or further testing based on the recommendations provided. We will also have partner hospitals and testing facilities to make this a easier process. The project will take roughly 14 weeks to be completed according to the plan with an estimate of $300,000.

The project is divided into 5 main phases and carefully analyses the WBS, Resource Allocation, Risk Mitigation plan, etc. It is a step into globalizing personalized healthcare and make the process easier and stress-free for each user and has huge potential to grow in the future.

13. Appendix

a. Mind Map

The Mind Map is used to divide the project into different phases and into tasks, and then subtasks to help in planning of the project. It helps us identify at the beginning of the project what are the different tasks required to complete each phase of the project and gives a high-level description of the project which can be used to track the progress. It also helps in building the WBS and the RACI matrix for resource and budget allocation.
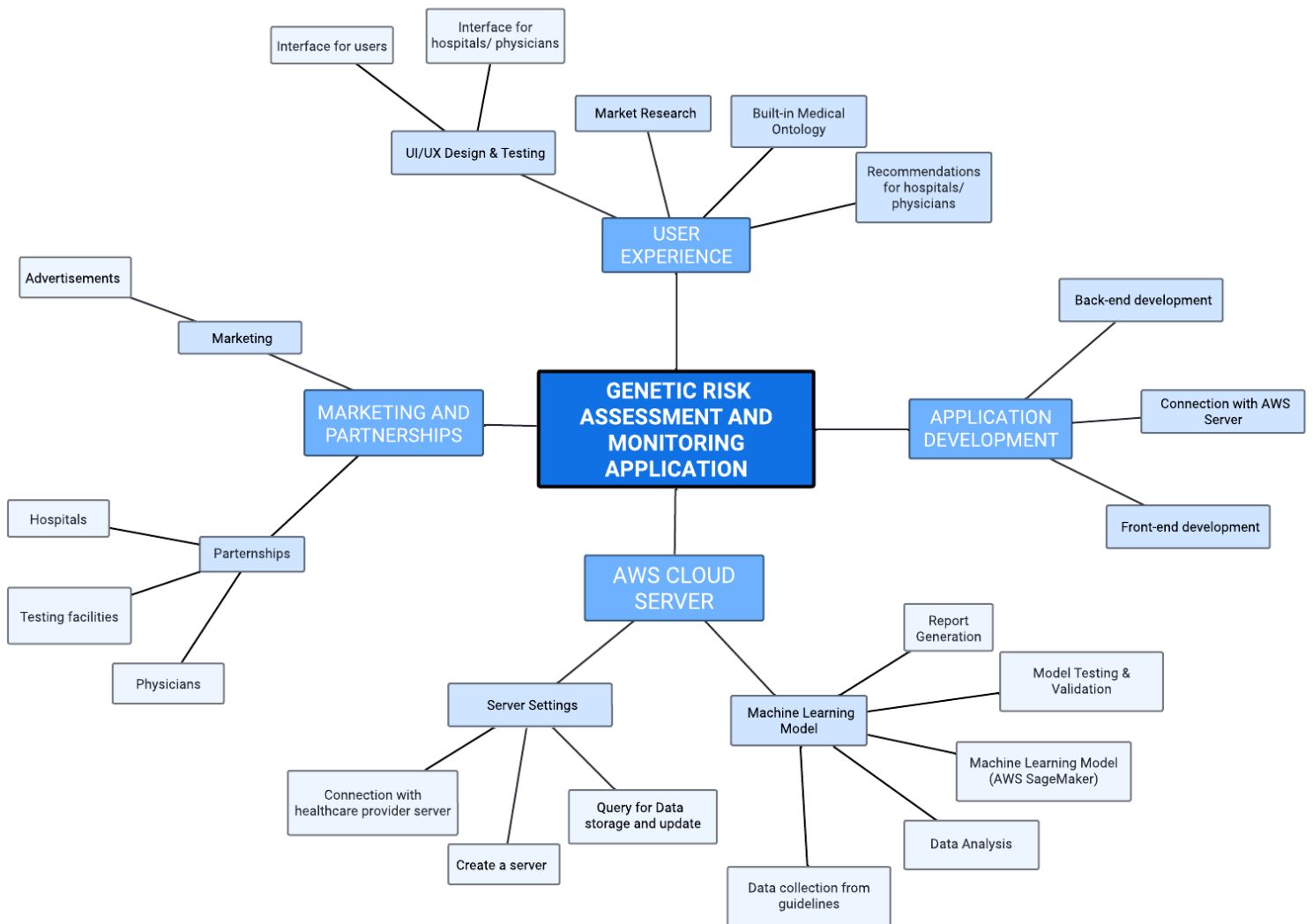

Figure 9: Mind Map

b. UML Activity Diagram

The UML Activity Diagram provides a view on the system describing the actions which are taking place in the process. In the following diagram, we can visually see the steps and the flow of the whole process that takes place in this application.
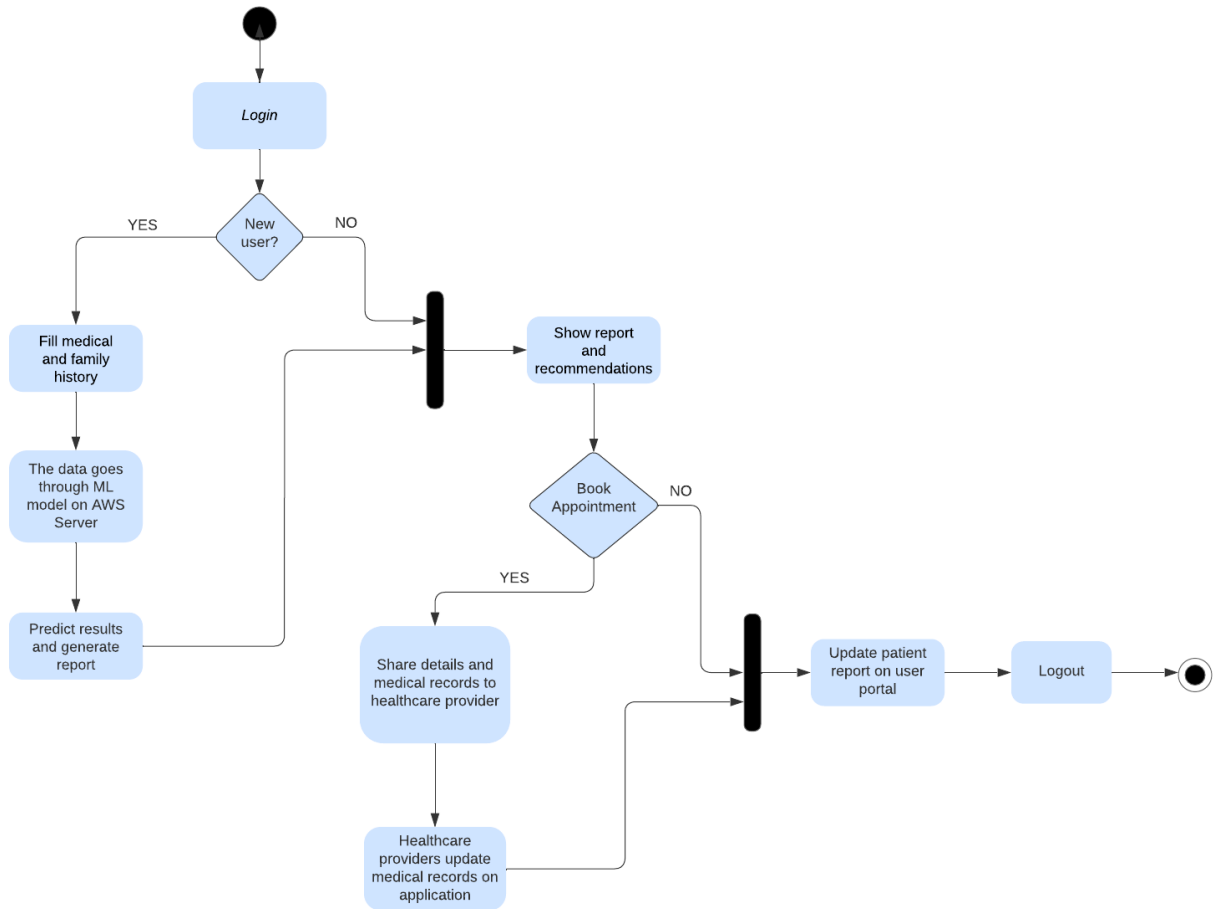

Figure 10: Activity Diagram

c. UML Sequence Diagram

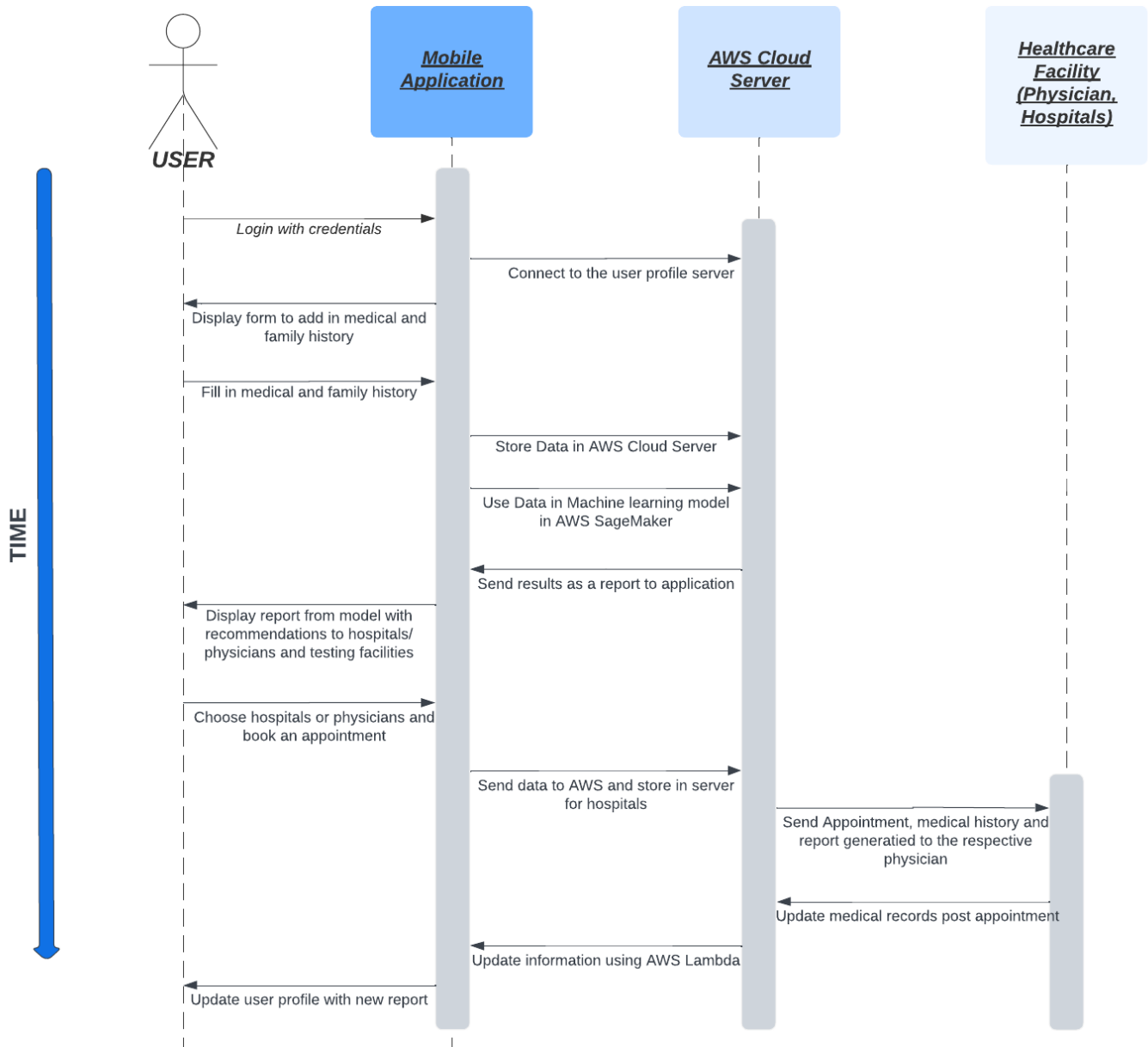The UML Sequence Diagram shows the interactions between the Users, Mobile Application, and the AWS Cloud Server.

Figure 11: Sequence Diagram