

Bishal Ghataney,  
Mark Hanson,  
Thriambak Giriprakash,  
Eric Schmitt,  
Chris Tague,  
Norfinn Norius

## 1.1 PROBLEM STATEMENT

What problem is your project trying to solve? Use non-technical jargon as much as possible.

The project is trying to solve the problem of predicting the occurrence and recurrence of cancer. Despite advancements in medicine, cancer remains a significant challenge due to the limited ability to predict its occurrence and recurrence. This project aims to leverage artificial intelligence (AI) to provide more accurate predictions than human doctors alone. It will involve building and training a simple AI model for cancer prediction, with the goal of improving cancer treatment. The project will also provide training for students on AI, equipping them with skills to build and train a medical AI diagnosis tool using common AI tools and libraries like Tensorflow and Keras. The ultimate goal is to create an AI model capable of recognizing and predicting the risk of cancer occurrence and recurrence, which will be continually improved through training on multiple sets of data provided by leading research institutes and hospitals.

## 1.2 REQUIREMENTS & CONSTRAINTS

List all requirements for your project . This includes functional requirements (specification), resource requirements, qualitative aesthetics requirements, economic/market requirements, environmental requirements, UI requirements, performance requirements, legal requirements, maintainability requirements, testing requirements and any others relevant to your project. When a requirement is also a quantitative constraint, either separate it into a list of constraints, or annotate at the end of requirement as “(constraint)”. Other requirements can be a single list or can be broken out into multiple lists based on the category.

Learning Simple AI principles and tools

- Familiarize ourselves with Tensorflow and Keras' existing AI models
- Learning how the image classification or detection works
- Understanding the frameworks of convolutional neural networks and deep neural networks
- Exploring the concept of transfer learning and its application to enhance model accuracy

Construct a rudimentary AI model fit for our needs

- Find multiple models which are capable of processing medical data
- Stitch them together in a logical sense in which there is a clear path for a set of data to follow from start to finish

Data Preprocessing Tools

- Use data preprocessing tools to clean, normalize, and transform raw data into a format suitable for model training

#### Train the Model

- Use transfer learning to train our model by providing data and expected outcomes. This will take time.

#### Security and Privacy

- Implement robust security measures to protect sensitive cancer-related data privacy regulations

#### Display the model

- Develop a UI interface which can be easily accessible for people to input their data and see results.

### 1.3 ENGINEERING STANDARDS

What Engineering standards are likely to apply to your project? Some standards might be built into your requirements (Use 802.11 ac wifi standard) and many others might fall out of design. For each standard listed, also provide a brief justification.

#### Backend Development Language: Python

- Most AI libraries are python-based

#### Data Format: SQL

- SQL can be easily integrated into python code
- We have the most experience with SQL

#### Front End development Language: JavaScript/ReactJS

- Our team has the most experience with JS and ReactJS

#### General Software Standards

- Clear and concise comments on code
- Properly named variables
- Files named properly
- Files must be placed and go through our shared Git repository (Minimal local work)

#### 1.4 INTENDED USERS AND USES

Who benefits from the results of your project? Who cares that it exists? How will they use it? Enumerating as many “use cases” as possible also helps you make sure that your requirements are complete (each use case may give rise to its own set of requirements).

The goal of the project is to advance research in the area of cancer prediction. This benefits researchers, hospitals, and laypeople. It gives researchers and hospitals a tool to utilize medical information and use it to predict cancer. It helps advanced researchers predict a layperson’s risk of cancer using their basic medical data and gives them the ability to do extra screening and take precautions, and potentially save their lives.

A layperson could also input their own data (images, symptoms) to the algorithm and receive a response as to the likelihood of cancer. However, without proper medical data acquired through actual medical tests, any results produced by the AI would be similar to google searching your symptoms. There will be many false positives, and negatives if a layperson attempts to use this algorithm without proper medical data and counsel.