

#Instructions

- Please make a copy before you edit it: File -> Make a copy.
- Please find the problem statement and detailed template below.
- From where the template starts you will be allowed only 3 pages for the solution summary
- Please submit the final solution document with an access link in the [submission form](#)

Girl Hackathon 2025 - Software Engineering Track

[Do not edit this section. This is read-only]

Amplifying Human Potential using AI

Theme:

AI capabilities can now be used to automate traditionally creative or complex tasks, resulting in enhanced productivity and efficiency across diverse domains. AI can not only produce creative and intelligent multimodal forms of output, but can also orchestrate workflows and perform tasks as an 'agent', in its own right. This can take automation across various verticals to whole new heights and unlock tremendous levels of productivity and efficiency for everyone. This hackathon challenges participants to develop AI-driven solutions that accelerate processes, optimize workflows, and unlock new levels of human potential.

Participants are encouraged to create an application or enhance existing platforms to achieve the following goals: **[Choose 1 of the following]**

Problem Statements: **[Choose 1 of the following]**

- **Finance: A Tax Assistant** that can automate tax filing processes, simplifying complex calculations, identifying deductions, and minimizing errors.
- **Medicine: A Pharmacist's assistant**, that automatically matches orders against handwritten prescriptions and creates orders for patients or **a diagnostic assistant**, a solution that can analyze medical images, patient data, and symptoms to assist healthcare professionals in diagnosing diseases accurately and efficiently (using mock data).
- **Enterprise: Intelligent Process Automation**, AI solutions that can automate repetitive and time-consuming business processes, such as data entry, document processing, and customer service interactions.
- **Developer Productivity: Intelligent IDE** - automated generation of code, tests with advanced debugging capabilities (automagically suggest bug fixes) leading on to largely automated continuous build and integration processes.

Goodluck!

Submission:

Participants are required to create a PDF document as the final submission. The document should contain the link to a public GitHub repository (accessible and open to all).

The repository should have all the collaterals of the code, along with a README file. The code can be written in any open-source programming language using standard open-source libraries.

The README file should cover how to generate the environment needed to run the code, how to run the code, and any other necessary information.

Evaluation Rubrics:

- Potential Impact of Proposed idea (25%)

- Usage of correct DS/Algorithm and AI technique (40%)
- Code Quality (20%)
- Testing (15%)

[Find Template to use below](#)

[\(3 Pages Maximum from the template below\)](#)

2025 Girl Hackathon Ideathon Round: Solution Submission

Project Name: **Heart Disease Detector**

Participant Name: Kukkala Aashritha Reddy

Participant Email ID: aashrithak232@gmail.com

Participant GOC ID: 145047969198

ReadMe File Links (Eg: Github) :

<https://raw.githubusercontent.com/aashritha-kr/Google-Hackathon/refs/heads/main/README.md>

Brief Summary

Please summarize your problem statement and solution in a short paragraph.

Cardiovascular diseases are the highest reason for deaths across the world. Early prediction of the disease may help medical professionals to take required actions for curing the disease. Several factors like RestingBP,Cholesterol ,RestingECG,gender,age and medical history of the patient are considered. This information is used to evaluate the chances of survival of the patient.

Medicine: A diagnostic assistant, a solution that can analyze medical images, patient data, and symptoms to assist healthcare professionals in diagnosing diseases accurately and efficiently (using mock data).

The project is an implementation of this problem statement. Here the mock data is restricted to heart disease.

Problem Statement

What are you doing, why, and for whom?

This project reads the key indicator values of heart disease either via manual user entry in a form or by uploading a pdf/image of the medical report or any other document containing these values.

The values are extracted from the report using Tesseract OCR.

The document could also very well just be a pdf file with the user explaining his test values.

After reading these values, it uses a pre-trained ML model to estimate the chances of the patient having a heart disease.

Users don't need expensive medical expertise to use the app. They can either upload a medical report or manually enter test results.

This app is especially beneficial for low-income individuals for whom healthcare is not affordable.

This is also very useful for users who live in remote areas where immediate healthcare is often not easily accessible.

This can also be used by medical professionals as a second opinion.

Design Idea and Approach

A short and sweet overview of your implementation ideas. You don't need to contain every detail of your implementation, and should omit code here specifically. Use a diagram that illustrates your solution when necessary.

You can discuss about this below but not limited to:

- Which technologies will you use?
- What new components will you write?
- What technologies will you use to write them?
- What are the dominant scaling parameters? (data sizes, qps estimates, etc.) Consider the range and maximum values.
- What is the general rollout strategy?
- What are your information security/privacy concerns and how will you address them?

The approach used to generate the algorithm.

Technologies used

Web Framework: Streamlit (for UI)

Machine Learning Model: Gradient Boosting Classifier (from scikit-learn)

Data Processing: Pandas & NumPy (for handling patient data)

OCR & Text Extraction: Tesseract OCR & PDF2Image (for extracting medical report data)

Improvements

Train using a more diverse and larger dataset for more accurate predictions

Expand the functionality to cover a larger range of diseases

Impact

How will the proposed project address a societal challenge, and to what extent? Is the application grounded in research and data about the problem and the solution? Is there a clear plan to deploy the AI model for real-world impact, and what are the expected outcomes?

The proposed project has the potential to revolutionize healthcare and make it more accessible and affordable for everyone.

More research could improve the quality of the database and thus improve the accuracy of the predictor.

Yes, there is a plan to deploy this. Once deployed it can be used by medical professionals as well as patients.

This could change the way healthcare works altogether

Feasibility

Does the team have a well-developed, realistic plan to execute on the proposal? Does the team have a plan to access meaningful dataset and technical expertise to apply AI to the problem? Have they identified the right partners and domain experts needed for implementation?

Yes, multiple hospitals could be contacted and asked for their data regarding how different parameters are affecting the probability of heart disease.

This large dataset could be stored in a database and queries can be used to extract information using a database query language.

Use of AI

Does the proposal apply AI technology to tackle the issue it seeks to address?

Yes, the prediction of disease is centered around AI concepts.

Alternatives Considered

Include alternate design ideas here which you are leaning away from.

Predicting disease based on x-rays and ecg charts.

References and Appendices

Any supporting references, mocks, diagrams or demos that help portray your solution.

Any public datasets you use to predict or solve your problem.

<https://github.com/aashritha-kr/Google-Hackathon.git>

This is the link to the GitHub Repository

<https://raw.githubusercontent.com/aashritha-kr/Google-Hackathon/refs/heads/main/README.md>

This is the link to the readme file

I have used a kaggle dataset for heart disease prediction.

https://drive.google.com/file/d/1qVqXLRdH09h2iWv-LDnVZ2v_HLg3CEF/view?usp=sharing

<https://drive.google.com/file/d/1s4jHv13P2BfRuB1tPKSrSCdObTfWIRF4/view?usp=sharing>

This is the test input pdf used

https://drive.google.com/file/d/1ZY3_uaqRQ2cQjiSfLv-OylXcqbwAHkg8/view?usp=sharing

This is the video link for demo

All of these links are also uploaded as files in the GitHub repo. Please refer to them in case there is any problem opening these links.