# 2   Project Plan

2.1 TASK DECOMPOSITION

In order to solve the problem at hand, it helps to decompose it into multiple tasks and subtasks and to understand interdependence among tasks. This step might be useful even if you adopt agile methodology. If you are agile, you can also provide a linear progression of completed requirements aligned with your sprints for the entire project. At minimum, this section should have a task dependence graph, description of each task, and a justification of your tasks with respect to your requirements. You may optionally also include sub-tasks.

Tasks:

- 1. Get an AI model from Keras and Tensorflow

   - 1.1 Search through the existing AI models and find the one that is the closest to the model we need

   - 1.2 Edit the model to accept and use the data we will get from the client

- 2. Data Collection and Preprocessing

   - 2.1 Clean, modify and extract the data to meet the project needs to make it ready for training

   - 2.2 Divide the data into two sets, training and testing, using an 80-20 split

   - 2.3 Address data privacy and HIPPA compliance when handling medical information

- 3 Train the AI model with data given by the client

   - 3.1 Utilize GPU to speed up model training

   - 3.2 Document the training process and any issues encountered

- 4 Test the AI model with data given by the client

   - 4.1 Validate the accuracy of the model with the data given by the client

   - 4.2 Maximize the accuracy and minimize the latency of the model using the transfer learning

- 5 User Interface and Visualization

   - 5.1 Create a user-friendly interface for healthcare professionals to input patient data

   - 5.2 Generate informative dashboards and reports to aid in decision-making
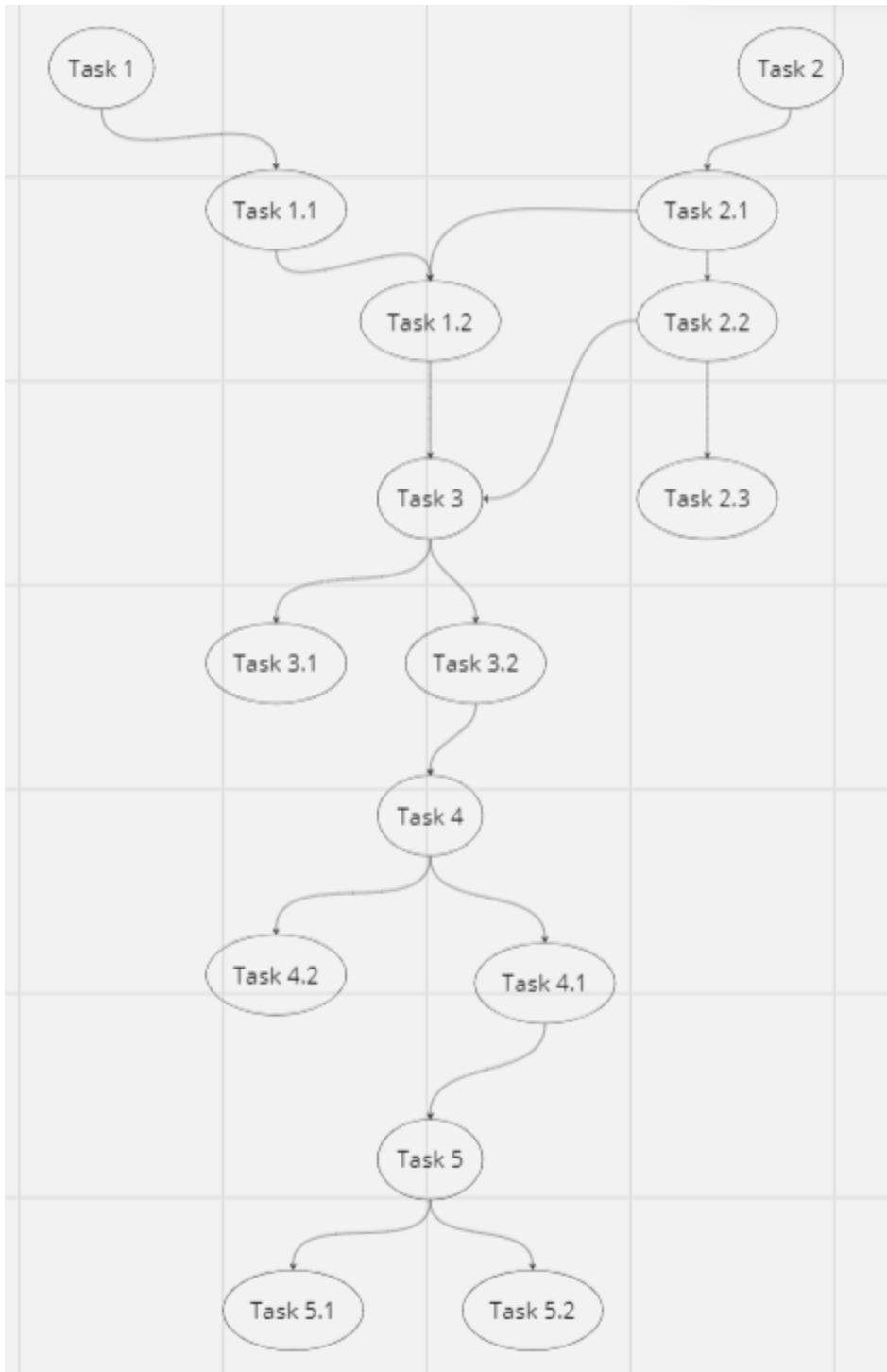
2.2 PROJECT MANAGEMENT/TRACKING PROCEDURES

We will be using the Agile project management style. Our goal is to develop an AI interface that accurately predicts the occurrence and recurrence of cancer. Our Agile management style will

allow us to ensure our AI is utilizing all the important factors for predicting cancer and is giving all the necessary feedback in a comprehensive way.

We will use the Gitlab issue board to track the progress of each part of the overall project. The Gitlab board will have 3 separate sections that we will move the respective story cards into as we continue through the development process. The sections will be "start" "in progress" and "done."

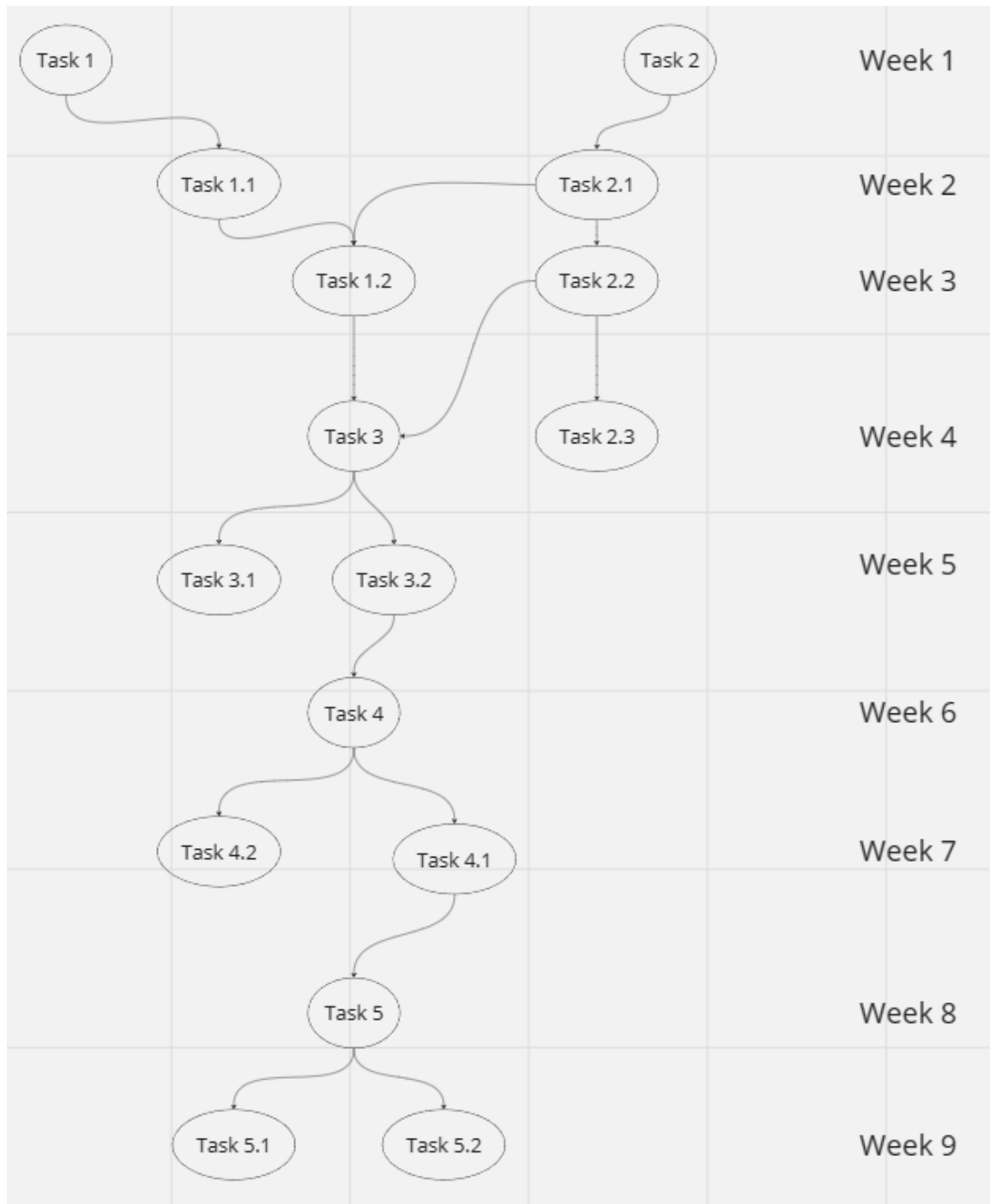## 2.3 Project Proposed Milestones, Metrics, and Evaluation Criteria

What are some key milestones in your proposed project? It may be helpful to develop these milestones for each task and subtask from 2.1. How do you measure progress on a given task? These metrics, preferably quantifiable, should be developed for each task. The milestones should be stated in terms of these metrics: Machine learning algorithm XYZ will classify with 80% accuracy; the pattern recognition logic on FPGA will recognize a pattern every 1 ms (at 1K patterns/sec throughput). ML accuracy target might go up to 90% from 80%.

In an agile development process, these milestones can be refined with successive iterations/sprints (perhaps a subset of your requirements applicable to those sprints).

1: Construct an AI model which is capable of processing a data point from start to finish

- Understand how image classification and detection work.

2: Gather a proper dataset found from either our professor or from open sources, and make any changes necessary to prepare it to be used in our algorithm.

3: Train our algorithm with our data, the algorithm should classify with 90% accuracy with a pattern recognition of 100 ms. Test furthermore.

4: Build a proper Interface for the algorithm. With minimal lag and stress under high loads of processing. We will need to use cloud computing for this. AWS ELB might be of use here.

2.4 Project Timeline/Schedule

• A realistic, well-planned schedule is an essential component of every well-planned project

• Most scheduling errors occur as the result of either not properly identifying all of the necessary activities (tasks and/or subtasks) or not properly estimating the amount of effort required to correctly complete the activity

• A detailed schedule is needed as a part of the plan:

– Start with a Gantt chart showing the tasks (that you developed in 2.2) and associated subtasks versus the proposed project calendar. The Gantt chart shall be referenced and summarized in the text.

– Annotate the Gantt chart with when each project deliverable will be delivered

• Project schedule/Gantt chart can be adapted to Agile or Waterfall development model. For agile, a sprint schedule with specific technical milestones/requirements/targets will work.

	Week 1
Task 1	Task 2
Task 1.1	Task 2.1	Week 2
Task 1.2	Task 2.2	Week 3
Task 3	Task 2.3	Week 4
Task 3.1	Task 3.2	Week 5
Task 4	Week 6
Task 4.2	Task 4.1	Week 7
Task 5	Week 8
Task 5.1	Task 5.2	Week 9

*The above week deadlines might be subject to change due to unforeseen events

2.5 RISKS AND RISK MANAGEMENT/MITIGATION

Consider for each task what risks exist (certain performance target may not be met; certain tool may not work as expected) and assign an educated guess of probability for that risk. For any risk factor with a probability exceeding 0.5, develop a risk mitigation plan. Can you eliminate that task and

add another task or set of tasks that might cost more? Can you buy something off-the-shelf from the market to achieve that functionality? Can you try an alternative tool, technology, algorithm, or board?

Learning how image classification or detection works

- **Risk:** Moderate risk due to complexity in computer vision. This moderate risk is due to the complexity involved in understanding image classification and detection techniques, which are essential for medical image analysis.
- **Risk Mitigation Plan:** To mitigate this risk, allocate more time for learning and provide support from experts in the field.

Understanding convolutional neural networks and deep neural networks

- **Risk:** Moderate risk due to the complexity of neural networks.
- **Risk Mitigation Plan:** Allocate more time for learning and provide access to relevant educational resources. Encourage team members to engage in hands-on practice with neural networks, possibly through coding exercises or small projects, to solidify their understanding.

Finding multiple models capable of processing medical data

- **Risk:** Moderate risk related to the availability and compatibility of models.
- **Risk Mitigation Plan:** Research and select models with a proven track record, consider custom model development if necessary. Conduct thorough research to identify existing models with a proven track record in processing medical data. Ensure compatibility by testing selected models with the project's data.

2.6 PERSONNEL EFFORT REQUIREMENTS

Include a detailed estimate in the form of a table accompanied by a textual reference and explanation. This estimate shall be done on a task-by-task basis and should be the projected effort in the total number of person-hours required to perform the task.

Names/Tasks	Thri	Chris	Eric	Bishal	Mark	Finn
Task 1	2.5hr	2.5hr	2.5hr	2.5hr	2.5hr	2.5hr
Task 2	2.5hr	2.5hr	2.5hr	2.5hr	2.5hr	2.5hr
Task 3	2.5hr	2.5hr	2.5hr	2.5hr	2.5hr	2.5hr
Task 4	2.5hr	2.5hr	2.5hr	2.5hr	2.5hr	2.5hr
Task 5	2.5hr	2.5hr	2.5hr	2.5hr	2.5hr	2.5hr

Task1: Get an AI model from Keras and Tensorflow

Task2:Data Collection and Preprocessing

Task3:Train the AI model with data given by the client

Task4:Test the AI model with data given by the client

Task5:User Interface and Visualization

We are expecting a fairly even distribution of effort across the board for all major tasks.

2.7 OTHER RESOURCE REQUIREMENTS

Identify the other resources aside from financial (such as parts and materials) required to complete the project.

- Unlimited platform to train AI with a GPU
- Appropriate initial AI model to start with
- Keras and Tensorflow libraries
- Data set relevant to predicting cancer
- Server to host the website on