AI AGENT FOR IDENTIFYING LEAF DISEASES IN CROPS

PREPARED FOR

Mini Project Proposal (IS 31230)

Department Of Computing & Information Systems

Faculty Of Applied Sciences

Sabaragamuwa University Of Sri Lanka

PREPARED BY

A.I. Hewarathna

17APC3081

APPROVAL OF MINI PROJECT

1. **Index Number:** 17APC3081
2. **Name with initials:** A.I. Hewarathna
3. **Email:** hewarathna@outlook.com
4. **Mobile No:** +94788136641
5. **Title of the Mini Project:** AI Agent For Identifying Leaf Diseases In Crops
6. **Name of the Mentor:** K.M.K.S.Karunanayake
7. **Mentor's designation:** Researcher
8. **Mentor's Organisation:**
9. **Name of the Internal Supervisor:** Mr P Vigneshwaran
10. **Internal Supervisor's designation:**

FOR OFFICE USE ONLY

- **Approved/Not approved**

- **Signature of the Mentor:** ________________________

- **Suggestions:**

- **Date:** ________/________/_______

Index

Acknowledgement

I would like to convey my thanks to my mentor K.M.K.S.Karunanayake and internal supervisor Mr P Vigneshwaran, Lecturer, Department of Computing and Information Systems, Faculty of Applied Sciences, the Sabaragamuwa University of Sri Lanka for their assistance, encouragement, guidance for this project and also for their continuous concern.

As well as I would like to acknowledge Professor S. Vasanthapriyan, Head of the Department and, W.V. Subodhi Kalpani Wasalthilake, Temporary Lecturer, Department of Computing & Information Systems, Faculty of Applied Sciences, Sabaragamuwa University of Sri Lanka, Mini Project Coordinator for her support and help regarding this project.

EXECUTIVE SUMMARY

One of the interesting study fields in agriculture is disease identification from plant photos, for which machine learning principles from computer vision may be utilised. By interfering with photosynthesis, leaf spot diseases harm plants and shrubs. Most leaf spot diseases impact just a tiny fraction of the tree's total leaf area and pose only a minimal threat to the tree's health. Leaf spot infections should be handled seriously if they cause moderate to full leaf loss for the second to fourth year in a row. Leaf loss across multiple growing seasons can lead to decreased development and greater vulnerability to pests and other diseases. Many leaf spot diseases affect a wide variety of native and ornamental plants and shrubs. Many leaf spot diseases have comparable biology and hence therapeutic possibilities.

Because of the trickiness of clearly identifying diseases, I decided to present a system for the detection and classification of rice diseases based on the images of infected plants. Plant diseases are one of the factors contributing to the decline in the quality and quantity of crops. Reductions in these characteristics might have a direct impact on a country's total agricultural yield. The main issue is a lack of continuous plant monitoring. Sometimes novice farmers are unaware of illnesses and their incidence periods. In general, diseases can strike any plant at any moment. Continuous monitoring, on the other hand, may help to avoid illness transmission. The diagnosis of plant disease is critical in agriculture. This AI agent seeks to apply Machine Learning and Image Processing techniques to the problem of autonomous disease detection and classification in rice, cassava, cherry, corn, grapes, potato, soybeans, strawberry, and tomato plants.

1. Project Objectives

There are a few challenges modern farmers or/and agricultural associations are facing. Which are climate changes, diseases, labour limitations and so on. So agriculture has to use as much technological advancement to move towards precision agriculture. With this project, I propose a system powered by artificial intelligence to identify and classify leaf diseases. Because content awareness about crops needs to identify these diseases and a reasonable amount of knowledge about diseases is much needed. With this AI agent, we can achieve those objectives to secure corpses and avoid transmission.

By uploading an image or a video to agents user's can get to know if their crops are affected or not and if they are affected what are the diseases and how much danger it is and what are the possible solutions for them. Also, it provides a lot more insight into the detected diseases and also plants. Such as what is suitable for the environment, what will be a good season for cultivating that plant and also the market situation of that crop.

Disease detection can be done on rice, cassava, cherry, corn, grapes, potato, soybeans, strawberry, and tomato plants and each plant has several disease groups. Which are,

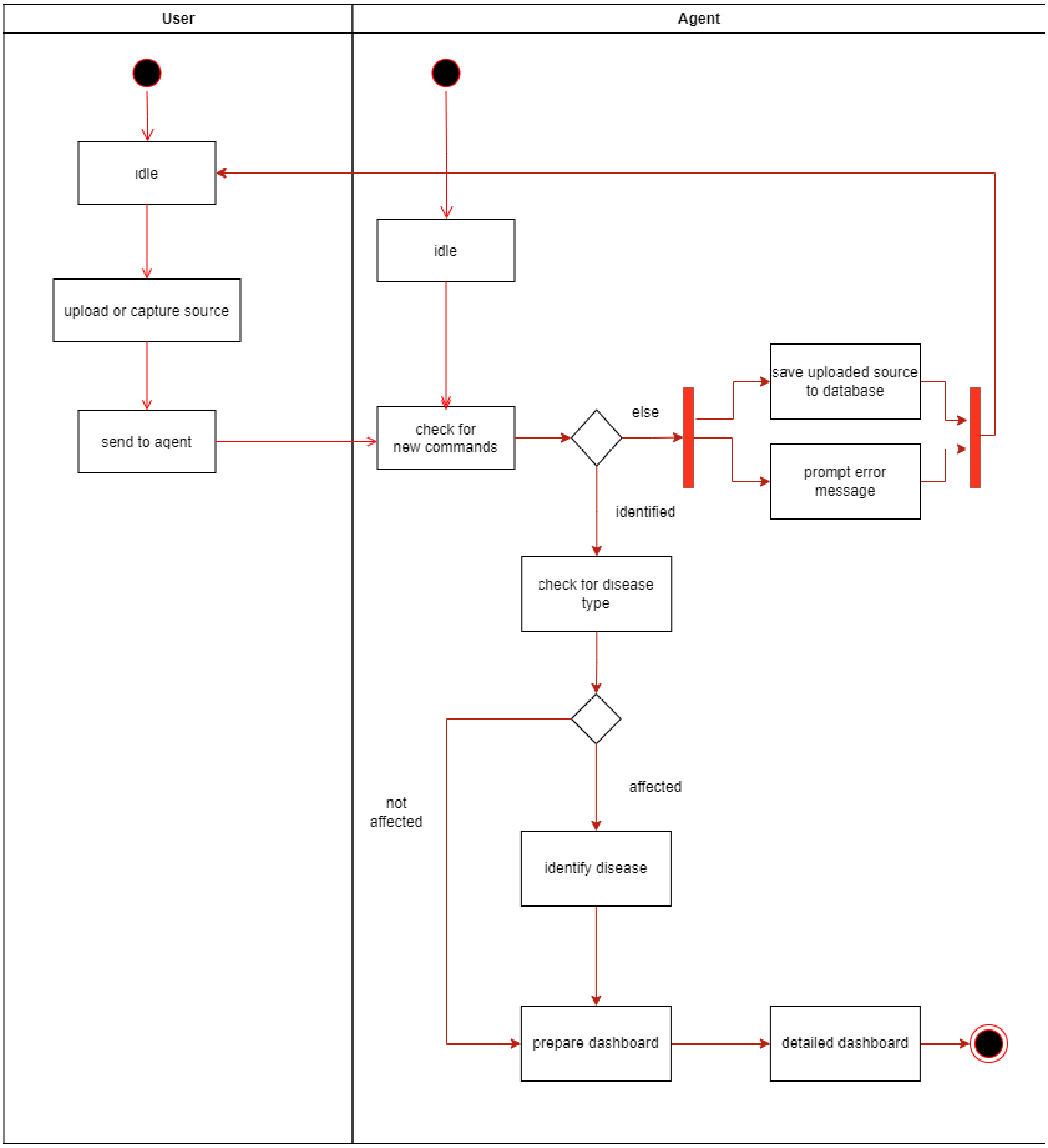
plant	diseases
Rice	Bacterial leaf blight
	Brown spot
	Leaf smut
Cassava	Bacterial Blight
	Brown Streak Disease
	Green Mottle
	Mosaic Disease
Apple	Apple scab
	Black rot
	Cedar apple rust

Cherry	(including_sour) Powdery mildew
Corn	(maize)Cercospora leaf spot Grey leaf spot
	(maize) Common rust
	(maize)Northern Leaf Blight
Grape	Black rot
	Esca (Black_Measles)
	Leaf blight (Isariopsis_Leaf_Spot)
Bell pepper	Pepper,_bell Bacterial spot
Potato	Early blight
	Late blight
Strawberry	Leaf scorch
Tomato	Yellow Leaf Curl Virus
	Mosaic virus
	Target Spot
	Spider mites Two-spotted spider mite
	Septoria leaf spot
	Leaf Mould
	Late blight
	Early blight
	Bacterial spot

Table 1.1

After detecting plant type it shows more details about plant and disease (if affected) such as what are suitable locations for this plant in the country based on historical records, how much water and what nutrients are the nutrients needed for it and suitable seasons to grow based on climate and market prices.

2. Activity diagram



3. Requirements (software & hardware)

- Kaggle kernel
- Pycharm, Dataspell
- Android studio
- Hosting server
- Tensorflow, Keras
- Pandas, NumPy, SciPy
- Opencv
- Python 3.9
- Cuda , cudnn, Nvidia GPU

4. Methodology

- **Data collecting**

Data for these neural networks are images. Those images are collected from several places and they are in different shapes and colour schemas.

- Shayan Riyaz’s rice plant dataset
- Dimitri Oliveira's cassava dataset
- Abdallah Ali’s plant village dataset

Lots of preprocessing are needed before these data are used for training neural networks. They are in different shapes so need to get them to the same common shape and colour schema. Then I need to create a 3d tensor from these data and normalise it and apply appropriate transformations to it.

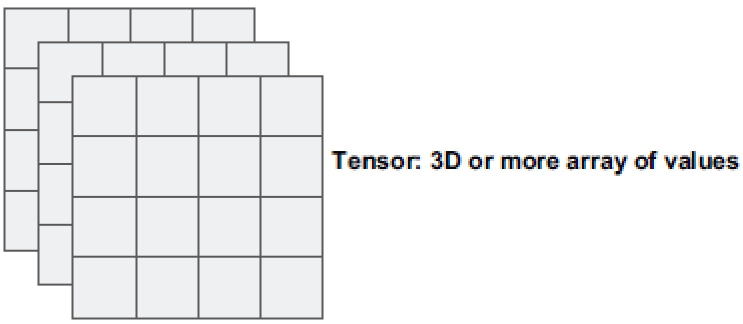


Figure 4.1

- **Create neural network architecture**

Go to solution for these kinds of jobs is a multi-label convolutional neural network but it needs to decide what are the activation functions should use and what are the regulation layers and how many convolutions use inside a layer, and how many hidden layers and nodes need to learn without starving and what is the strategy for generalising the model.

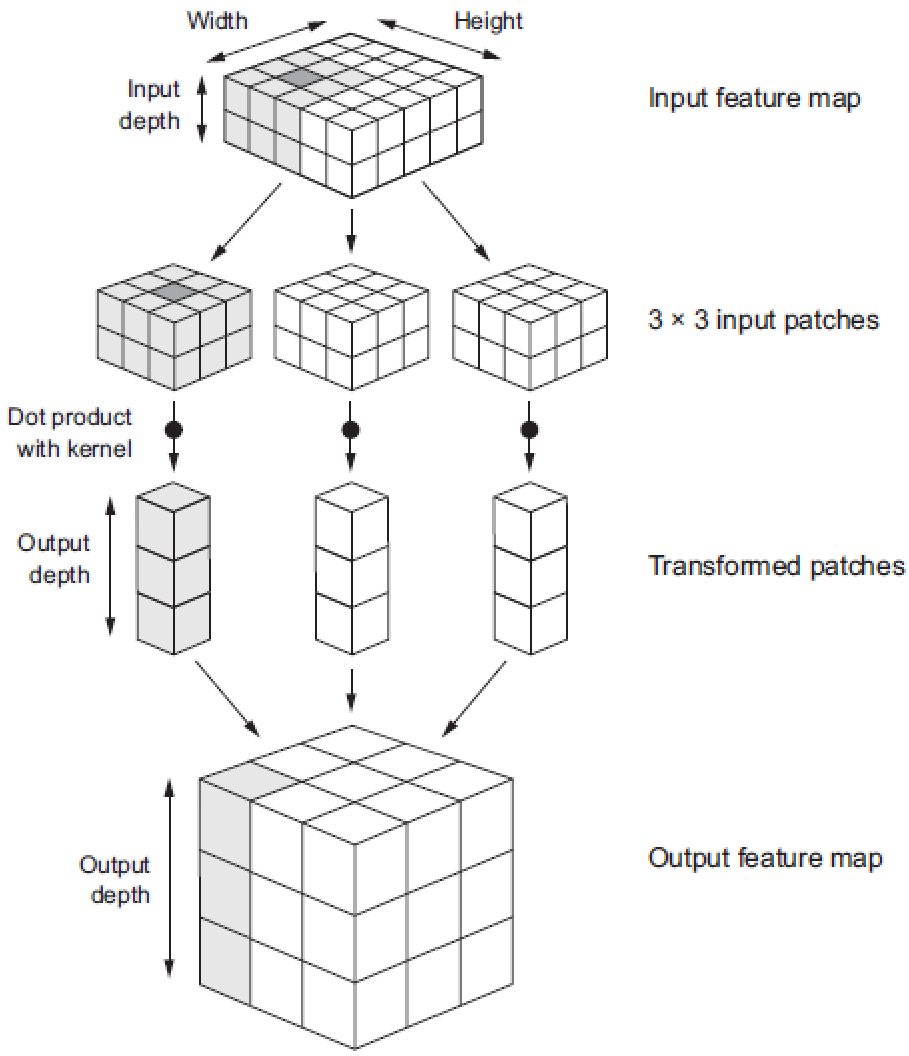


Figure 4.2 - convolution

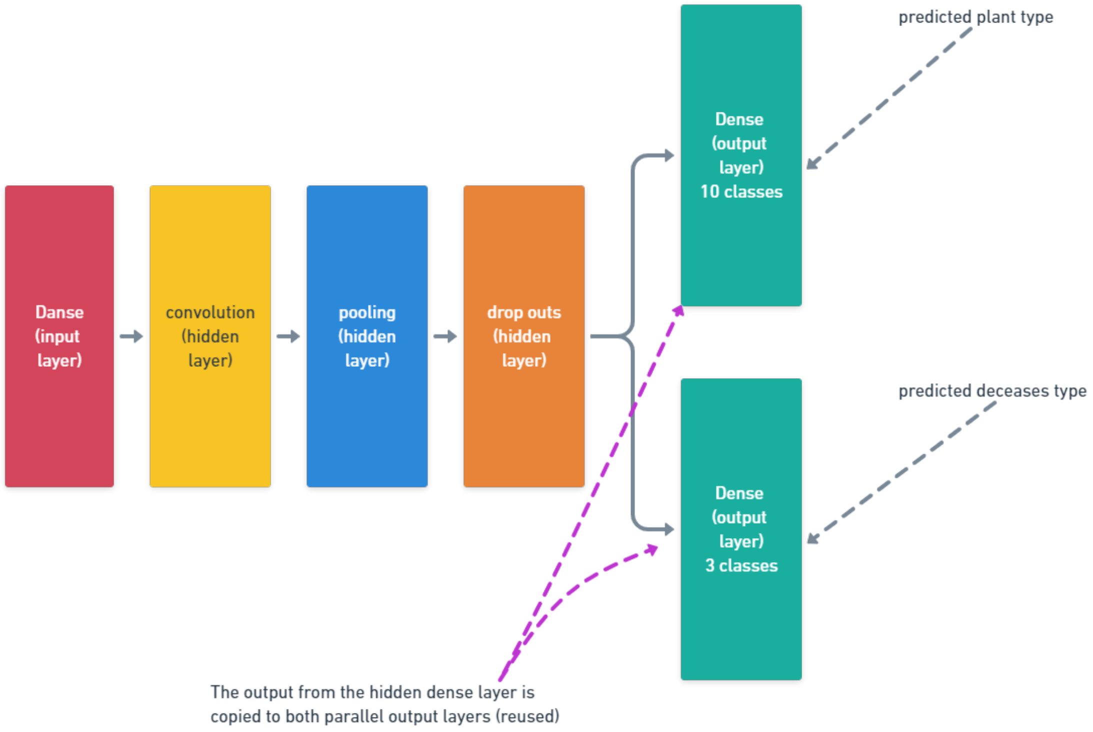


Figure 4.3 - high-level model structure

- **Train, validation and hyper-parameter tune**

The whole data is split into basically two main groups, training and testing. Model train with training data and test metrics on the test dataset. When training, use mini-batches with K-Fold cross-validation to prevent the model from learning unwanted patterns from the validation batch.

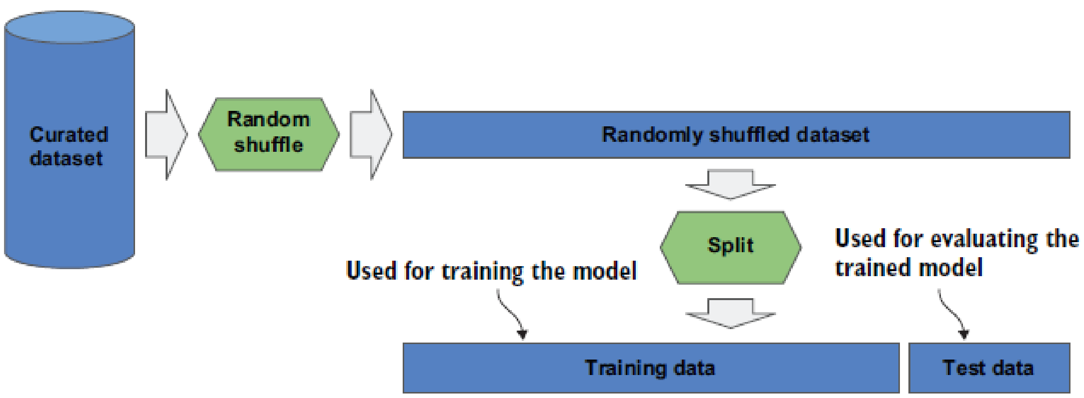


Figure 4.4 - randomly split the whole dataset

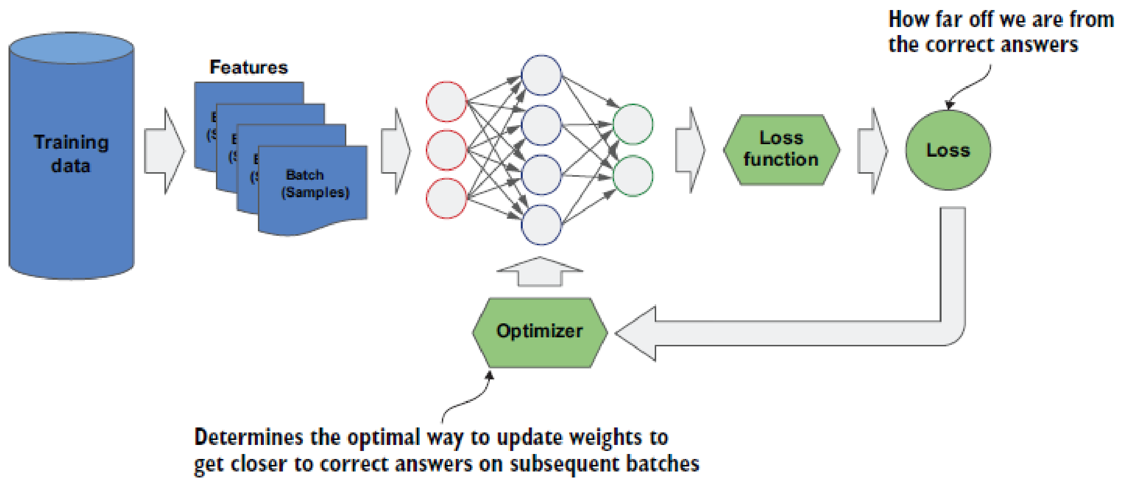


Figure 4.5 - training using backward propagation

Training will be done with backward propagation technique which is first to randomly assign weights to nodes in the neural network and update those weights after each iteration of training according to loss score which derives from loss function and input it to the optimizer to find global minima to loss score in hypothetical space.

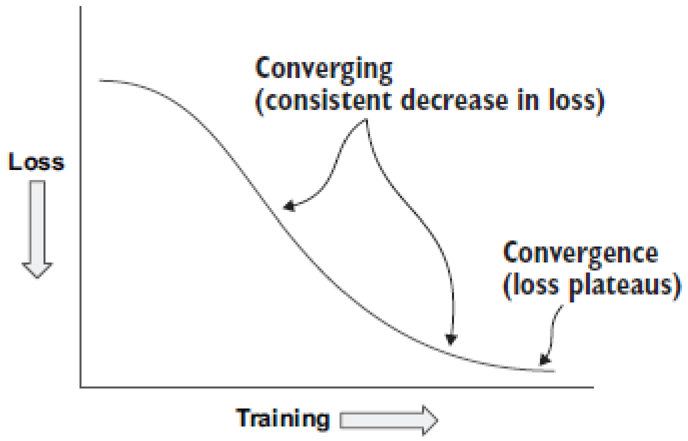


Figure 4.6 - loss vs. training

- **User interface**

UI consists of mainly two sections one is the input getting interface which the user can choose an available option (upload an image or real-time detection) to give inputs and after agent detecting then automatically redirected to the dashboard for more analytical information the user needs to access.

5. The functional requirements of the system

1. Any user can fully access
2. Available as much as plant types
3. Most common disease types
4. Present possible solutions
5. Indicate seriousness level
6. Provide more details about deceases
7. Provide market insights
8. Dynamic dashboards with filters to retrieve much-needed details about plant
9. Captured images saved for further model training

6. The non-functional requirements of the system

1. **Performance -** faster detection (less time to classification)
2. **Accuracy -** achieve accuracy on test data at least 60 - 80%
3. **Minimize FN and FP rate -** reduce false negative and false positive rates

7. Milestones and Reporting

Total estimation of man-hours: 1280

	WEEK														
TASK	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Find data sets															
Pre-process data															
Choose learning strategy															
Design neural network															
Train and evaluation															

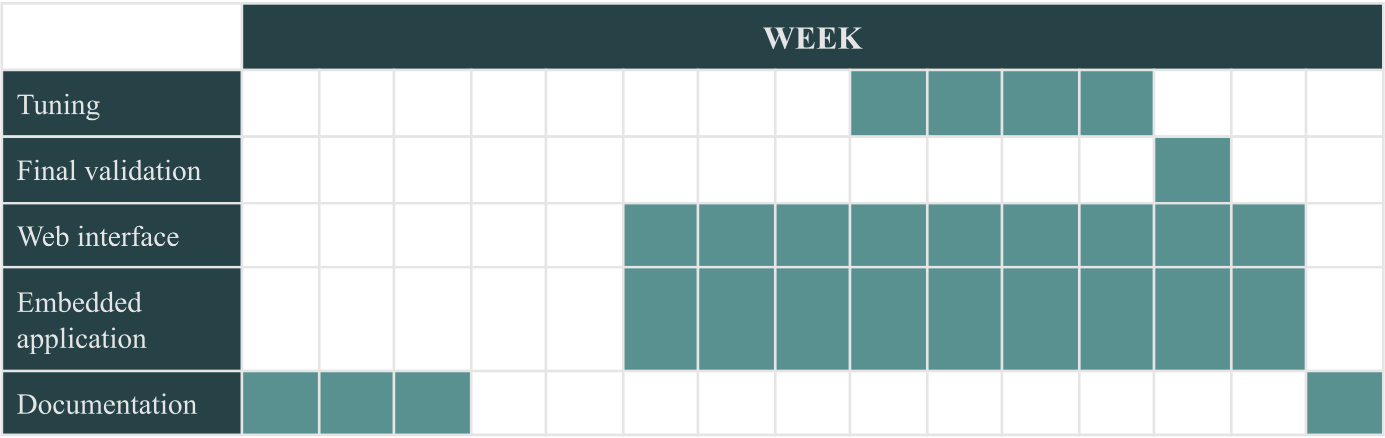


Table 7.1 project milestones