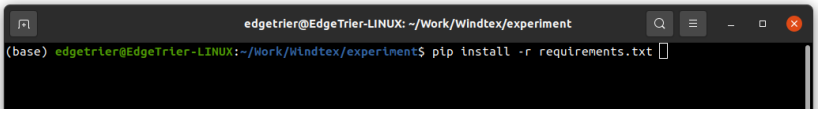
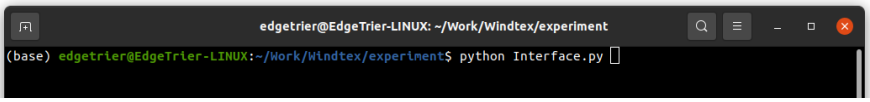
Machine Learning Monkey
Instruction & Information
Version 1.0-alpha

1. **Machine Learning Monkey Installation**
    1. **Python v 3.8+**
    2. **Python Package Prerequisites**
       In terminal, please enter **pip install -r requirements.txt** to install the required packages.

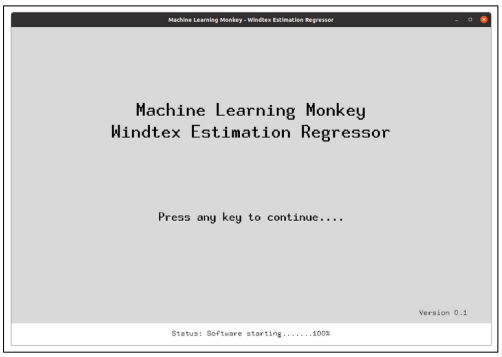


Or manually install the packages by pip/conda
1. numpy
2. scipy
3. scikit-learn
4. scikit-image
5. matplotlib
6. shapely
7. pillow
8. opencv-python
9. tqdm

2. **Start Machine Learning Monkey**
    1. In terminal, please enter **python Interface.py** .

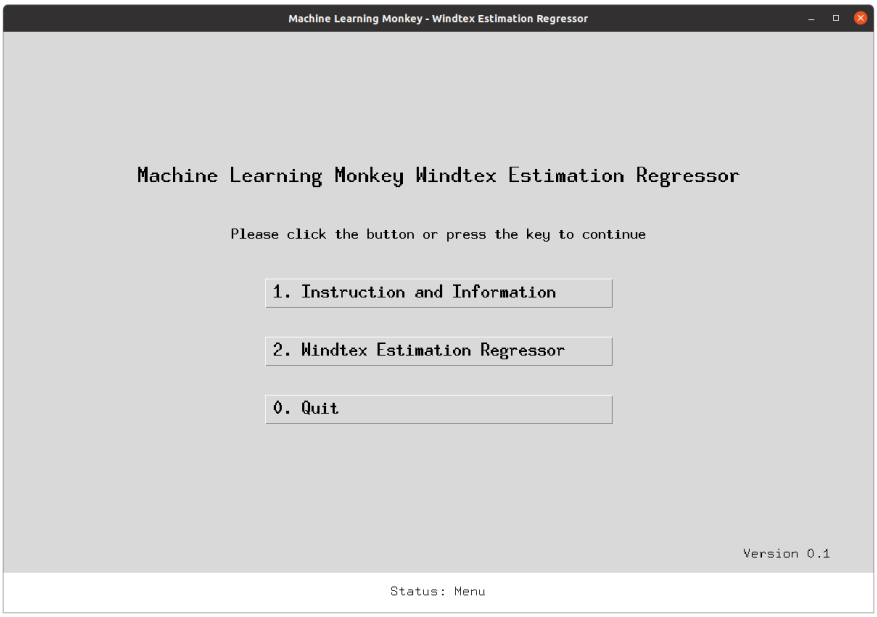


    2. Then, a window will pop-up as shown below



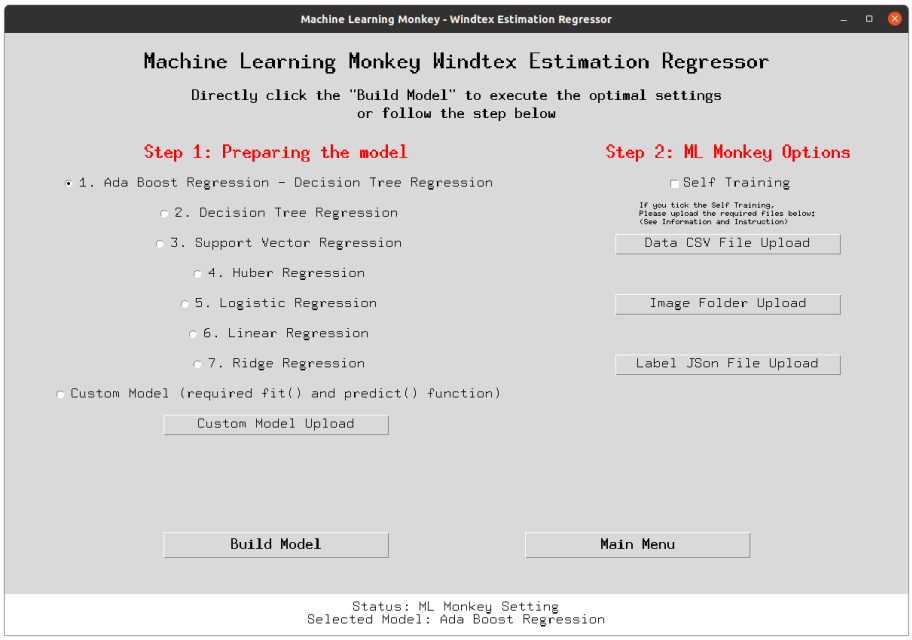
    3. Press **any** key when the loading is completed; **Esc** key to close the window

3. **Menu Page**



1. **Instruction and Information** – open user manual
2. **Windtex Estimation Regressor** – start Windtex Estimation Regressor Model
3. **Quit** – close the windows

4. Windtex Estimation Regressor



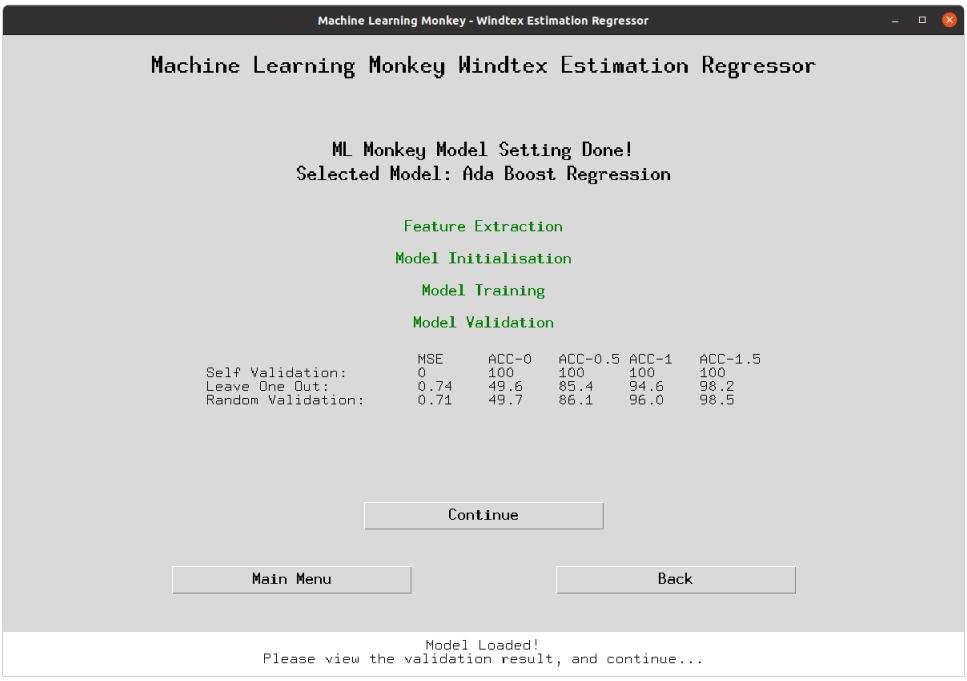
1. The optimal settings have been loaded as default option.
   Please press **Build Model** button to continue.

2. Step 1: Preparing the model
   – The list shows 7 trained models, simply click one of those options and continue.
   – **Custom Model** option: using the uploaded model (custom model needs to have fit() and predict() functions and save as a pickle file) to train and execute regression. Detail please refers to Section 8.

3. Step 2: ML Monkey Options
   – Self Training option: model will be trained with uploaded data, including images, data and annotation label. Detail please refers to Section 9.

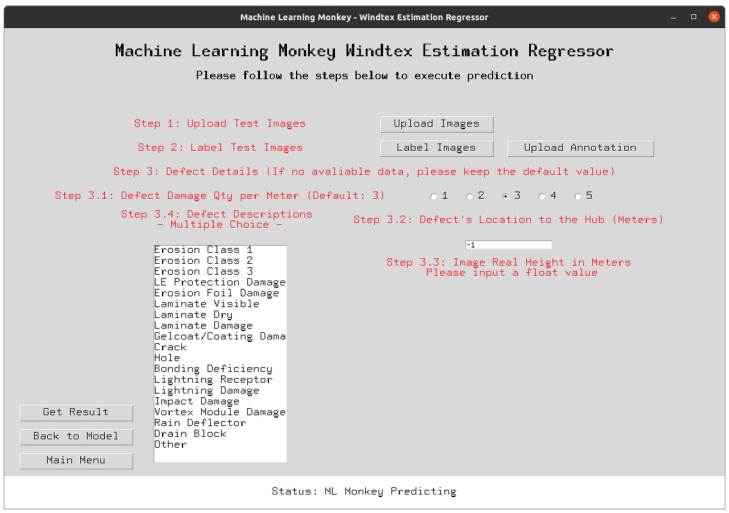
4. Click **Build Model** button to continue.

5. Model Validation Page

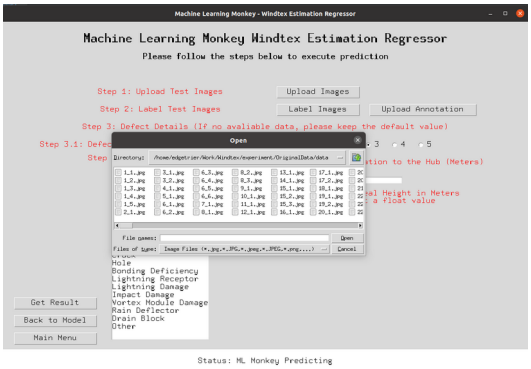


1. There are 4 steps for loading a model
   – Feature Extraction, Model Initialisation, Model Training and Model Validation.
2. When the text turns to **blue** colour, ML Monkey is working on this step.
3. When the text turns to **green** colour, ML Monkey has completed this step.
4. After the completion of **model validation**, a evaluation table will show on the screen.
   1. MAE = Mean Absolute Error
   2. RMSE = Root Mean Squared Error
   3. ACC-0 = Accuracy when considering no errors as true positive predictions
   4. ACC-0.5 = Accuracy when considering ± 0.5 weeks errors as true positive predictions
   5. ACC-1 = Accuracy when considering ± 1 weeks errors as true positive predictions
   6. ACC-1.5 = Accuracy when considering ± 1.5 weeks errors as true positive predictions
5. Click **Continue** button to continue.

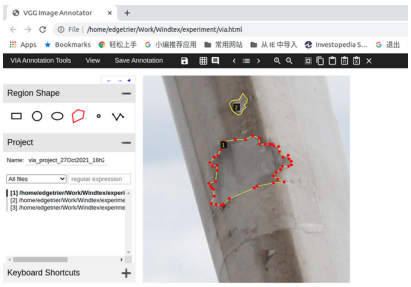
6. Prediction Execution Page

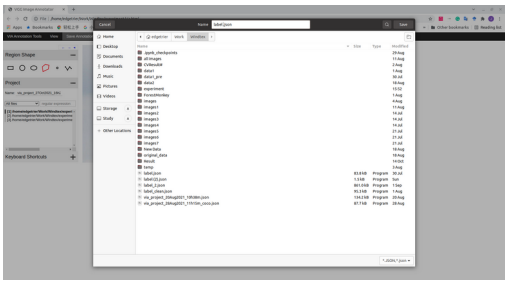


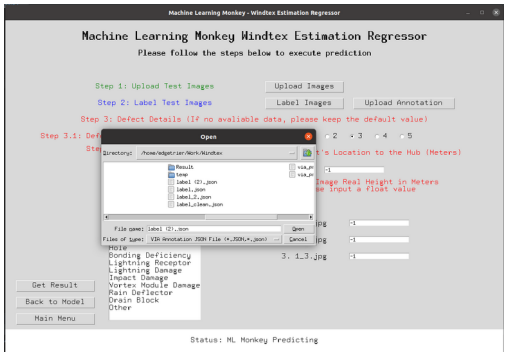
1. Step 1: Upload Test Images – select the testing images from file system



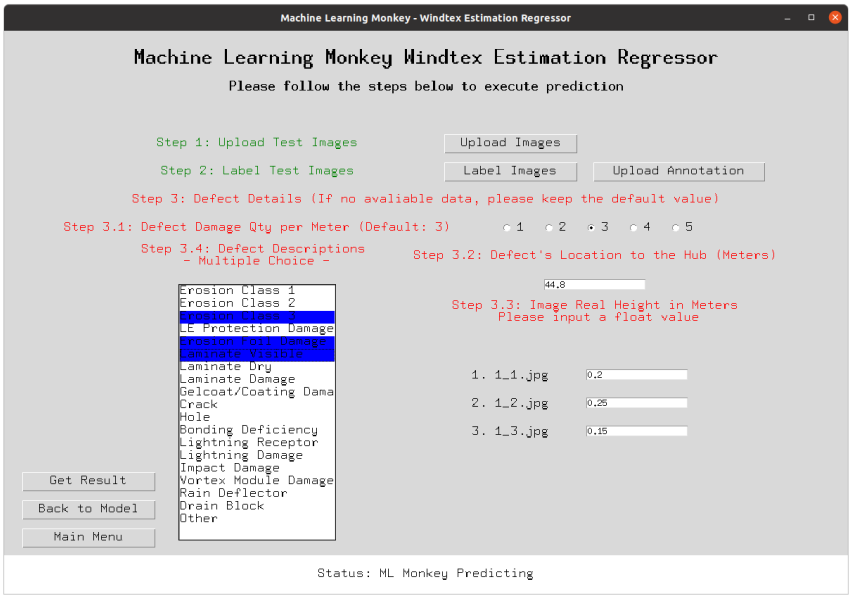
2. Step 2: Annotate Test Images
   – System will automatically open VIA annotation tool;
   – Use polygon to annotate all defects in all images;
   – Save annotations;
   – Upload the annotation file;





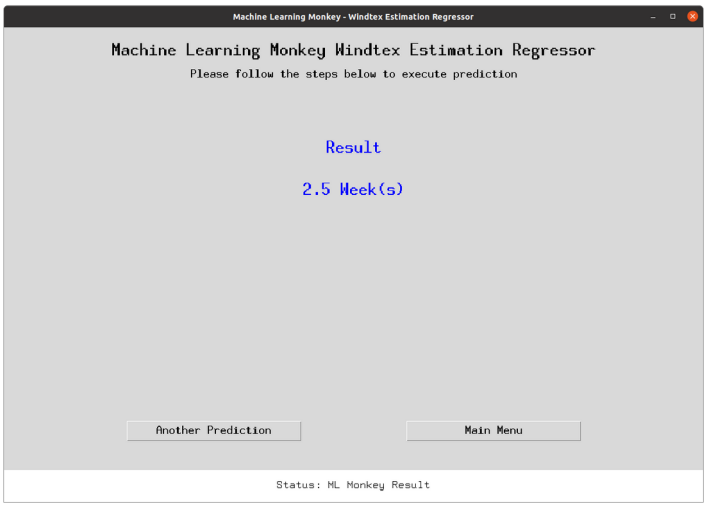


3. Step 3: Input Defect Information – manual input the data from the defect report
   – Defect Qty per Meter (1-2-3-4-5);
   – Defect Location to Hub – positive float (unit: meters);
   – Image Height in real world – positive float (unit: meters): users need to define the real height of each image by their own-decision;
   – Defect Description – multiple selection: select the relative choices, the selected will turn blue colour;



4. Step 4: Click **Get Result** button to continue.

7. Result Page



1. Result will be shown on the screen.
2. Click **Another Prediction** button to test another data with same regression model.

8. **Custom Model**
    1. Custom model is loaded by reading the **pickle file** in ML Monkey.
    2. Custom model can be an algorithm from **sklearn** package.
    3. Custom model also can be other algorithms which contain **fit()** and **prediction()** functions in their built-in classes.
    4. Custom model should be untrained and initialised.
    5. How to save a custom model in **python**:
        1. model = YOUR_MODEL(PARAMETERS)
        2. with open(“name.pkl”, “wb”) as pk_file:
            Import pickle
            pickle.dump(model, pk_file)
            pk_file.close()
    6. Upload to ML Monkey, and the custom model will be trained and validated with the default dataset.

9. **Self Training**
    1. An example is shown in **Example Data** folder.
    2. **Image folder** – All images which are extracted from report
        1. **Filename** - <ID>_<index>.jpg .
        2. **ID** should match with the ID in Data.csv.
        3. **index** should start from 1.
    3. **Data.csv** – All defect information which are extracted from report
        1. Fill the **related values** in each row, one row presents one task.
        2. **ID** must be any integer, but should be **unique**.
        3. **Description** - use “,” to split the multiple descriptions
        4. **Length (meters)** – Real Heights of each image in meter;
            use “+” to split the values of each image (e.g. ID_1+ID_2+ID_3).
        5. **Position** - use “,” to split the multiple positions.
    4. **Label.json –** All defect shape information which are annotated through **VIA Annotation Tools**
        1. Download and Open **VIA Annotation Tools**
        2. Add Files
        3. Annotate all Defects
        4. Export Annotation as **json**