Ali, Aboelela,22398556

Project 1

Leaf Classification

Part 1: Data Preparation.

1. Data cleaning: the data is clean with no missing values or duplicates.

```
df.isnull().sum()

species      0
margin1      0
margin2      0
margin3      0
margin4      0
            ..
texture60    0
texture61    0
texture62    0
texture63    0
texture64    0
Length: 193, dtype: int64
```

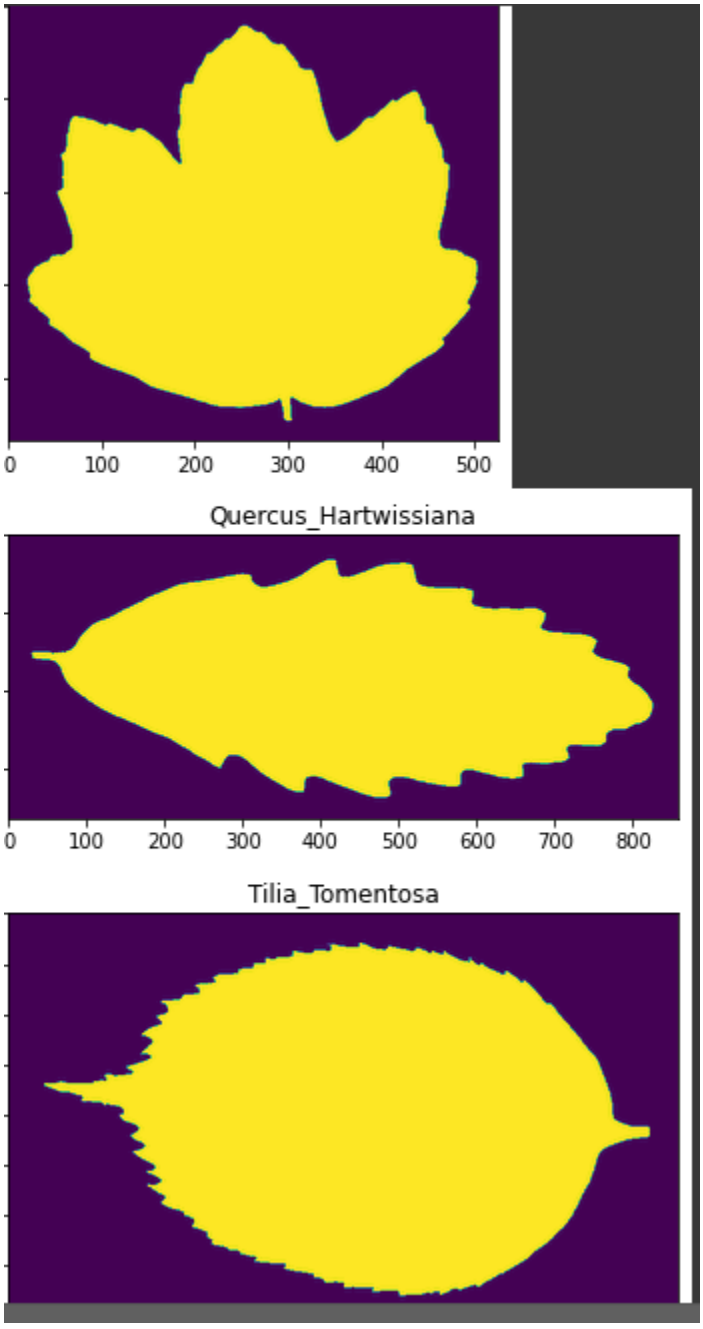
```
df.duplicated().sum()

0
```

2. Preprocessing: the data is already normalized so there is no need for standardization.

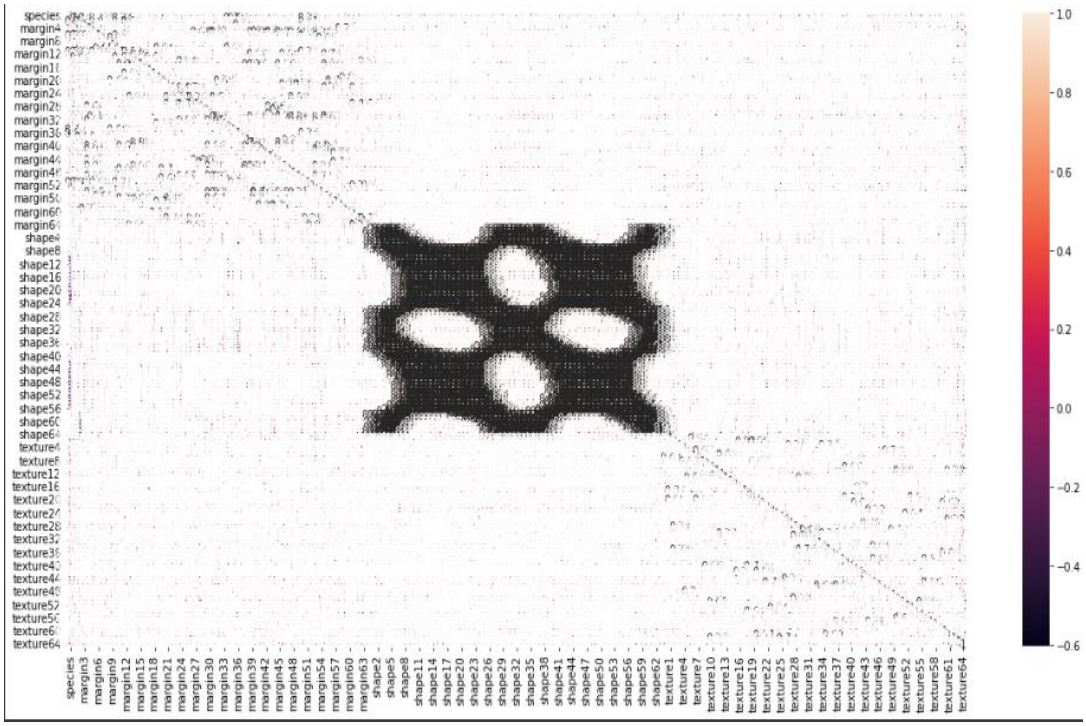
df.describe()

	margin1	margin2	margin3	margin4	margin5	margin6	margin7	margin8	margin9	margin
count	990.000000	990.000000	990.000000	990.000000	990.000000	990.000000	990.000000	990.000000	990.000000	990.0000
mean	0.017412	0.028539	0.031988	0.023280	0.014264	0.038579	0.019202	0.001083	0.007167	0.0186
std	0.019739	0.038855	0.025847	0.028411	0.018390	0.052030	0.017511	0.002743	0.008933	0.0160
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.0000
25%	0.001953	0.001953	0.013672	0.005859	0.001953	0.000000	0.005859	0.000000	0.001953	0.0058
50%	0.009766	0.011719	0.025391	0.013672	0.007812	0.015625	0.015625	0.000000	0.005859	0.0156
75%	0.025391	0.041016	0.044922	0.029297	0.017578	0.056153	0.029297	0.000000	0.007812	0.0273
max	0.087891	0.205080	0.156250	0.169920	0.111330	0.310550	0.091797	0.031250	0.076172	0.0976

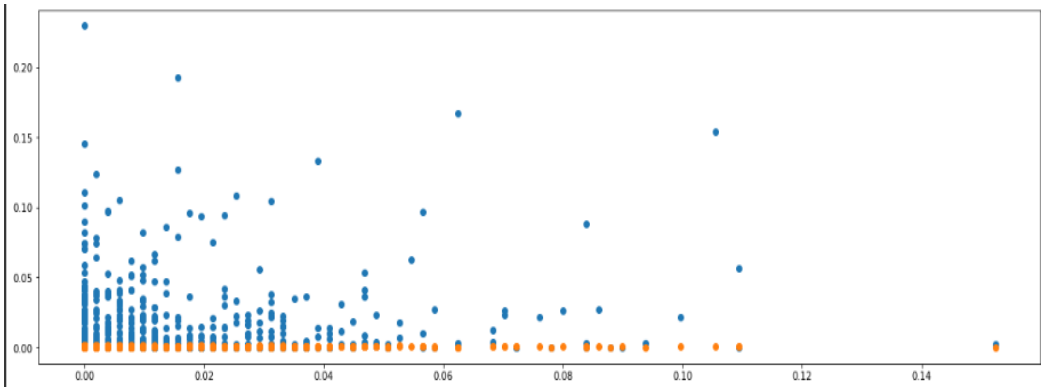
3. Image Drawing.



4. Correlation analysis and Data visualization.



Scatterplot between margin, texture and shape.



Part 2 : Model Training.

1. Tuned hyperparameters :

- hidden units
- Dropout
- Learning rate
- l1_penalty_hidden
- l2_penalty_hidden
- l2_penalty_hidden_bias

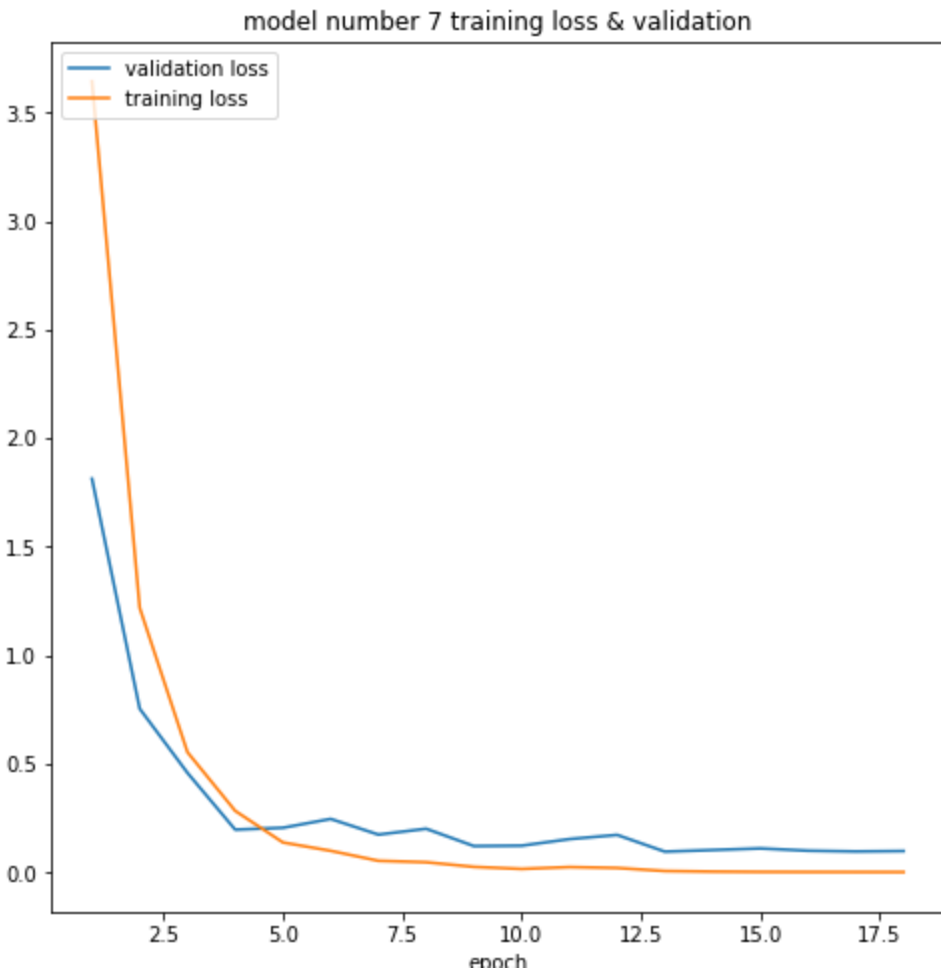
2. Best Hyper parameters.

Learning rate	Best hidden units number	Best hidden units L2	Best hidden L1	Best hidden L2 bias	Best Dropout Rate
0.019587	160	0.0	0.0	0.0030	0.45
0.016519	224	0.0	0.0	0.0030	0.10
0.016370	128	0.0	0.0	0.0015	0.35
0.006996	512	0.0	0.0	0.0045	0.00
0.023062	512	0.0	0.0	0.0060	0.00
0.010809	96	0.0	0.0	0.0030	0.10
0.031854	512	0.0	0.0	0.0090	0.05
0.023469	32	0.0	0.0	0.0000	0.45
0.068585	64	0.0	0.0	0.0000	0.40
0.006447	128	0.0	0.0	0.0045	0.45
0.066470	256	0.0	0.0	0.0015	0.20
0.061102	224	0.0	0.0	0.0090	0.30

3. **Best models.**

```
25/25 [==============================] - 0s 3ms/step - loss: 0.0328 - accuracy: 1.0000
7/7 [==============================] - 0s 5ms/step - loss: 0.1163 - accuracy: 0.9697
25/25 [==============================] - 0s 4ms/step - loss: 0.0087 - accuracy: 1.0000
7/7 [==============================] - 0s 3ms/step - loss: 0.1111 - accuracy: 0.9697
25/25 [==============================] - 0s 3ms/step - loss: 0.0200 - accuracy: 1.0000
7/7 [==============================] - 0s 4ms/step - loss: 0.1090 - accuracy: 0.9848
25/25 [==============================] - 0s 4ms/step - loss: 0.0232 - accuracy: 1.0000
7/7 [==============================] - 0s 5ms/step - loss: 0.1186 - accuracy: 0.9697
25/25 [==============================] - 0s 5ms/step - loss: 0.0058 - accuracy: 1.0000
7/7 [==============================] - 0s 5ms/step - loss: 0.1050 - accuracy: 0.9747
25/25 [==============================] - 0s 4ms/step - loss: 0.0662 - accuracy: 1.0000
7/7 [==============================] - 0s 4ms/step - loss: 0.1830 - accuracy: 0.9646
25/25 [==============================] - 0s 3ms/step - loss: 0.0031 - accuracy: 1.0000
7/7 [==============================] - 0s 4ms/step - loss: 0.1134 - accuracy: 0.9596
25/25 [==============================] - 0s 3ms/step - loss: 0.0911 - accuracy: 0.9962
7/7 [==============================] - 0s 4ms/step - loss: 0.1771 - accuracy: 0.9697
25/25 [==============================] - 0s 3ms/step - loss: 0.0158 - accuracy: 0.9987
7/7 [==============================] - 0s 4ms/step - loss: 0.1482 - accuracy: 0.9596
25/25 [==============================] - 0s 3ms/step - loss: 0.1586 - accuracy: 0.9912
7/7 [==============================] - 0s 4ms/step - loss: 0.2725 - accuracy: 0.9495
25/25 [==============================] - 0s 3ms/step - loss: 0.2160 - accuracy: 0.9470
7/7 [==============================] - 0s 3ms/step - loss: 0.5286 - accuracy: 0.8889
25/25 [==============================] - 0s 3ms/step - loss: 0.0470 - accuracy: 0.9912
7/7 [==============================] - 0s 4ms/step - loss: 0.2778 - accuracy: 0.9444
```

4. **Overall best model is model 7 with accuracy of 98.48%**


model number 7 training loss & validation
validation loss
training loss
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
epoch