------------------------------------------------------------------------------------------------------------------------------------------

Lab Assignment: 1

Full name: Gardyan Priangga Akbar

Student number: 7695883

Email address: [gardyanakbar99@gmail.com](mailto:gardyanakbar99@gmail.com)

------------------------------------------------------------------------------------------------------------------------------------------

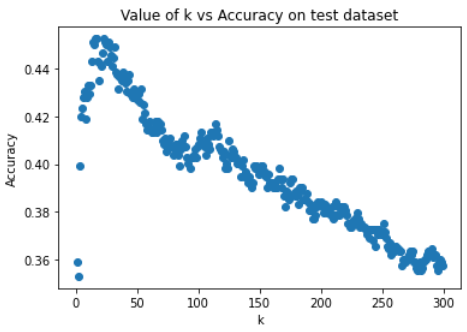
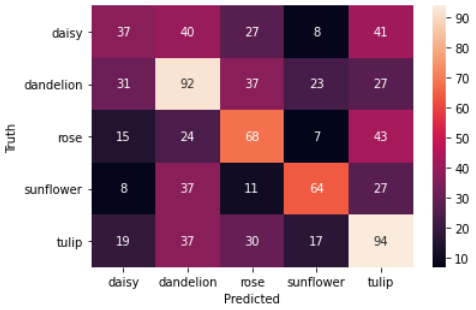
**Flower Image Recognition by KNN, MLP and CNN**

This lab assignment aims to compare and analyze the various image classification approaches, namely KNN, MLP, and CNN. The dataset used for this assignment is a dataset consisting of over 4000 images of flowers spread across 5 different classification labels (daisy, dandelion, rose, sunflower, and tulip).

**KNN**

1. The python modules used for this task are as follows:

|  |  |
| --- | --- |
| Module | Usage |
| Pandas | To convert the resulting confusion matrix into a DataFrame alongside its proper labels |
| OpenCV | To load the flower images and generate the color histogram |
| NumPy | To convert the images into input for KNN |
| Matplotlib | To visualize the color histogram, images, and confusion matrix |
| Sklearn | To split the dataset into training, validation, and test sets, generate confusion matrix, and perform KNN |
| Seaborn | To generate a heatmap of the confusion matrix |
| os | To get list of files from a directory |
| time | To track the time taken to evaluate the test set |

1. I created a list of k to test and then compare them to see which one has the highest accuracy on the validation dataset. Plotting the results visualizes where the value of k is best.   
     
   I did not go beyond 300 for the value of k as the accuracy seem to continue going downhill as shown by the graph. By further utilizing Python’s max() and index() functions I was able to determine the best value for k is 16, with it having the highest accuracy score of 0.4525462962962963.
2. The classification accuracy of the KNN model on the test set was 0.41087962962962965.  
     
   According to the confusion matrix, the model struggles the most in classifying roses and tulips.
3. It took 0.09799885749816895 seconds for the KNN classifier to classify all the samples in the test set.