4. Below is the screen-print of Eigen Values and vectors:

As we check the Eigenvalues = [4.22484077 0.24224357 0.07852391 0.02368303],

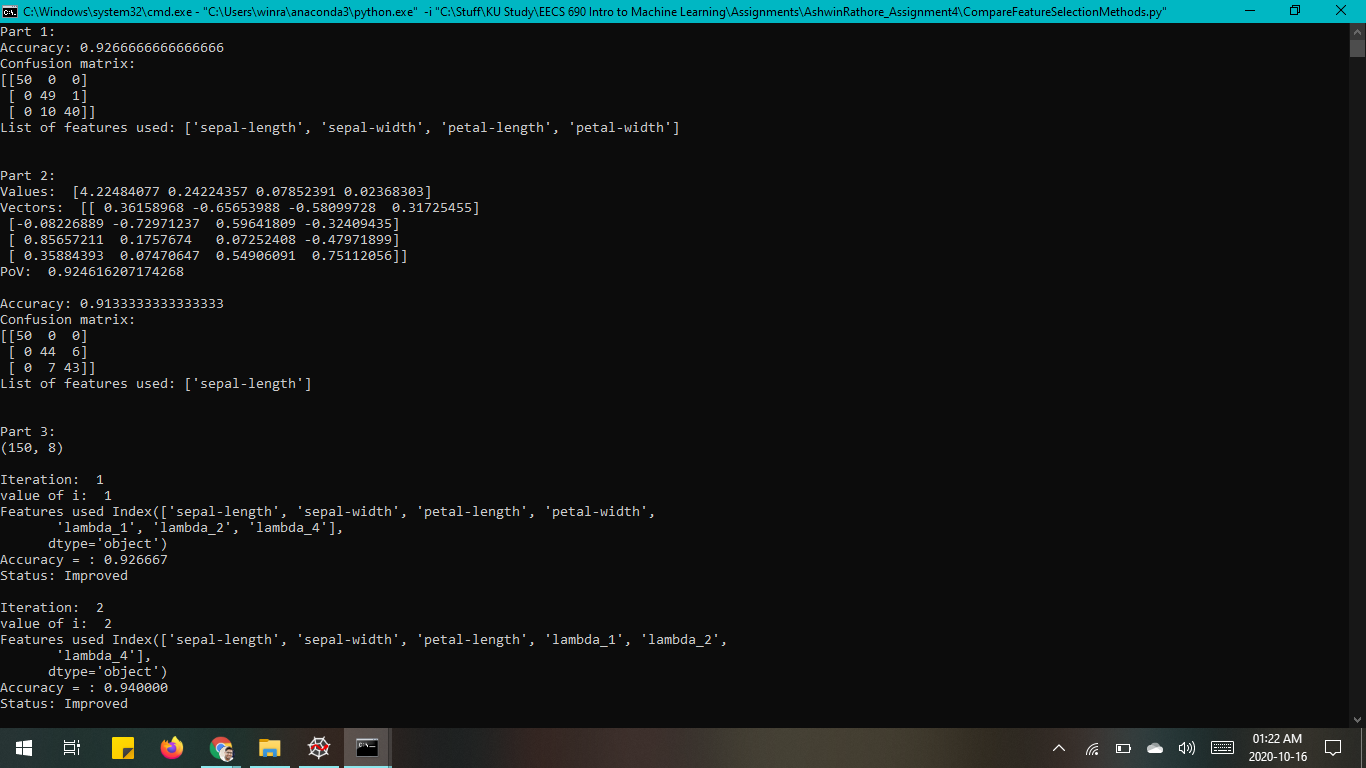
the first one is 4.225 and the POV for that = 4.224/(4.224 + 0.242 + 0.079 + 0.023)

= **0.924**

The program calculated the same value of PoV

So, we can take the first eigenvector to get a PoV > 0.90

i.e. [[ 0.36158968 -0.65653988 -0.58099728 0.31725455]



5. a) Genetic Algorithm method worked best in our program with an accuracy of 96%.

The other three methods gave the accuracy of

Part 1, Original features: 92.67%

Part 2, PCA: 91.33%

Part 3, Simulated Annealing: 93.33

5. b) The Original features gave the accuracy of 92.67% which is not as good as Genetic Algorithm because GA produces “close” to optimal results and it is a non-probabilistic method which makes no assumptions about the statistics of the feature set. Also, GA covers the original set of features if the other set don’t give better results.

PCA assumes the features are linearly correlated while GA doesn’t makes assumptions about problem space. Also, PCA doesn’t go through all kinds of combinations of features like GA does with most of them.

Simulated Annealing doesn’t jump very far and may still find a sub-optimum solution but GA iterates over almost every possible set of features and calculates the best accuracy based on them.

5. c) Features produced in Part 2: ['sepal-length']

Features produced in Part 3: ['sepal-width', 'petal-width', 'lambda\_1']

Features in Part 2 and Part 3 are different because Part 2 just uses the features which are transformed and also it doesn’t iterate over every combination of features while Part 3 uses both the original and transformed features by trying most of the possible combinations of set.

5. d) Features produced in Part 2: ['sepal-length']

Features produced in Part 4: ['lambda\_1', 'sepal-width', 'sepal-length', 'lambda\_4']

Features in Part 2 and Part 4 are different because Part 2 just uses the features which are transformed and also it doesn’t iterate over every combination of features while Part 4 uses the original, and transformed features by trying most of the possible combinations of set using mutation of set of features.