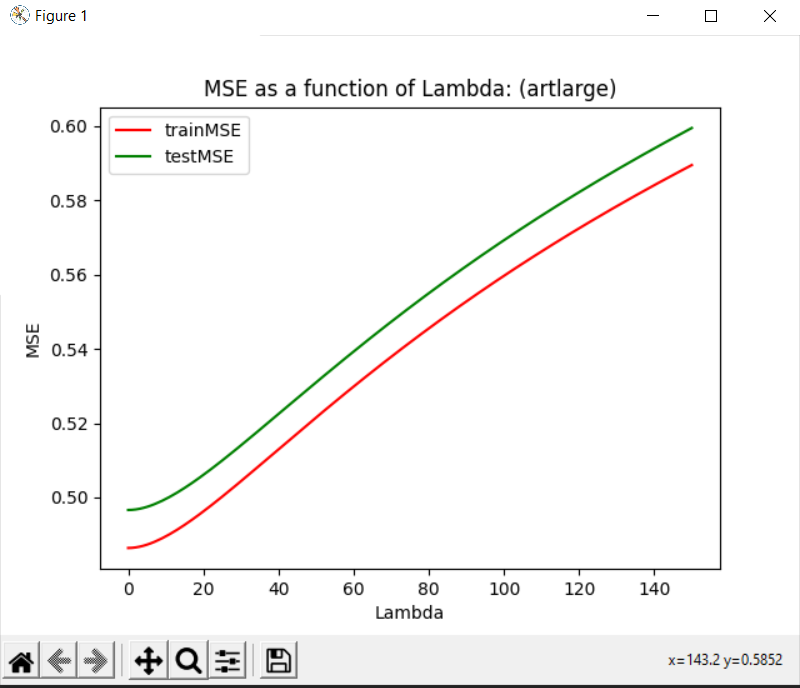
**Programming Assignment 2 Report**

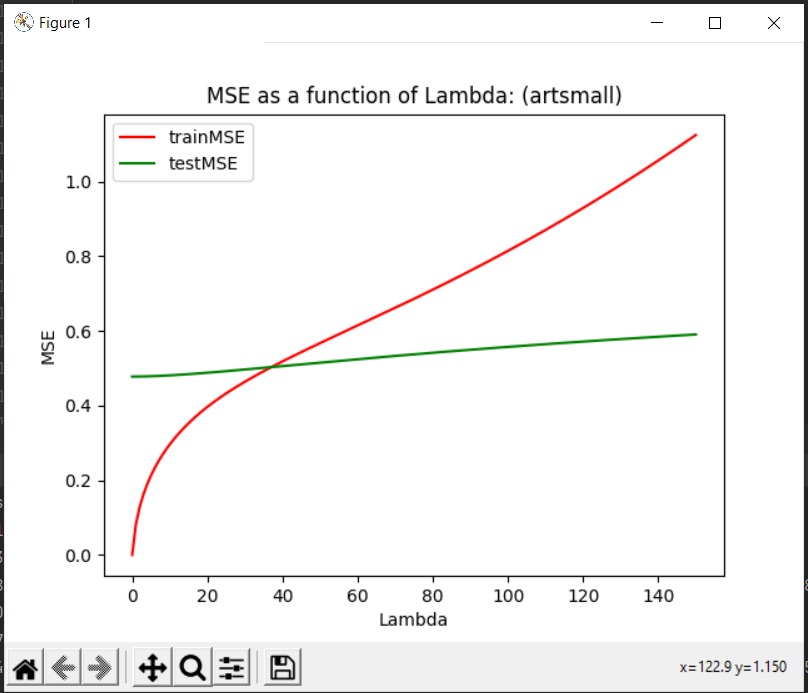
**Part 1: REGULARISATION**

The plots of train and test MSE as a function of Lambda for different datasets are given below.

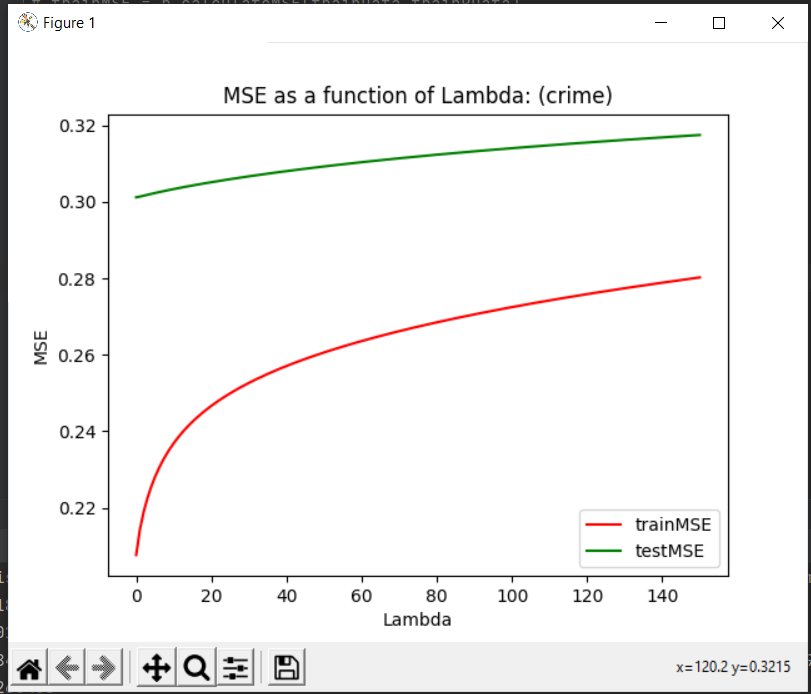
**ARTLARGE:**



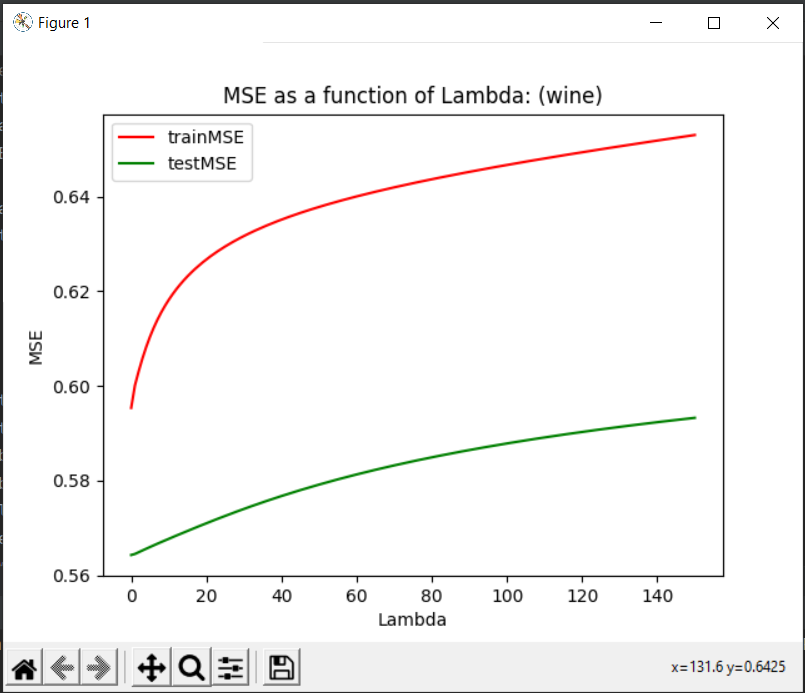
**ARTSMALL:**



**CRIME:**



**WINE:**



Discussions:

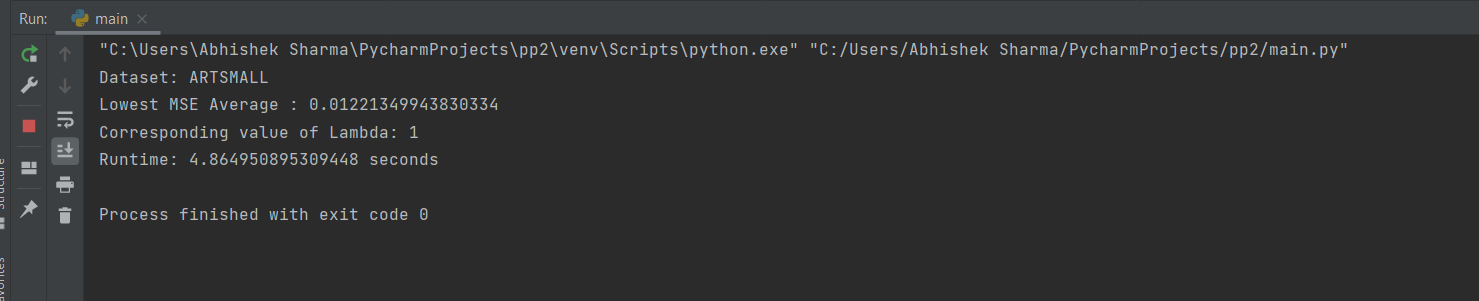
1. Why can’t the training set MSE be used to select λ?

As can be seen in the plots above, the test set MSE is higher than the training set MSE . This means that the test set data is being Over-fitted which causes the MSE to increase. Therefore the training set MSE should not be used to select λ since it will cause over-fitting in the dataset.

1. How does λ affect error on the test set?  
   The test set MSE for each dataset increases on increasing the value of λ, but as can be seen, it doesn’t vary as vastly as the MSE of the train set. Thus, the effects of regularisation are evident.
2. Does this differ for different datasets?  
   Yes, the range of MSE is different for different datasets.

**PART 2: Model Selection using Cross Validation**

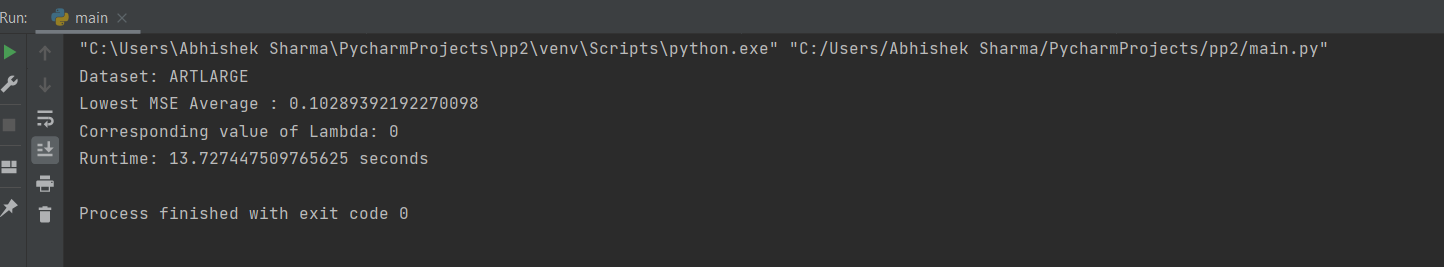
**DATASET: ARTSMALL**



Lowest MSE: 0.012

Corresponding lamda value: 1

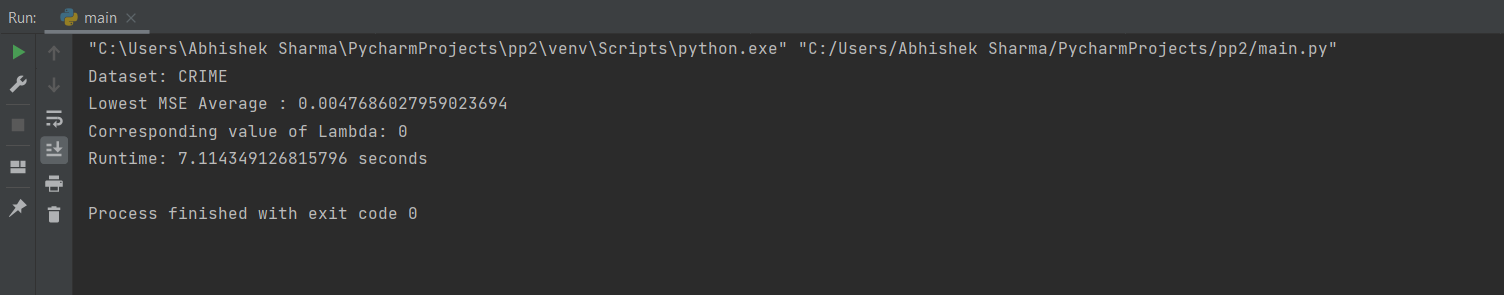
**DATASET: ARTLARGE**



Lowest MSE: 0.102

Corresponding lamda value: 0

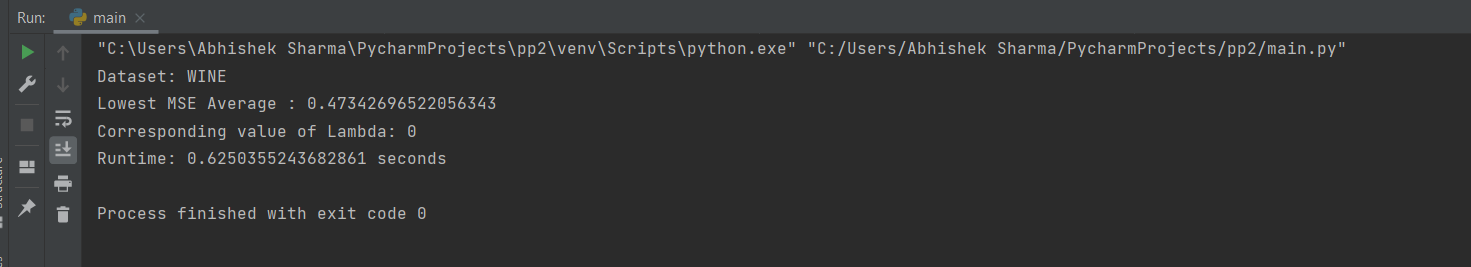
**DATASET: CRIME**



Lowest MSE: 0.004

Corresponding lamda value: 0

**DATASET: WINE**



Lowest MSE: 0.47

Corresponding lamda value: 0

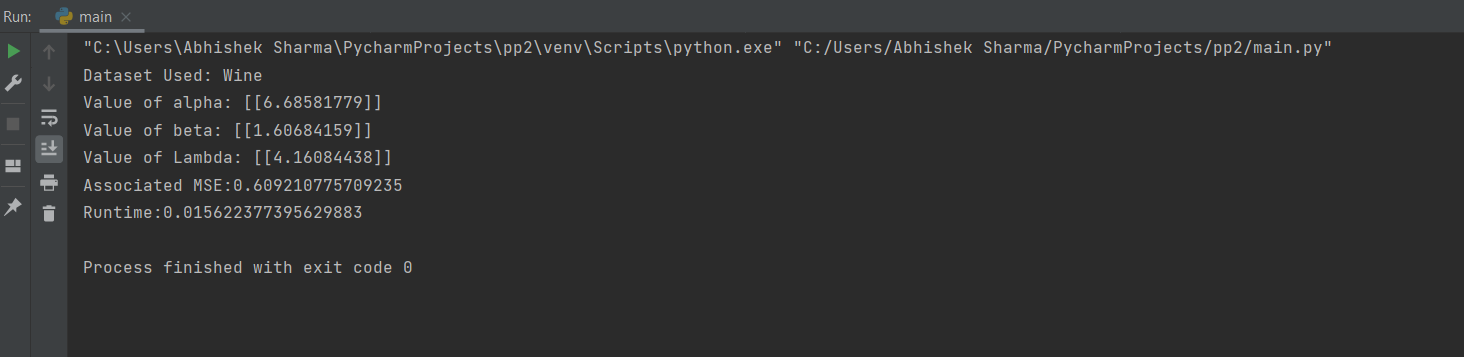
**Observations**:

The MSE in case of cross validation increases as the value of lambda increases. The lower the value of lambda, the lower the MSE.  
However, the runtime increases a lot in case of cross validation. The bigger the dataset size, the runtime is bigger.

Compared to the values of lambda and associated MSE in part 1, the values of lambda and MSE both are lower in case of cross validation.

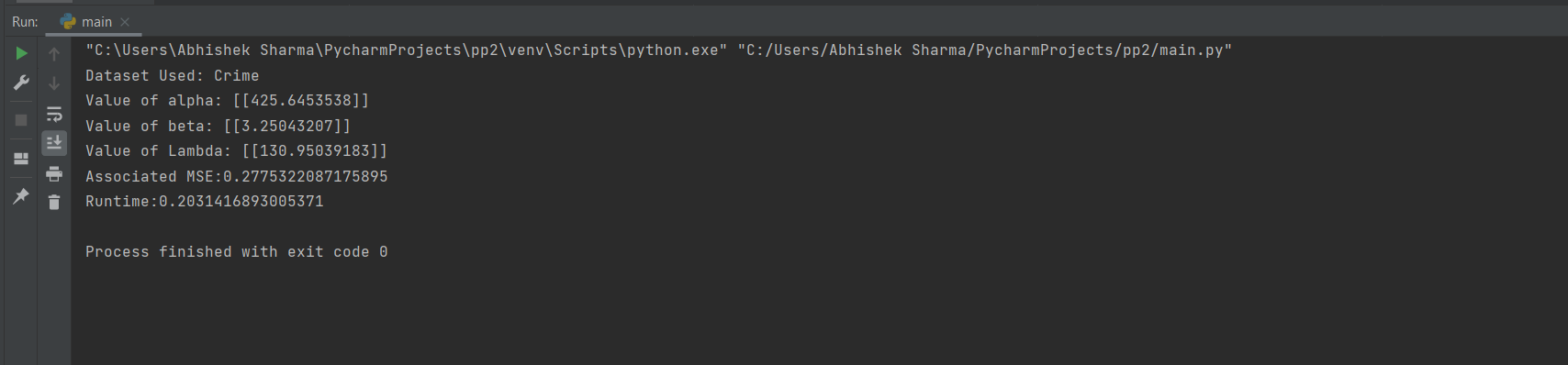
**Part 3: BAYESIAN MODEL SELECTION**

WINE:



The observations are as follows:

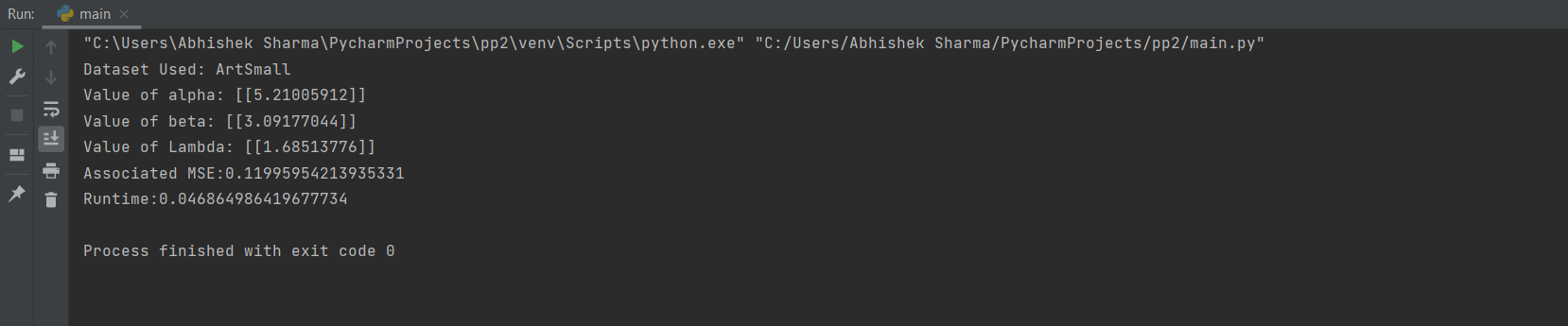
For Wine Dataset, the value of alpha : 6.68, beta = 1.60, optimal λ = 4.16, Associated MSE = 0.6 and the runtime is 0.15 seconds.

CRIME:  


The observations are as follows:

For Crime Dataset, the value of alpha : 425.6, beta = 3.25, optimal λ = 130.9, Associated MSE = 0.27 and the runtime is 0.20 seconds.

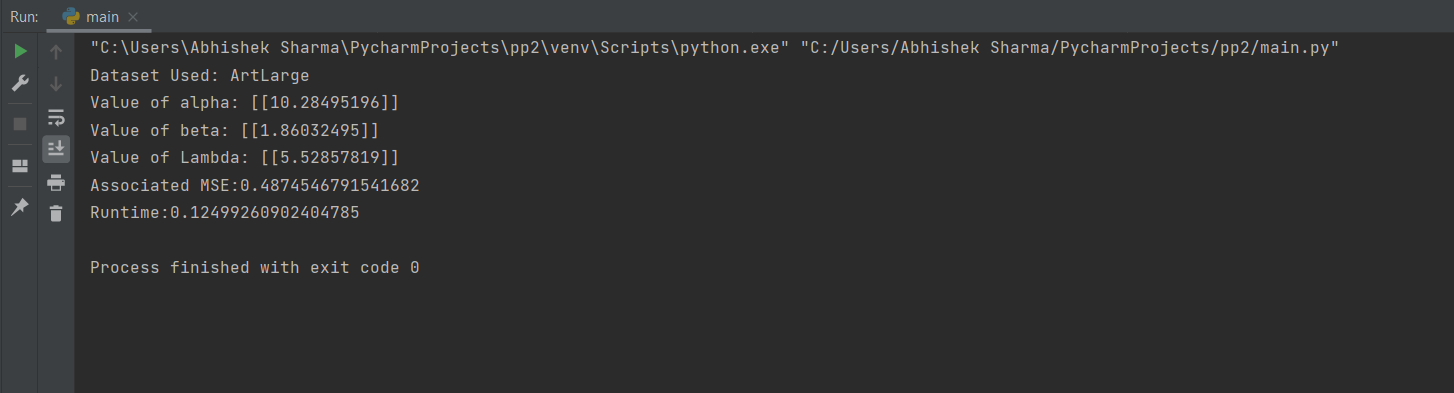
ARTSMALL:



The observations are as follows:

For artsmall Dataset, the value of alpha : 5.21, beta = 3.09, optimal λ = 1.68, Associated MSE = 0.11 and the runtime is 0.04 seconds.

ARTLARGE:



The observations are as follows:

For artlarge Dataset, the value of alpha : 10.28, beta = 1.86, optimal λ = 5.5, Associated MSE = 0.48 and the runtime is 0.12 seconds.

**PART 4: DISCUSSION OF RESULTS**

**Q1. How do the two model selection methods compare in terms of effective λ, test set MSE and run time?**

The two models vary with each other in terms of lambda, test set MSE and run time as follows:

**In terms of lambda:**

Values of lambda in cross fold model remain low compared to other model.  
However, lambda has higher values in Bayesian Model Selection.

**In terms of MSE:**

Test set MSE values are also lower for cross fold method as compared to those in the Bayesian Model Selection.

This shows that cross fold method is better at prevention of the issue of over fitting as is evident from the values of MSE from both the models.

**In terms of Runtime:**

The runtime of cross fold model goes a lot higher than that in the Bayesian model. This is because the whole process of cross fold validation, though results in lower values of MSE, take large amount of time.

**Q2.** **Do the results suggest conditions where one method is preferable to the other?**

The results from this assignment suggest that both the models, namely the cross fold validation model and the Bayesian model selection give low MSE and lambda values when tested on test data. (The values being lower in the cross fold model).

The difference, however lies in the runtime of both the models. For large datasets, the cross fold model can take up large amounts of time which may sometimes result in huge costs. The runtime of the Bayesian Model selection stays below a few mili seconds for the given datasets and therefore results in faster results.

So, we can say that in conditions where getting results FAST is the priority, the Bayesian Model selection may be the better choice.

However, if the priority is to get as accurate results as possible, with minimum MSEs, then the better option would be to go for Cross fold Validation model.