ECE 517 HW 5.2-5.3 Construction of a Regression Algorithm

Joseph Hilland

November 2023

1   Outline

The data that was provided contained a training and test set dataset. Each set consisted of predictor inputs Xtrain and Xtest for regressors Ytrain and Ytest. The data corresponded to the response of y of a physical model to a vector signal x of 19 dimensions. The train and test outputs consisted of 81 samples. The are depicted in figure 1 that was given with the assignment.
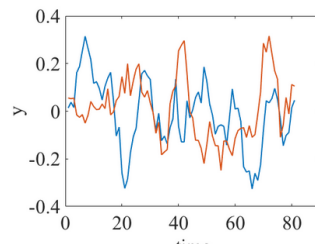

Figure 1: Given dataset

The algorithms listed below were all required to use 20 percent of the training data to validate. The validation procedure consisted of:

- Training the predictor with 80 percent of the training data and give value for gamma (epsilon, nu).

- Running a test for the rest of the training data.

- Computing the mean square error (MSE) of the prediction.

- Repeat for a reasonable range of gamma.

```
Xtr=Xtrain(1:64,:);      % 64 samples is roughly 80% of the samples xtr
Xval=Xtrain(65:end,:);   % Remaining 20% of data
ytr=ytrain(1:64);        % 64 samples is roughly 80% of the samples ytr
yval=ytrain(65:end);     % Remaining 20% of data
```

Figure 2: Matlab Data Seperation

2 Regression

Figure 3 depicts my results, that the optimal value for gamma was to be 0.004.
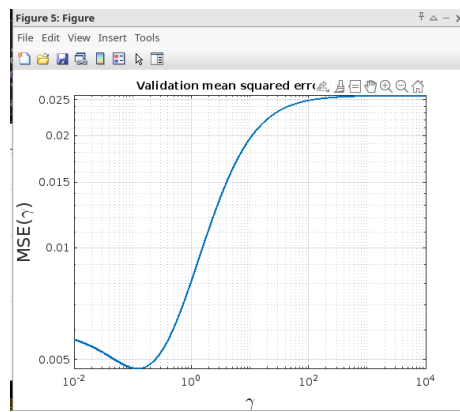As a result, you can see that the MSE based on gamma was 0.008.

Figure 3: Validation MSE Regression

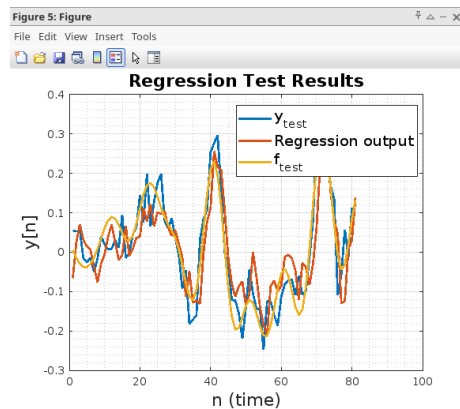Figure 4 shows a graph of the result of the real vs predicted data.

Figure 4: Regression Test Results

3 SVR

The above experiment needed to be repeated utilizing an SVR. After completing the experiments, Figure 5 and 6 depict the MSE validation as well as the SVR test results.
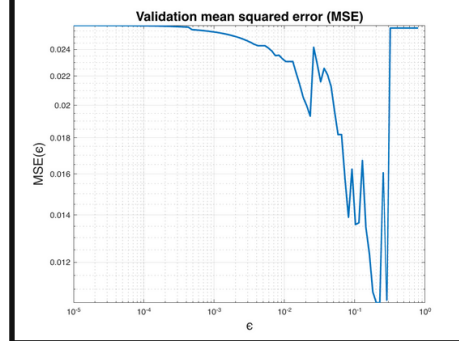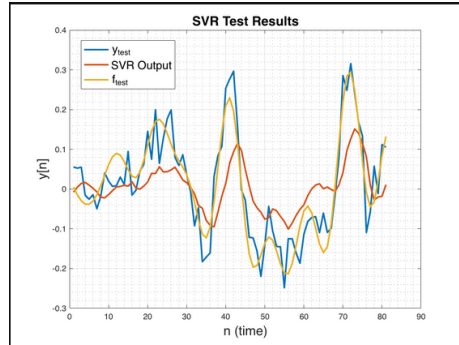

Figure 5: Validation MSE SVR

Figure 6: SVR Test Results

The optimal value for epsilon was determined to be 0.01. C was equal to 5 which led to an MSE of 0.006.

4 nu-SVR

The above experiment needed to be repeated utilizing an nu-SVR. After completing the experiments, Figure 7 and 8 depict the MSE validation as well as the nuSVR test results.

The optimal value for nu was determined to be 0.04. C was equal to 10 which led to an MSE of 0.006.
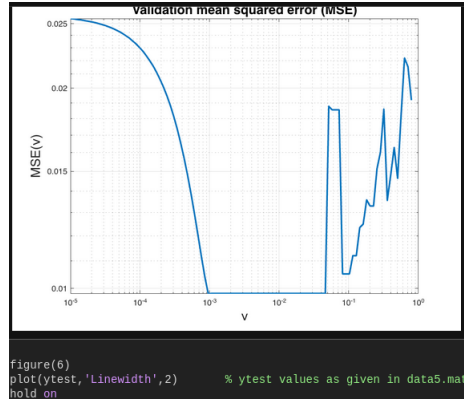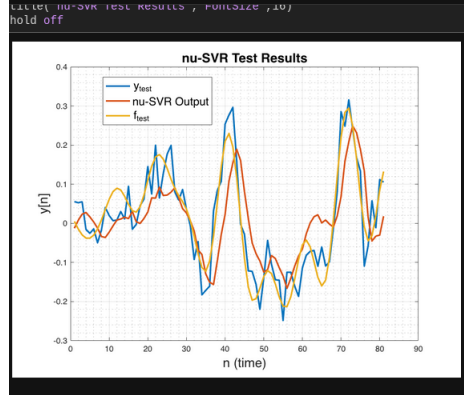
Figure 7: Validation MSE nu-SVR

Figure 8: nu-SVR Test Results

5   Conclusion

As you can see from the results stated above, the results are all very similar. Since we are utiliing the same dataset for all three algorithms, this would lead to similar results.

Attempting to identify the optimal value may be the reason why results are similar as well, choosing the incorrect optimal values would also lead to larger differences in the results of the experiments.

The 20 percent test data was the same for all three algorithms, having provided a different percentage, or perhaps different values within that 20 percent may have also caused discrepencies in the results between the models.