ENPM808A - Introduction to Machine Learning Learning Model Evaluation and Selection

Shon Cortes

Introduction

Models Evaluated:

- Linear Regression
- Decision Regression Tree (3 different max depths)
- Sequential Neural Network

Steps taken:

1. Preprocess data
2. Train models
3. Evaluate models by comparing out of sample errors and examine the learning and validation curves
4. Select best performing model

Data Preprocessing

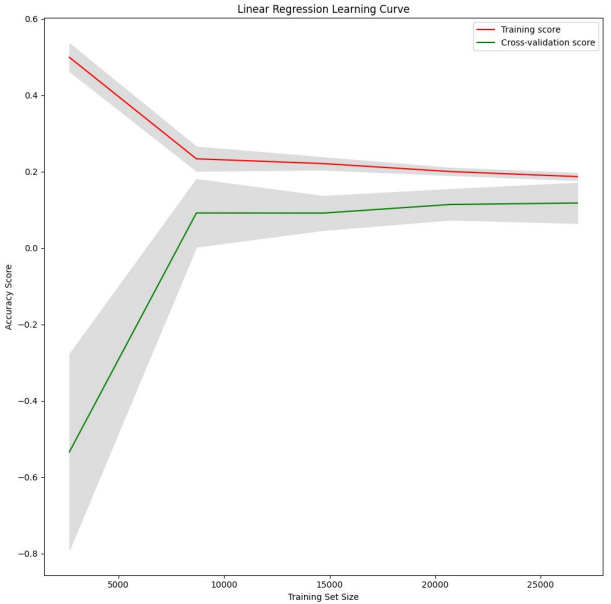
- LIDAR Data simplified into 4 columns that were averages of data within 67.5 degree segments of the original 270 degree FOV data. This reduced 1080 features to 4.
- Local goal and robot current position data was combined by finding the deltas of their respective x, y, qk, and qr values. This reduced 8 features to only 4.
- The feature data was padded with a 1 to give a feature list of 9 elements.
- The data was regularized by dividing each feature set by the maximum in their column. This scales all features to be between 0 and 1.
- The Final goal information was not considered in any training models and was therefore discarded.

Linear Regression

Linear Regression Weights: [[0.00000000e+00 0.00000000e+00]
[-6.41612957e-01 -1.29485114e-01]
[5.79072690e-02 -1.35470964e-01]
[5.31883990e-01 1.33194893e-01]
[-7.12046332e-01 6.69709921e-02]
[1.44051771e-02 1.80931171e-02]
[-2.64140785e-02 1.66382072e-03]
[-1.13852007e-01 1.45684322e-02]
[-6.27623154e-03 -4.89986573e-04]]

Linear Regression Out of Sample Error: 0.12049004638645117
Linear Regression In Sample Error: 0.07156458766910313
Linear Regression Out of Sample Accuracy Score: -0.15649031778890493
Linear Regression In Sample Accuracy Score: 0.1812224925740104

The success rate was 87.9% when using the test data (out of sample error 0.120).



The plot above shows the learning and validation curves converging which indicates that adding more training data may have a marginal effect on the model performance.

Decision Regression Tree Max Depth = 1

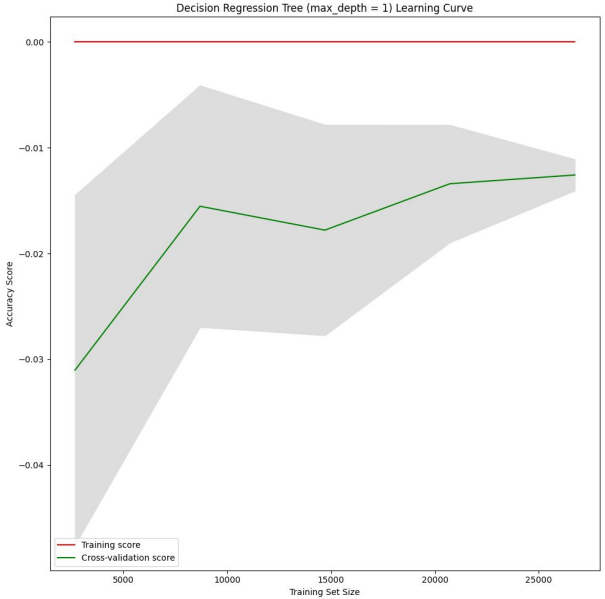
DTR Out of Sample Error (max depth= 1): 0.10358445541261233

DTR In Sample Error (max depth= 1): 0.09541276098000764

DTR Out of Sample Accuracy Score (max depth= 1): -0.029599475999058766

DTR In Sample Accuracy Score (max depth= 1): 0.0

The out of sample error was 89.7% and the plot on the right shows that more training data could improve the model performance.



Decision Regression Tree Max Depth = 10

DTR Out of Sample Error (max depth= 10): 0.16056006687573898

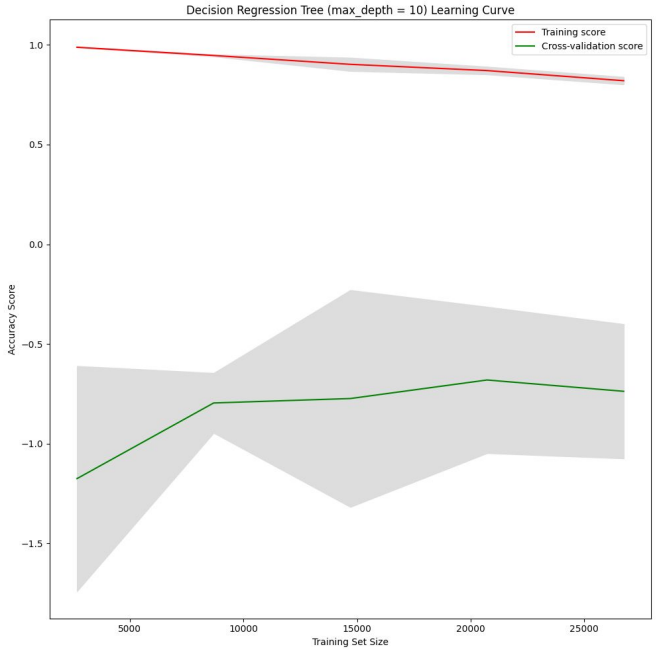
DTR In Sample Error (max depth= 10): 0.012906480492195158

DTR Out of Sample Accuracy Score (max depth= 10):

-0.688004772502052

DTR In Sample Accuracy Score (max depth= 10): 0.8238384103059477

The out of sample error was 83% and the plot on the right shows that more training data could improve the model performance.



Decision Regression Tree Max Depth = Default

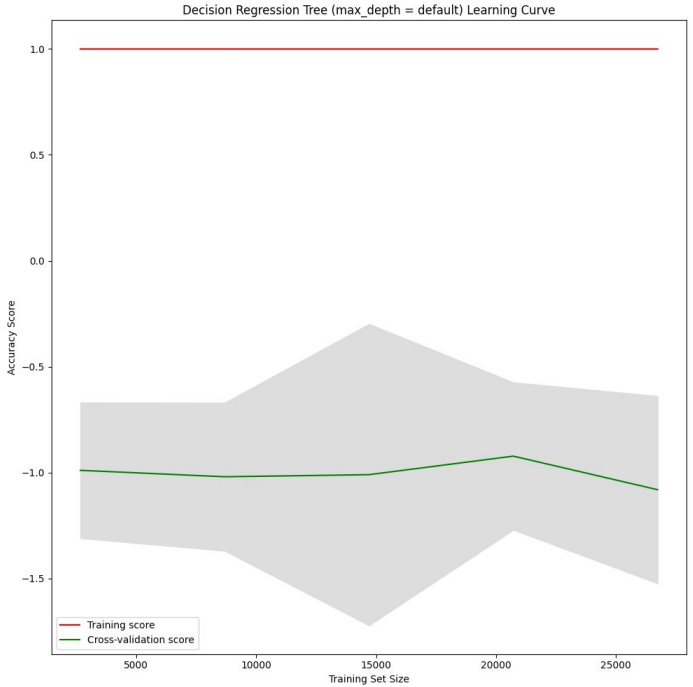
DTR Out of Sample Error (max depth= default): 0.1826566532802652

DTR In Sample Error (max depth= default): 1.5336673387216882e-07

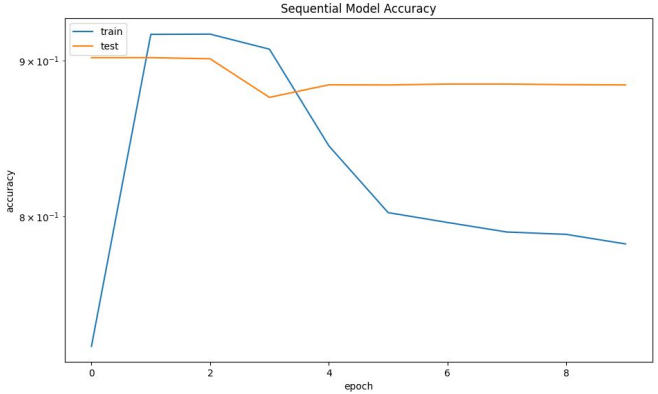
DTR Out of Sample Accuracy Score (max depth= default): -0.9135885418106986

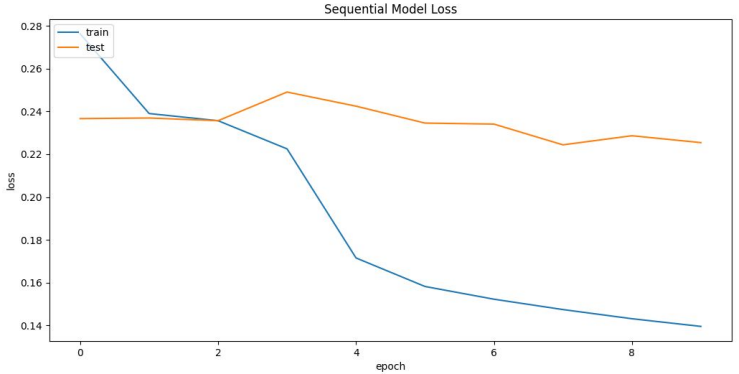
DTR In Sample Accuracy Score (max depth= default): 0.9999969728718665

The out of sample error was 81.4% and the plot on the right shows that more training data could improve the model performance.



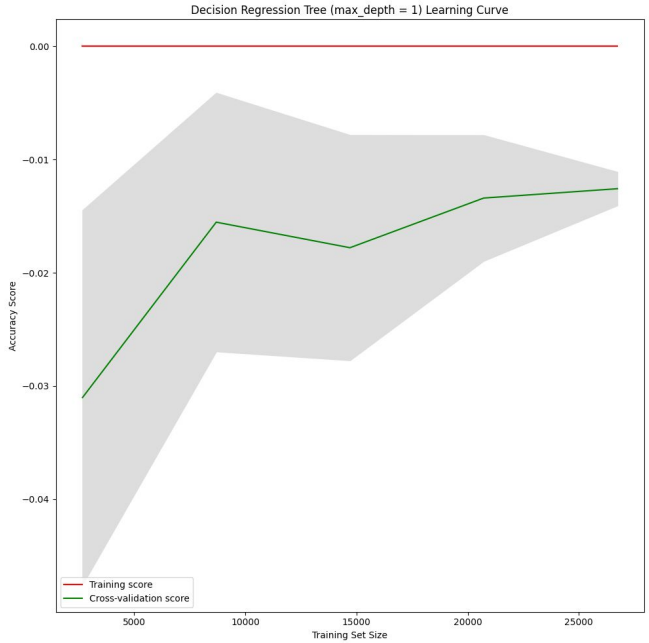
Sequential Recurrent Neural Network





The sequential model had a worse accuracy the more data it was trained on, as seen in in the plots above, however its out of sample performance was not greatly affected by the amount of data it was trained on. The plot on the right shows the model did experience less loss as more data was used for training. The out of sample error was 0.119 meaning the model had an 88.1% accuracy.

Conclusion



We strive for a model to be generalizable in optimal model selection. The model which generalizes best using out of sample data was chosen to be the best performing. Decision regression tree training using a maximum depth of one returned a model that successfully predicted the correct output 89.7% of the time. The learning and validation curves for this model indicate that more training data would improve the model performance. For future improvement, the model would be trained on the full data set provided. With the full training data set being used, all models should be re-evaluated as their performances may change as well.