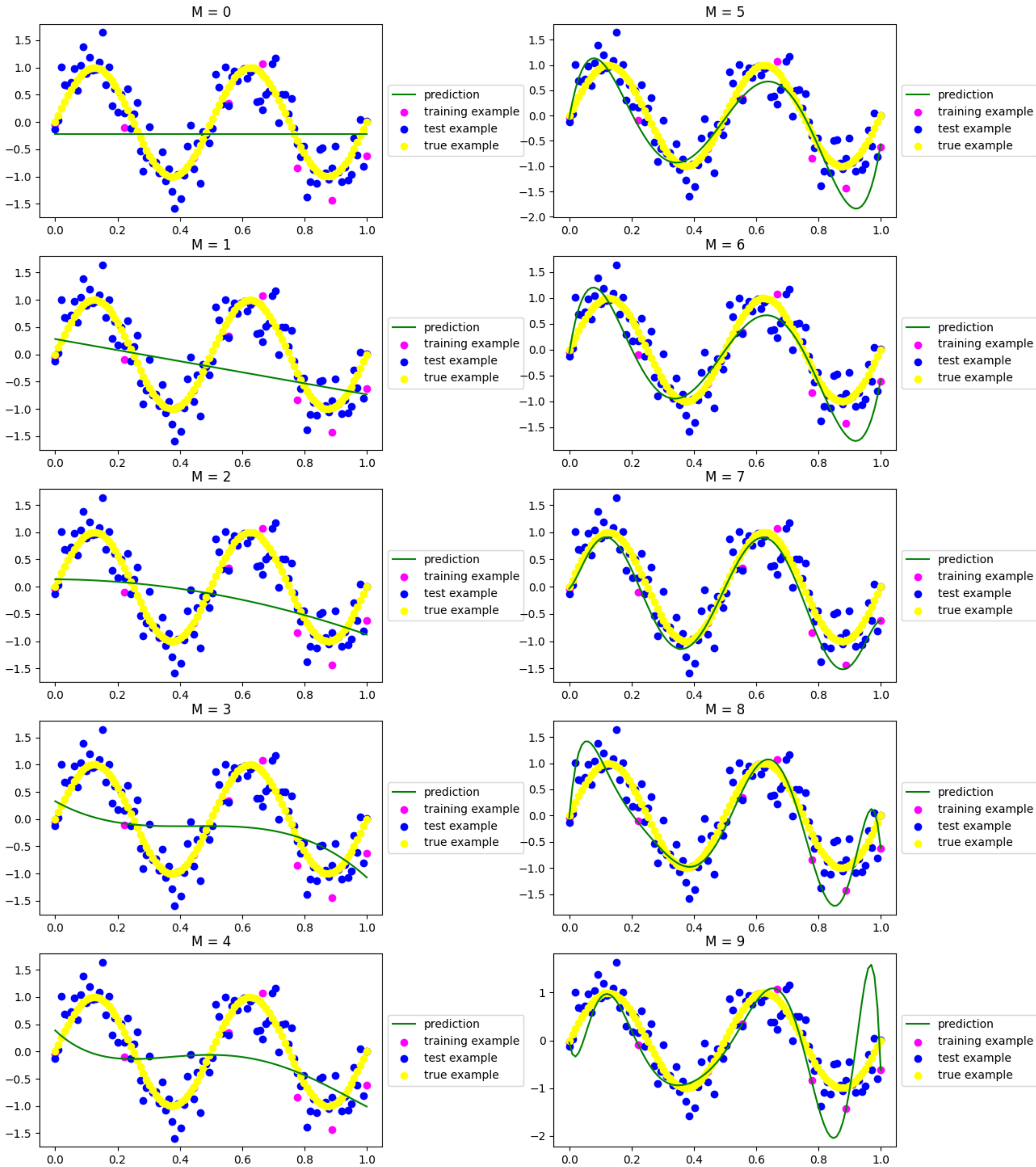
McMaster University

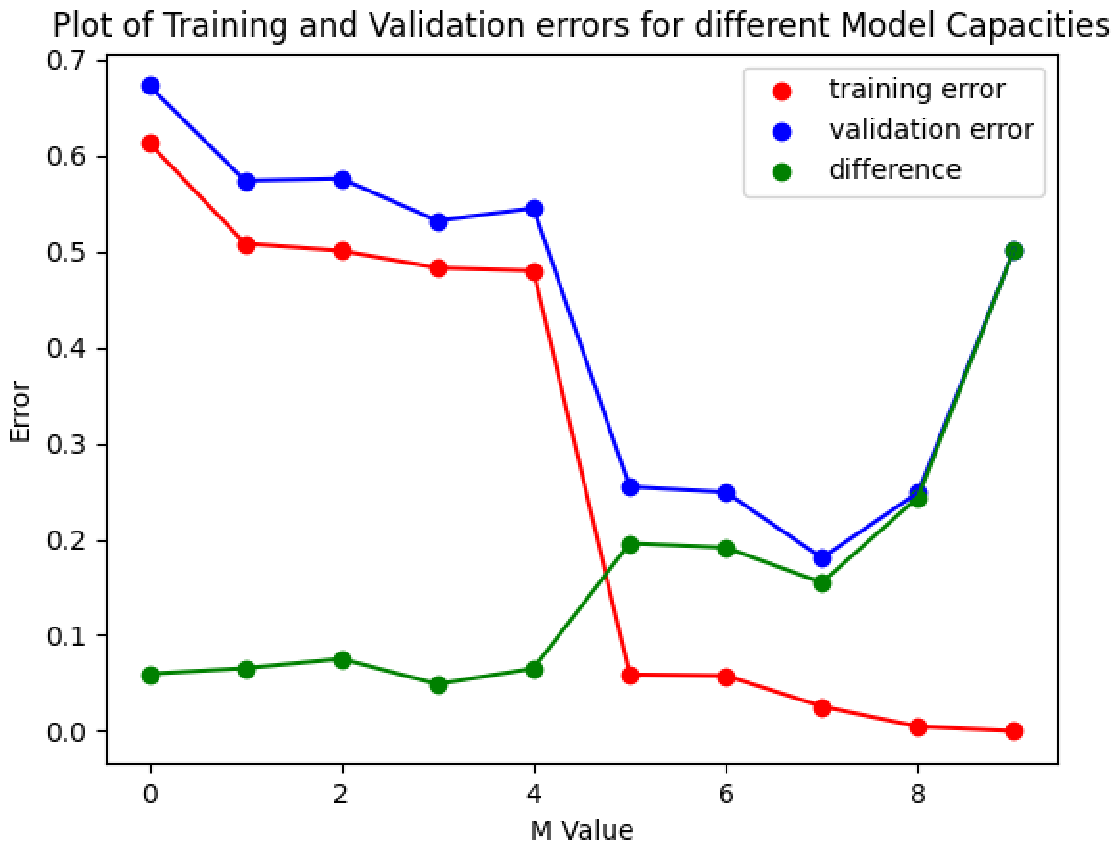
4SL3 Assignment 1

Linear Regression

Uday Sharma - 400139248
10-2-2021

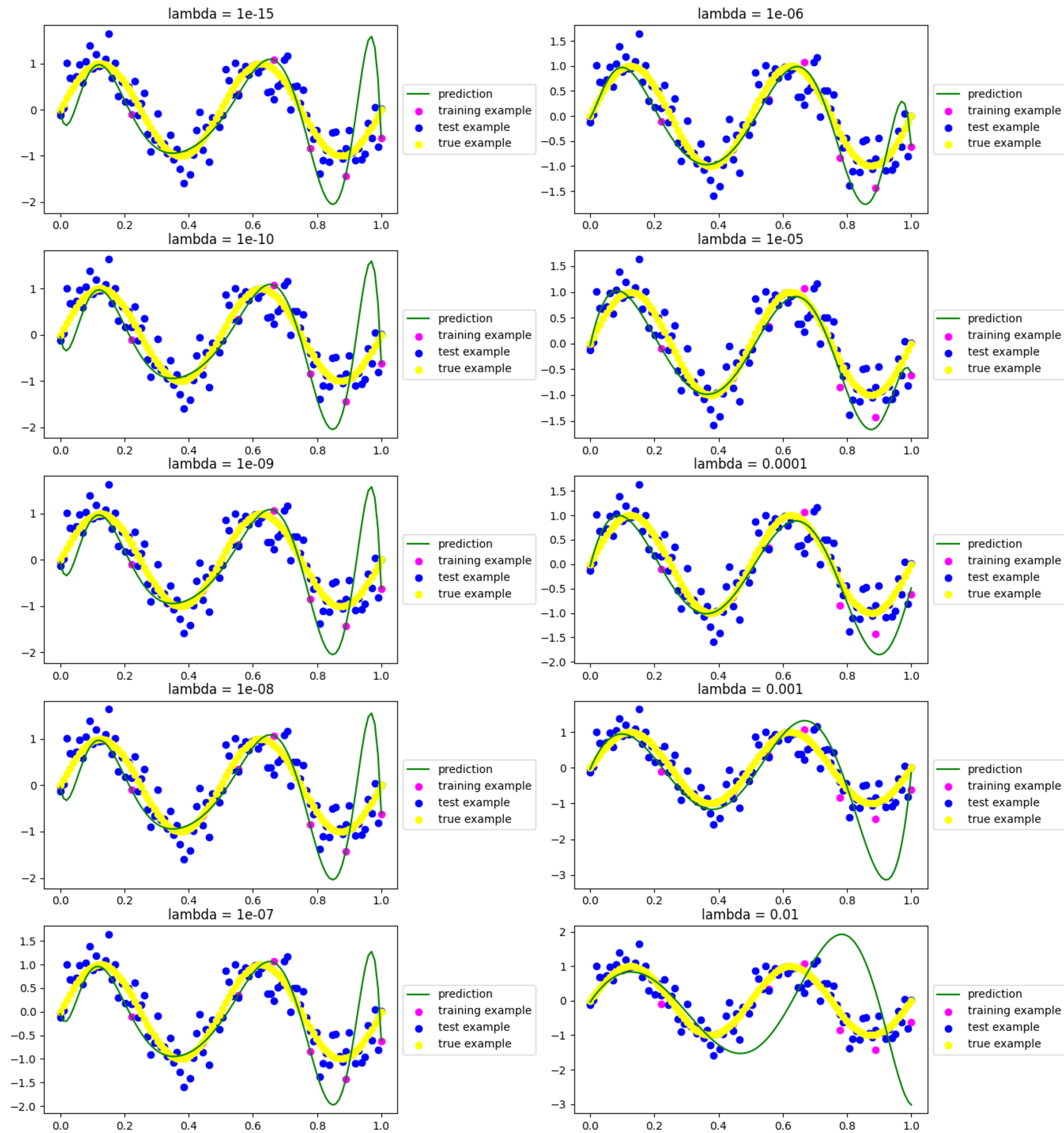
Plots with different model capacities.





With the plots shown above. We can see that as we increase the model capacity, we significantly decrease the training error. As we are training the model against the training set, and our goal is to minimize the error between our true targets and predicted targets by increasing the model capacity. However, we also notice that as the model capacity is increased, we become prone to overfitting, this is evident when the model capacity is 9 (M=9). We are severely overfitting the curve while retaining minimum training error. For the M=9 case, our predicted curve passes directly through all the points in our training set.

Controlling Overfitting with Regularization:



For the model capacity M=9, we notice a large case of overfitting. This can be controlled using L^2 Regularization.

$$\bar{\mathcal{C}}(\mathbf{w}) = \frac{1}{N}\sum_{i=1}^{N}\left[\sum_{k=0}^{D} w_k x_k^{(i)} - t^{(i)}\right]^2 + \lambda \sum_{k=1}^{D} w_k^2.$$

Where we apply penalties to larger values in the predictor. Reference the plots above for different values of lambda. We can see that for λ = 1e-4 we eliminate any overfitting and for λ = 1e-3 we begin to see some underfitting.