ISE-537
Homework 4
Lucas (Deuce) Palmer

Problem #1 *(all calculations performed in the first half of HW-4.ipynb)*

(a)

$$\hat{\sigma}^2 = \frac{1}{M}\sum_{k=1}^{M}(S_k - S_{k-1})^2$$

where M = entire dataset (231 entries)

$$\hat{\sigma}^2 = \frac{1}{231}34.3342$$

$$\hat{\sigma}^2 = 0.1486$$

(b)

ε = Spread / 2 ε = 0.2 / 2 **ε = 0.1**	η = Spread / (0.0001 * average daily volume) η = 0.2 / (0.0001 * 157,120,451.95) **η =1.27 x 10^{-5}**	γ = Spread / (0.001 * average daily volume) γ = 0.2 / (0.001 * 157,120,451.95) **γ =1.27 x 10^{-6}**

(c)

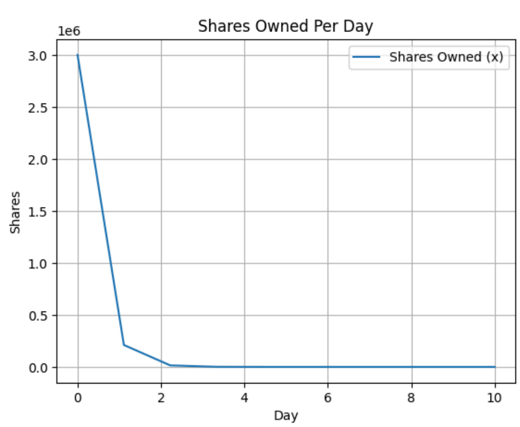
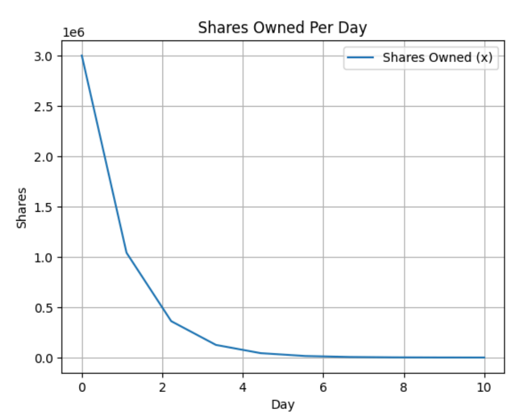
- $\tau = T/N = 1$
- $\tilde{n} = \eta - \frac{1}{2}\gamma\tau = 1.21$
- $\kappa\_tilde^2 = \frac{\lambda\hat{\sigma}^2}{\tilde{n}} = \frac{\lambda(0.1486)}{1.21}$ where λ is 0.001 or 0.0001
- $\frac{2}{\tau^2}(\cosh(\kappa\tau) - 1) = \kappa\_tilde^2$

$$\cosh(\kappa\tau) = \frac{\kappa\_tilde^2\tau^2}{2} + 1$$

$$\kappa\tau = \cosh^{-1}\left(\frac{\kappa\_tilde^2\tau^2}{2} + 1\right)$$

$$\kappa = \frac{\cosh^{-1}\left(\frac{\kappa\_tilde^2\tau^2}{2}+1\right)}{\tau}$$

Use for loop to calculate x[j] & n[j] for j = 1, 2, …, N

when λ = 0.001	when λ = 0.0001
	

(d)

When taking the first-order derivative of the objective function $U_\lambda(x)$ with respect to x, εX becomes zero and therefore epsilon does not appear in the solution of the optimal trading list.

Problem #2 *(all calculations performed in the second half of HW-4.ipynb)*

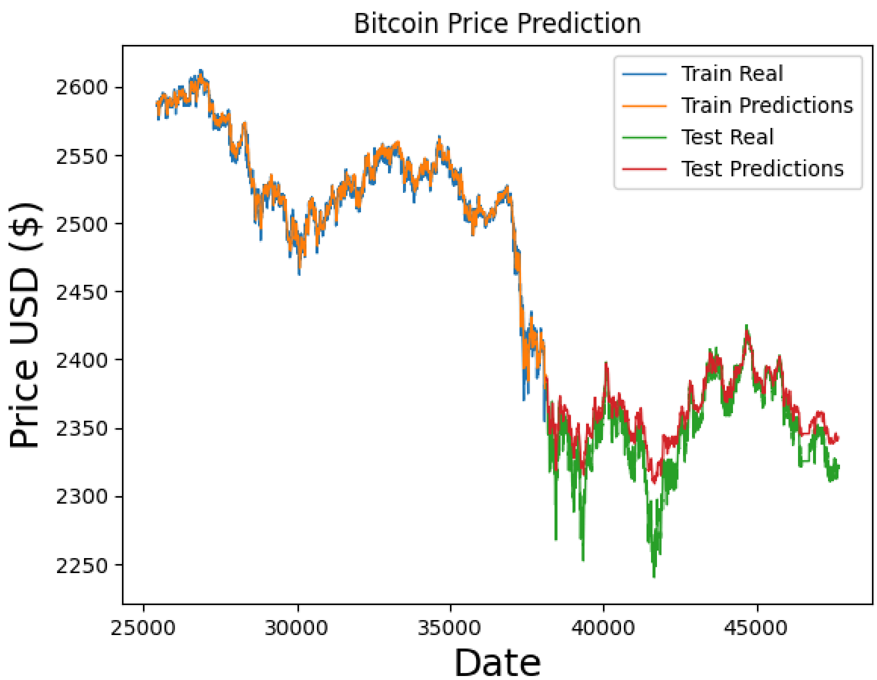
(a)

RMSE on training set = 5.13

RMSE on test set = 18.24

As the error on the training set is substantially lower than that of the unseen data in the test set, there is evidence that the model is over-fitting. The performance of the results, while subject, seems to be quite poor as the RSME is relatively high. It would be very difficult to achieve success when trading Bitcoin based on a model that has an RSME of 18.24 on unseen data, therefore I would conclude that the performance is not satisfactory.

(b)



The graph acts as a visual representation of what was shown by the root-mean-squared error calculations. The performance is much better for the training set, as expected, and error substantially increases during the test set. In particular, the model fails to detect three large price drops early on in the test data and has a higher prediction when compared to the true value throughout most of the test set.