Berke Ataseven

0054326

**DASC521 HW#4 Report**

In this homework, we were asked to implement 3 different non-parametric regression algorithms. These three algorithms are: regressogram, running mean smoother and kernel smoother.

Data points required to draw the relevant curves are generated by the functions “*calculate\_regressogram ()*”, “*calculate\_mean\_smoother ()*”, “*calculate\_mean\_smoother ()*”. In these functions, I have discretized the x-axis, then calculated the   values for every discretized x value. Following equations are used in the implementation of the functions:

|  |  |  |
| --- | --- | --- |
|  |  |  |
| where, | where, | where, |
|  |  |  |
| **REGRESSOGRAM** | **RUNNING MEAN SMOOTHER** | **KERNEL SMOOTHER** |

Notice that in the equation for kernel smoother, function is the formula for the normal distribution whose mean () is 0 and covariance parameter () is 1. After the curve functions are generated, I have calculated the root mean squared error (RMSE) for the test data. Calculated RMSE values turned out to be slightly different from what is given in the homework description. I reckon the reason for this is the size of the increment in the discretization of the x-axis. Resulting RMSE values are given below:

|  |
| --- |
|  |

I have commented my code for the sake of intelligibility.