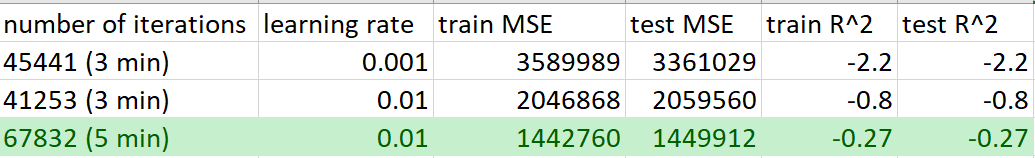
**Part 1**

The optimization algorithm I used is the Adam optimizer:

Adam stands for Adaptive Moment Estimation. Its corrected moments allows the gradient to move towards the general trend without oscillating a lot or overshooting.

**Python code**

Best set of parameters:  
Test dataset error values:   


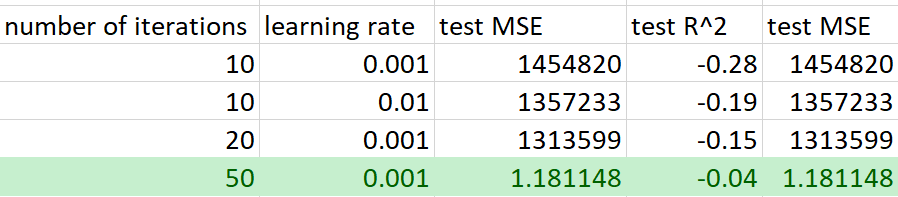
**Am I satisfied that I have found the best solution?**

I am not satisfied, as my model seemed to take forever to converge, and it never did. I was able to cut down a lot of time by doing feature selection, but the performance still paled in comparison to sklearn’s adam optimizer.

**Part 2**

**Python code**

Best set of parameters:  
Test dataset error values:

****

**Am I satisfied that the package has found the best solution?**

I am satisfied, even though I did not do several iterations. I can see that increasing the number of iterations would give me good results, as it gave lower R^2 and MSE values.