**CSCE 623: Machine Learning**

**Spring 2019**

**HW 5 GRADING WORKSHEET**

Due Tuesday, 21 May at 2359

Submit via Canvas

**(**This Homework is worth 5 points toward your final grade**)**

|  |  |
| --- | --- |
| Course points earned | 5.0 |

**Student Name: NEWLIN**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Step | Step Evaluation Criteria | Avail points | Student performance | Student Score |
| 1 | log transform data | 3 | Achieved | 3 |
| 3 | Explore data (plot & discuss) relationships between features and Salary | 6 | Achieved | 6 |
| 4 | Make a prediction on RndF with respect to #maxDepth, #maxFeatures | 3 | Achieved | 3 |
| 5 | Set up k-fold on non-test, decide n-splits, maxD (1,20), maxF (1,p); collect MSE for each setting | 8 | Achieved | 8 |
| 6 | Convincing visual evidence on MSE vs maxF and maxD | 5 | Achieved | 5 |
| 7 | Report params with lowest MSE. Discuss sqrt(p) vs p/3 for maxF | 5 | Achieved | 5 |
| 8 | Fit new model with best parameters on all non-test data | 3 | Achieved | 3 |
| 9 | Determine/Report performance model for salary prediction in real dollars | 5 | Achieved | 5 |
| 10 | Scatterplot residuals & discuss patterns where prediction would be poor | 7 | OK… but according to your residuals code, positive values mean you underpredicted. | 7 |
| 11 | Variable importance (feature importance) w numerical and visual evidence | 5 | Achieved | 5 |
| TOTAL |  | 50 |  | 50 |