|  |  |
| --- | --- |
| A description... | **Project Grading Rubric**  **BUDT 758T** |

***Note: one point will be deducted from the scores reported below for every evaluation above and beyond the official ones.***

There are three components of the project grade:

1. Main project deliverables (85 points)
2. Contest performance (5 points)
3. Effort and participation (10 points)

**Main Project Deliverables (up to 85 out of 100 points)**

*If you are unable to complete the tasks required to receive 60 points, I will score your project on a sliding scale.*

|  |  |  |
| --- | --- | --- |
| **WORTH 60 POINTS** | | |
|  | **Task** | **Achieved?** |
| 1 | Demonstrate an insight about at least one feature in the dataset with an appropriate and clear chart or table. | N |
| 2 | Correctly implement a train/validation split for model evaluation/selection. | Y |
| 3 | Train a model that predicts your chosen target variable. | Y |
| 4 | Generate and submit predictions that are better than the baseline for your chosen target variable. | Y |
| 5 | Turn in well-documented, concise R code that accomplishes your data understanding/preparation, modeling, prediction generation, and evaluation. | Y |
| 6 | Turn in a clear, professionally written report describing the necessary information about all of the other tasks (see report template for required sections and other guidance). | N |
| 7 | Provide peer evaluations for your group members. | N |

|  |  |  |
| --- | --- | --- |
| **WORTH 75 POINTS** | | |
|  | **Task** | **Achieved?** |
| 1 | Everything required to get 60 points. |  |
| 2 | Clean and/or create at least 30 feature variables for use in your models (not including dummy variables). | Y |
| 3 | Demonstrate insights about at least three features in the dataset with appropriate and clear charts or tables. |  |
| 4 | Compare at least three different types of models and describe them as specified in the report template. |  |
| 5 | Generate and submit predictions that are better than the baseline for your chosen target variable, using a model that includes at least 10 feature variables (not including dummies). |  |
| 6 | Receive an average peer evaluation score of at least 6/10 from your teammates. |  |

|  |  |  |
| --- | --- | --- |
| **WORTH 80 POINTS** | | |
|  | **Task** | **Achieved?** |
| 1 | Everything required to get 75 points. |  |
| 2 | Clean and/or create at least 50 feature variables for use in your models (not including dummy variables). |  |
| 3 | Use at least one unstructured text field to create new features. Provide evidence in your report that the features derived from this field result in a performance improvement. |  |
| 4 | Demonstrate insights about at least five features in the dataset with appropriate and clear charts or tables. |  |
| 5 | Correctly implement cross-validation for model evaluation/selection. |  |
| 6 | Compare at least four different types of models and describe them as specified in the report template. |  |
| 7 | Tune complexity parameters as appropriate for all models tried and provide fitting curves in your report. |  |
| 8 | Generate and submit predictions that score within 2 standard deviations of the winning group’s performance on your chosen target variable. The model that produces these predictions must include at least 25 feature variables (not including dummies). |  |

|  |  |  |
| --- | --- | --- |
| **WORTH 85 POINTS** | | |
|  | **Task** | **Achieved?** |
| 1 | Everything required to get 80 points. |  |
| 2 | Demonstrate insights about at least ten features in the dataset with appropriate and clear charts or tables. |  |
| 3 | Incorporate an external data source (beyond the Airbnb data that I have provided) to create at least one additional feature variable. Provide evidence in your report that the new feature variable(s) result in a performance improvement. |  |
| 4 | Compare at least six different types of models and describe them as specified in the report template. |  |
| 5 | Include at least one learning curve and describe an insight or decision you’re your team made based on this graph. |  |
| 6 | Generate and submit predictions that score within 1 standard deviation of the winning group’s performance on your chosen target variable. The model that produces these predictions must include at least 40 feature variables (not including dummies). |  |
| 7 | Receive an average peer evaluation score of at least 8/10 from your teammates. |  |

**Contest Points (up to 5 out of 100 points)**

* First place for each target variable gets 5 points
* Second place gets 3 points
* Third place gets 1 point

**Excellence Points (up to 10 out of 100 points)**

The remaining 10 points will be assigned based on above-and-beyond effort and performance on any aspect of the project. You must truly impress me to earn these points. Examples of how to earn some excellence points include, but are not limited to:

* Submitting predictions to all evaluation opportunities.
* Attending and actively participating in project meetings.
* Trying new methods (modeling or feature engineering) beyond what was covered in class.
* Demonstrating an unusual amount of care in modeling, hyperparameter tuning, or feature engineering.
* Writing an extremely professional report.
* Writing highly organized, efficient code.
* Using a rigorous nested holdout design for evaluation.
* Communicating innovative insights regarding the data and/or predictive problem.