|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Instructions:**  Evaluate the homework against the outlined criteria in the below rubric, assigning a rating to each criterion. Add points earned across all criteria and convert the total points to a letter grade, assigning a “+” or “-” letter grade designation at your discretion. | |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | A (+/-) | 90+ | C (+/-) | 40-64 | F (+/-) | <15 | | B (+/-) | 65-89 | D (+/-) | 15-39 |  |  | |
| **Notes:**  The deployed assignment utilizes the **sklearn** library to train models on a set of data and used to make predictions. The source code should also be deployed to **Github** or **Gitlab**. There are more models that could have been used for this HW, the 3 given solutions are only a select few. Therefore, if a student uses a different model that we did not provide as a solution, they will not be docked any points as long as they still meet the requirements specified in the rubric. |  |

**Rubric for Exoplanet Exploration:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Mastery**  **20 points** | **Approaching Mastery**  **15 points** | **Progressing**  **10 points** | **Emerging**  **5-0 points** | **Incomplete** |
| **Data Preprocessing** | The submission does all of the following:  ✓ Unnecessary Columns are removed.  ✓ All rows containing NaN are removed.  ✓ Data is correctly split into a training and test set.  ✓ Numerical data is scaled accordingly. | The submission does 3 of the following:  ✓ Unnecessary Columns are removed.  ✓ All rows containing NaN are removed.  ✓ Data is correctly split into a training and test set.  ✓ Numerical data is scaled accordingly. | The submission does 2 of the following:  ✓ Unnecessary Columns are removed.  ✓ All rows containing NaN are removed.  ✓ Data is correctly split into a training and test set.  ✓ Numerical data is scaled accordingly. | The submission does 0-1 of the following:  ✓ Unnecessary Columns are removed.  ✓ All rows containing NaN are removed.  ✓ Data is correctly split into a training and test set.  ✓ Numerical data is scaled accordingly.  -OR-  ✓ No preprocessing done. | No submission was received  -OR-  Submission was empty or blank  -OR-  Submission contains evidence of academic dishonesty |
| **Model Creation & Feature Selection** | The submission does all of the following:  **Model Creation:**  ✓ Creates, trains, and tests at least 2 different classification models  ✓ Correctly sets X and y (koi\_disposition) variables  **Feature Selection:**  ✓ Uses some form of feature selection method to identify insignificant variables (feature\_importance, RFE, backwards elimination, etc.)  ✓ Remove insignificant variables and retrain models with the significant features | The submission does 3 of the following:  **Model Creation:**  ✓ Creates, trains, and tests at least 2 different classification models  ✓ Correctly sets X and y (koi\_disposition) variables  **Feature Selection:**  ✓ Uses some form of feature selection method to identify insignificant variables (feature\_importance, RFE, backwards elimination, etc.)  ✓ Remove insignificant variables and retrain models with the significant features | The submission does 2 of the following:  **Model Creation:**  ✓ Creates, trains, and tests at least 2 different classification models  ✓ Correctly sets X and y (koi\_disposition) variables  **Feature Selection:**  ✓ Uses some form of feature selection method to identify insignificant variables (feature\_importance, RFE, backwards elimination, etc.)  ✓ Remove insignificant variables and retrain models with the significant features | The submission does 0-1 of the following:  **Model Creation:**  ✓ Creates, trains, and tests at least 2 different classification models  ✓ Correctly sets X and y (koi\_disposition) variables  **Feature Selection:**  ✓ Uses some form of feature selection method to identify insignificant variables (feature\_importance, RFE, backwards elimination, etc.)  ✓ Remove insignificant variables and retrain models with the significant features  -OR-  ✓ Only uses non-classification models. |
| **Model Tuning** | The submission does all of the following:  **Model Tuning:**  ✓ Uses GridSearch or some hyperparameter tuning to find the best parameters for the model  ✓ The tuned model is used to make the final exoplanet prediction | The submission does 2 of the following:  **Model Tuning:**  ✓ Uses GridSearch or some hyperparameter tuning to find the best parameters for the model  ✓ The tuned model is used to make the final exoplanet prediction | The submission does 1 of the following:  **Model Tuning:**  ✓ Uses GridSearch or some hyperparameter tuning to find the best parameters for the model  ✓ The tuned model is used to make the final exoplanet prediction | The submission does 0 of the following:  **Model Tuning:**  ✓ Uses GridSearch or some hyperparameter tuning to find the best parameters for the model  ✓ The tuned model is used to make the final exoplanet prediction |
| **Model Accuracy** | ✓ Model scores greater than 85% accuracy on test data. | ✓ Model scores between 85% and 75% accuracy on test data. | ✓ Model scores between 75% and 50% accuracy on test data. | ✓ Model scores less than 50% accuracy on test data. |
| **Reporting** | The submission does all of the following:  **Reporting**  ✓ README compares each of the models’ performances and predictions  ✓ README summarizes the findings and makes assumptions based on the data and their models.  ✓ README discusses the predictions of the possible exoplanets with their models. | The submission does 2 of the following:  **Reporting**  ✓ README compares each of the models’ performances and predictions  ✓ README summarizes the findings and makes assumptions based on the data and their models.  ✓ README discusses the predictions of the possible exoplanets with their models. | The submission does 1 of the following:  **Reporting**  ✓ README compares each of the models’ performances and predictions  ✓ README summarizes the findings and makes assumptions based on the data and their models.  ✓ README discusses the predictions of the possible exoplanets with their models. | The submission does 0 of the following:  **Reporting**  ✓ README compares each of the models’ performances and predictions  ✓ README summarizes the findings and makes assumptions based on the data and their models.  ✓ README discusses the predictions of the possible exoplanets with their models.  -OR-  ✓ Does not submit a README |