**Machine Learning Homework Analysis**

**Goal:** Create a README that reports a comparison of each model's performance as well as a summary about your findings and any assumptions you can make based on your model (is your model good enough to predict new exoplanets? Why or why not? What would make your model better at predicting new exoplanets?).

**Analysis:**

*Model* 1: Using logistic regression, this model output produced a best score of 0.86 (see Notebook, (model\_1.ipynb). This result is considered adequate. All the false positives in the data gives us the ability to know that the data points indicate a potential exoplanet. However, using gridsearchcv, the result didn’t change a significant amount from the initial data scores (0.85), so a second model was created to hopefully produce a better predictive level for models (0.90 or better).

*Model 2*: Using a logistic regression CV model, it was hoped changing only a few parameters would result in a better model (model\_2.ipynb) versus Model 1. Changing from MinMaxScaler to a standard scaler netted only slightly better results than the provided scaler. Then the only other metric that was changed and saved was the logistic regression model cv, even this small change produced a better result than before at 0.88. However, our max cv was 5 compared to Model 1’s 10. This model still did not net the result hoped for a 0.90 or better % score but was close enough to this benchmark to be considered slightly more successful than Model 1.

I then tried to “classify” my processed and tested model by using the *sklearn.preprocessing import LabelEncoder* feature. I tried to create a new data frame and sort by features (KOI “duration, “impact” etc.), I had selected for “importance”. However, I could not get this code to run (or understand the implications of the limited code that I did successfully run). Thus, this code was not included in my Model 2 ipynp file submission.

The problem I believe was due to my unfamiliarity working with Sklearn Packages *(LogisticRegression*, *KNeighborsClassifier*, and model selection functions) and Sklearn Evaluation Metrics. With input from my group, I am hoping to apply what I have learned to the final group (machine learning) project.

**Conclusion**: While additional fine-tuning may create a model with higher predictive power than these (R2 of 90% or better), these models appear to have good predictive value in predicting whether there are exoplanets out there for humans to live on.