Sea surface temperature (SST) is a fundamental quantity for understanding weather and climate dynamics. Although sensors aboard satellites provide global and repeated SST coverage, a characterization of SST retrieval error precision and bias is necessary for determining the suitability of SST retrievals in various applications. Guidance on how to derive meaningful error estimates is still being developed. Previous methods estimated retrieval uncertainty based on geophysical factors, eg. season or “wet”, “dry” atmospheres, but the discrete nature of these bins led to spatial discontinuities in SST maps. Recently, a new approach clustered retrievals based on the terms (excluding offset) in the statistical algorithm used to estimate SST, resulting in over 600 clusters - too many to understand the geophysical conditions that influence retrieval error. Using MODIS and buoy SST matchups (2002 - 2016), we use machine learning algorithms (recursive and conditional trees, random forests) to gain insight into geophysical conditions leading to the different signs and magnitudes of MODIS SST residuals (satellite SST estimates minus buoy SSTs). MODIS retrievals were first split into three categories: < -0.4 C, -0.4 C <= residual <= 0.4 C, and > 0.4 C. These categories are heavily unbalanced, with residuals > -0.4 C being much less frequent. Because performance of classification algorithms is affected by imbalance, we therefore test rebalancing algorithms (oversampling, undersampling, combinations of the two). We consider multiple features for the decision tree algorithms: regressors from the MODIS SST algorithm, proxies for temperature deficit, and measures of spatial homogeneity in brightness temperatures (BTs) such as the range of values in the 11 um channel over a 25 km^2 area centered on the buoy/retrieval area. These features and a rebalancing of classes led to a classifier that accurately identified the SST retrievals in the < -0.4 C and -0.4 C <= residual <= 0.4 C categories with 83.6% accuracy. Spatial homogeneity in BTs consistently appears among the most important variables for classification of SST residuals, suggesting that unidentified cloud contamination still is one of the causes leading to larger SST residuals. Precision and accuracy of error estimates from our decision tree classifier are enhanced using this knowledge.