Predictive analysis of the influence of the chemical composition and pre-processing regimen on structural properties of steel alloys using machine learning techniques. NARAYANAN KRISHNAMURTHY, SIDDHARTH MADDALI, VYACHESLAV ROMANOV, National Energy Technology Laboratory (Pittsburgh), JEFFREY HAWK, National Energy Technology Laboratory (Albany) — We present some structural properties of multi-component steel alloys as predicted by a random forest machine-learning model. These non-parametric models are trained on high-dimensional data sets defined by features such as chemical composition, pre-processing temperatures and environmental influences, the latter of which are based upon standardized testing procedures for tensile, creep and rupture properties as defined by the American Society of Testing and Materials (ASTM). We quantify the goodness of fit of these models as well as the inferred relative importance of each of these features, all with a conveniently defined metric and scale. The models are tested with synthetic data points, generated subject to the appropriate mathematical constraints for the various features. By this we highlight possible trends in the increase or degradation of the structural properties with perturbations in the features of importance. This work is presented as part of the Data Science Initiative at the National Energy Technology Laboratory, directed specifically towards the computational design of steel alloys.