The following thesis explores the impact of the dataset distributional properties on classification performance. We use Gaussian copulas to generate 1000 artificial dataset and train classifiers on them. We train Generalized linear models, Distributed Random forest, Extremely randomized trees and Gradient boosting machines via H2O.ai machine learning platform accessed by R. Classification performance on these datasets is evaluated and empirical observations on influence are presented. Secondly, we use real Australian credit dataset and predict which classifier is possibly going to work best. The predicted performance for any individual method is based on penalizing the differences between the Australian dataset and artificial datasets where the method performed comparatively better, but it failed to predict correctly.