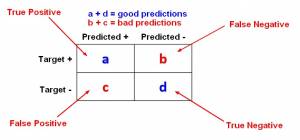
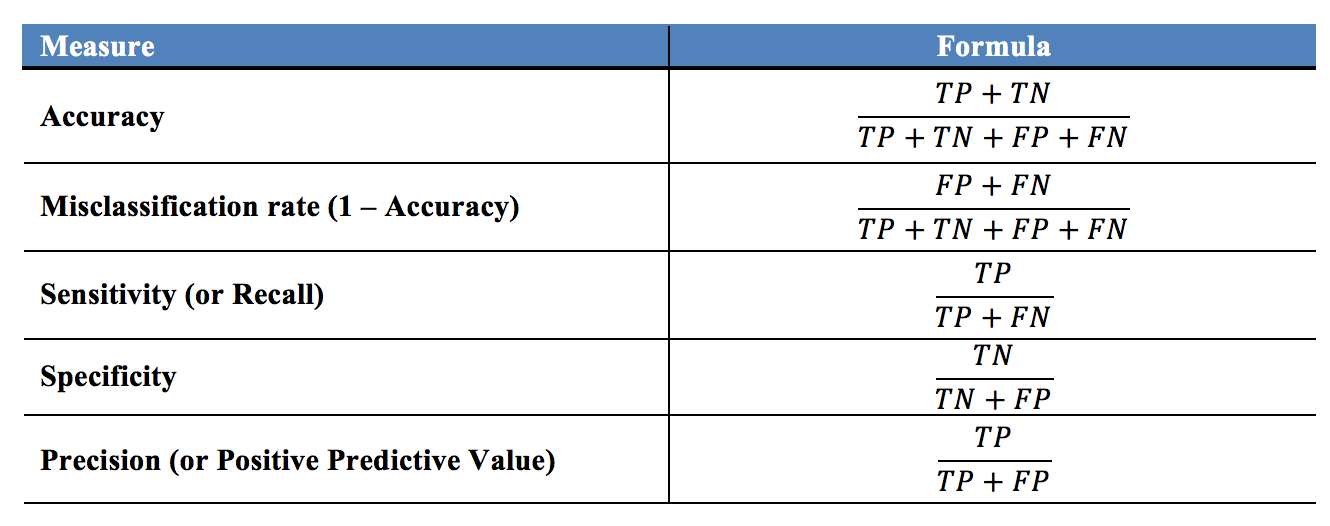
***Model Evaluation***

We used System R for data preparation and analysis. The two classiﬁcation methods (Decision Tree, Random Forest) were applied on data before and after applying the SMOTE algorithm to balance the data. For each model, we used 75%ofthedatainstancesfortraining, 15%forvalidation, and 15% for testing. For each run, we calculated some of the standard performance measures (statistics) to evaluate the performance of the algorithms, including error rate and speciﬁcity. Classiﬁcation results have four possible outcomes: true positive (TP), true negative (TN), false negative (FN), and false positive (FP), as given in Table.





*Decision Tree –*

* Accuracy = (868 + 1651 + 310) / total = **0.58**
* Sensitivity = **0.67**
* Error rate = **0.42**

*Random Forest –*

* Accuracy = (1230 + 1739 + 704) / total = **0.74**
* Sensitivity = **0.79**
* Error rate = **0.26**