

Competition result

	Train error		Test error		Brier difference (test-train error) (higher values indicate more overfitting)
	Brier	MCR	Brier	MCR	
GAM (two variables less)	0,1669	0,2440	0,1691	0,2444	0,0022
SVM linear kernel	-	0,2416	-	0,2486	
SVM radial basis kernel	-	0,2103	-	0,2486	
tree	0,1706	0,2457	0,1739	0,2389	0,0033
random forest	0,0243	0,0000	0,1723	0,2389	0,1480
bagging	0,0247	0,0000	0,1765	0,2389	0,1518
gradient boosting	0,1502	0,2211	0,1676	0,2375	0,0174
gradient boosting tuned	0,1759	0,2503	0,1352	0,1875	-0,0407

Boldfaced values indicate the top three methods. Boosting outperforms all methods.

Random forests and bagging might profit from tuning their parameters (probably, tree depth and mtry). Currently, they seem to overfit quite dramatically.

The GAM does pretty well out of the box, and used less variables than the other methods so perhaps the comparison is unfair.

The tree, given its simple structure, does quite well.