Updates (train\_MA\_Apr19)

* Cell 1: imports (#MA)
* Correlation heatmap (after preprocess)
* Preprocess
* Add to features
* Bagging: BaggingClassifier(KNeighborsClassifier()
* Very end: R code (markdown?)

Ada Boost

precision recall f1-score support

False 0.97 0.78 0.86 752

True 0.18 0.64 0.28 58

avg / total 0.91 0.77 0.82 810

precision recall f1-score support

False 0.96 0.78 0.86 743

True 0.20 0.61 0.30 67

avg / total 0.89 0.77 0.82 810

Gradient Boost

precision recall f1-score support

False 0.98 0.76 0.86 781

True 0.09 0.62 0.16 29

avg / total 0.95 0.76 0.83 810

precision recall f1-score support

False 0.98 0.76 0.86 783

True 0.07 0.56 0.13 27

avg / total 0.95 0.75 0.83 810

Logistic Regression

precision recall f1-score support

False 0.98 0.77 0.86 777

True 0.10 0.64 0.18 33

avg / total 0.94 0.76 0.83 810

precision recall f1-score support

False 0.98 0.77 0.86 772

True 0.12 0.63 0.20 38

avg / total 0.94 0.76 0.83 810

Random Forest

precision recall f1-score support

False 0.95 0.77 0.85 744

True 0.16 0.50 0.25 66

avg / total 0.88 0.75 0.80 810

precision recall f1-score support

False 0.95 0.77 0.85 753

True 0.14 0.51 0.22 57

avg / total 0.90 0.75 0.81 810

Bagging (Decision tree)

precision recall f1-score support

False 0.94 0.76 0.84 752

True 0.12 0.41 0.18 58

avg / total 0.88 0.74 0.80 810

precision recall f1-score support

False 0.94 0.76 0.84 748

True 0.12 0.39 0.18 62

avg / total 0.87 0.73 0.79 810