New Case Study Eight

You have been given the following banking data. There is no missing data:

<https://smu.box.com/s/bizpcq4a3dtid04tke1x7nz7359r79y5>

There is no metadata or data dictionary. Your class target is labeled “target”

1. Build 3 tuned models: An XGBoost, a Random Forest, and an SVM.
2. Show the log loss and accuracy for XGBoost, and Random Forest models on out of fold predictions. Show the Accuracy of the SVM on a validation set
3. Time how long it takes to do a sample of 1000, 2000, 5000, and 10,000 rows in the SVM. What is the rough scaling of SVM with sample size?