

Overall Survival of Clinical Trials

Predicting early trial termination

Beth Baumann



Why do trials end early?



Low patient recruitment



Safety concerns



Lack of efficacy



Loss of funding



Project Process



Phase I

Interventional
Cancer Trials
from
clinicaltrials.gov API

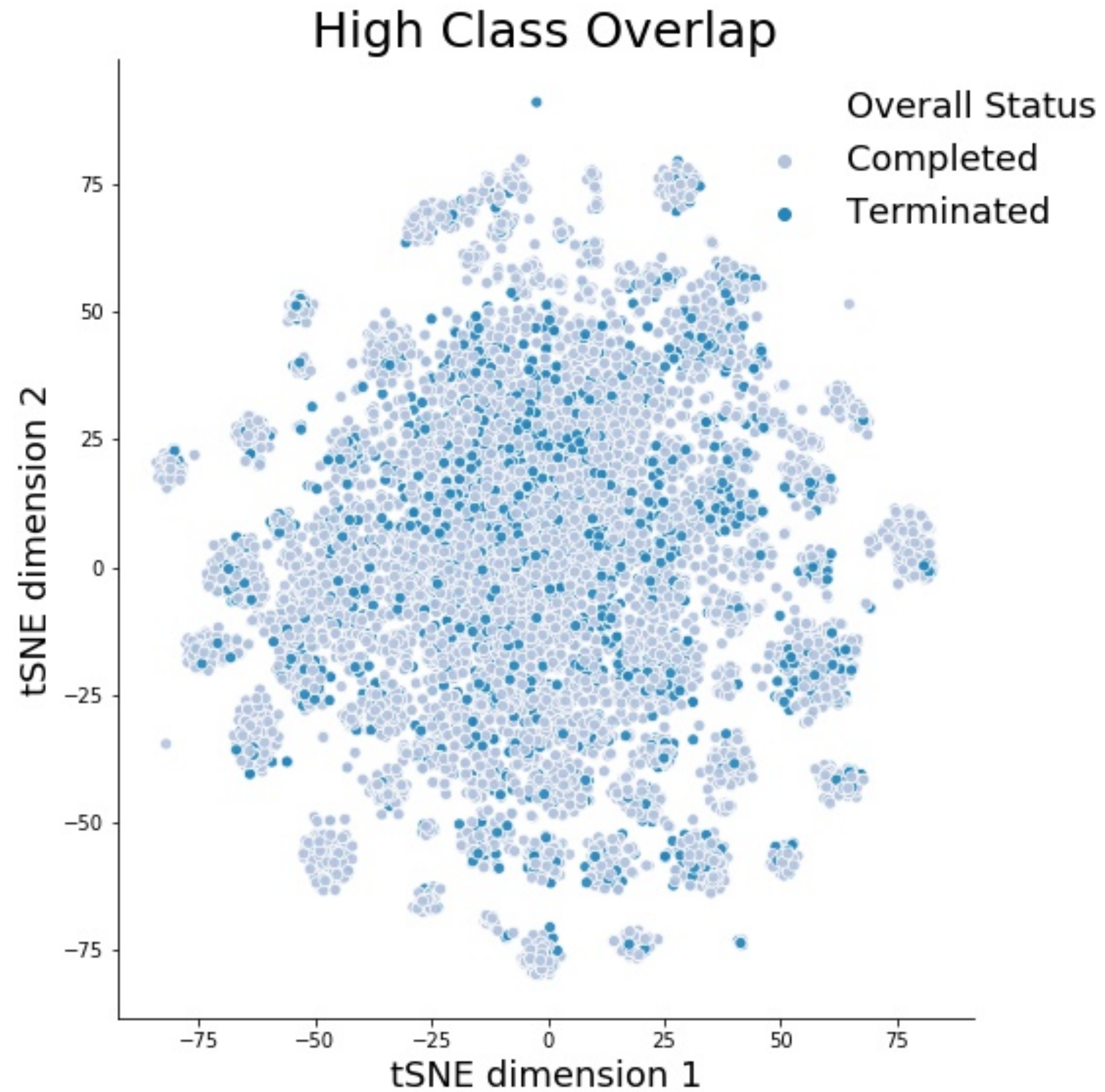
Phase II

Machine Learning
Models Trained
to Classify
Terminated Trials

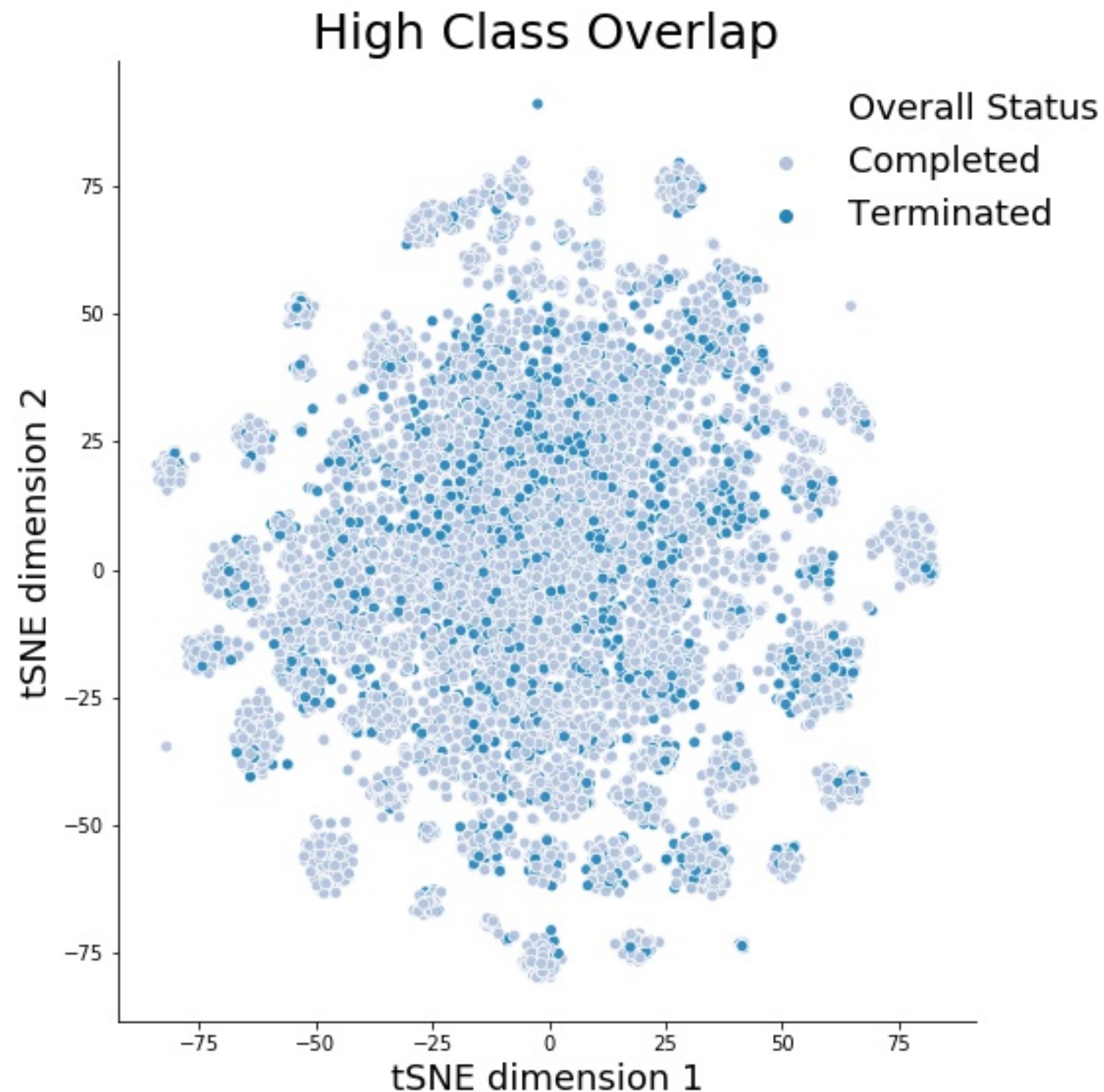
Phase III

App for
Trial Designers
to Estimate
Termination Risk

A difficult classification problem



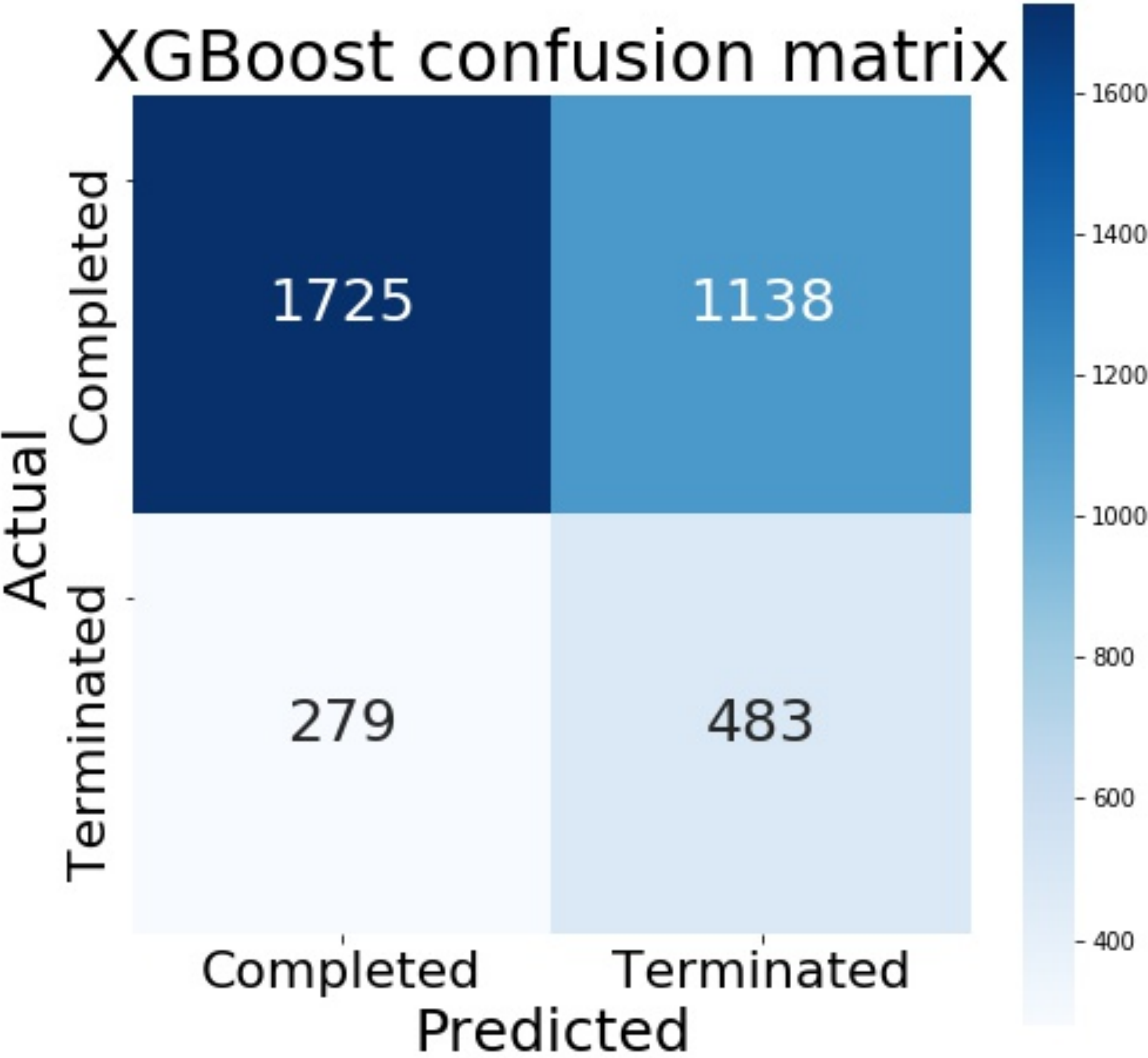
A difficult classification problem



- Optimizing sensitivity results in default model that calls all 'Terminated'
- F1 score = harmonic mean of sensitivity and positive predictive value
- Optimizing F1 score marginally improved class separation
- F1 score improved from 0.35 in default model to 0.41 in best model 'XGBoost'

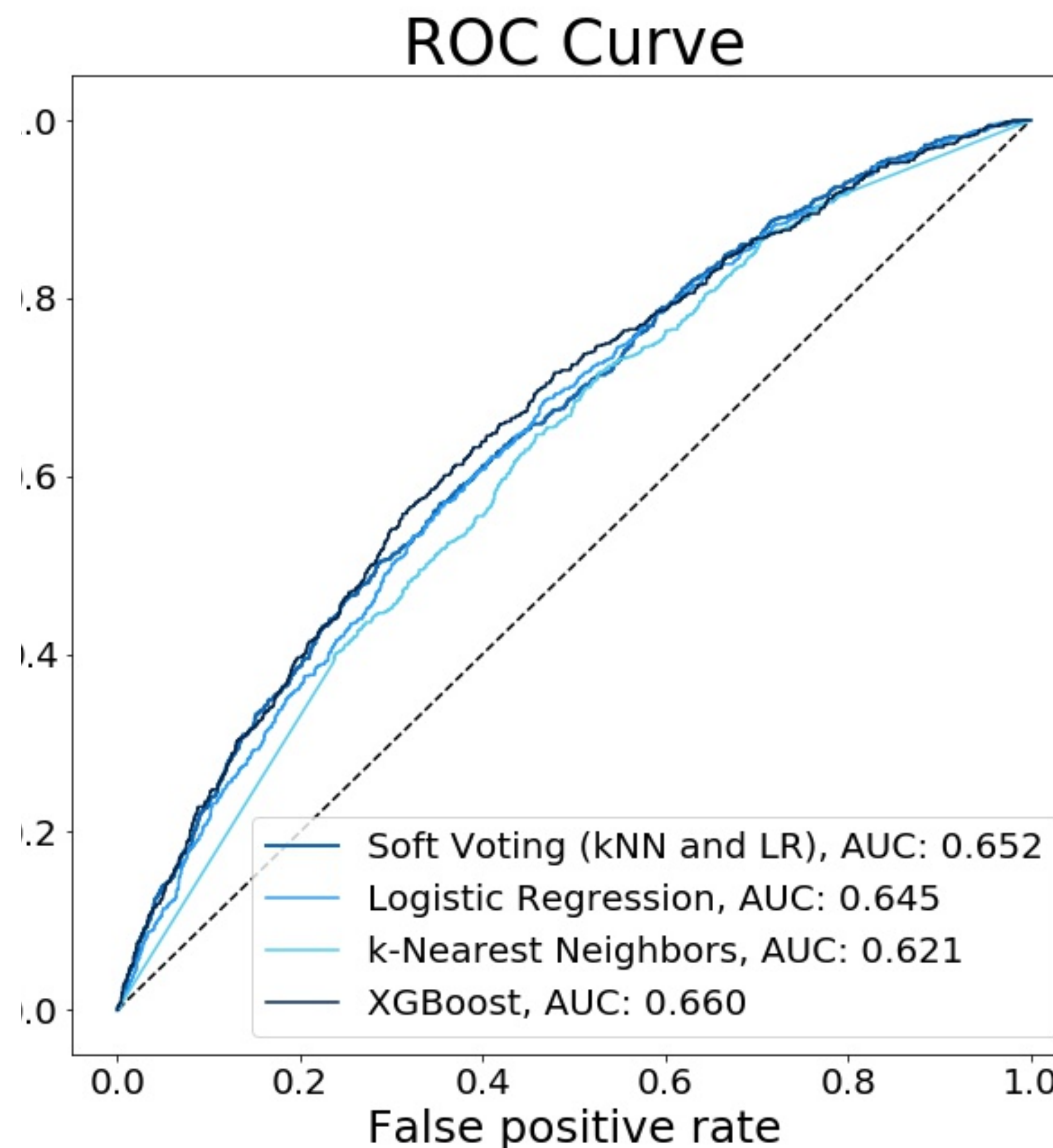
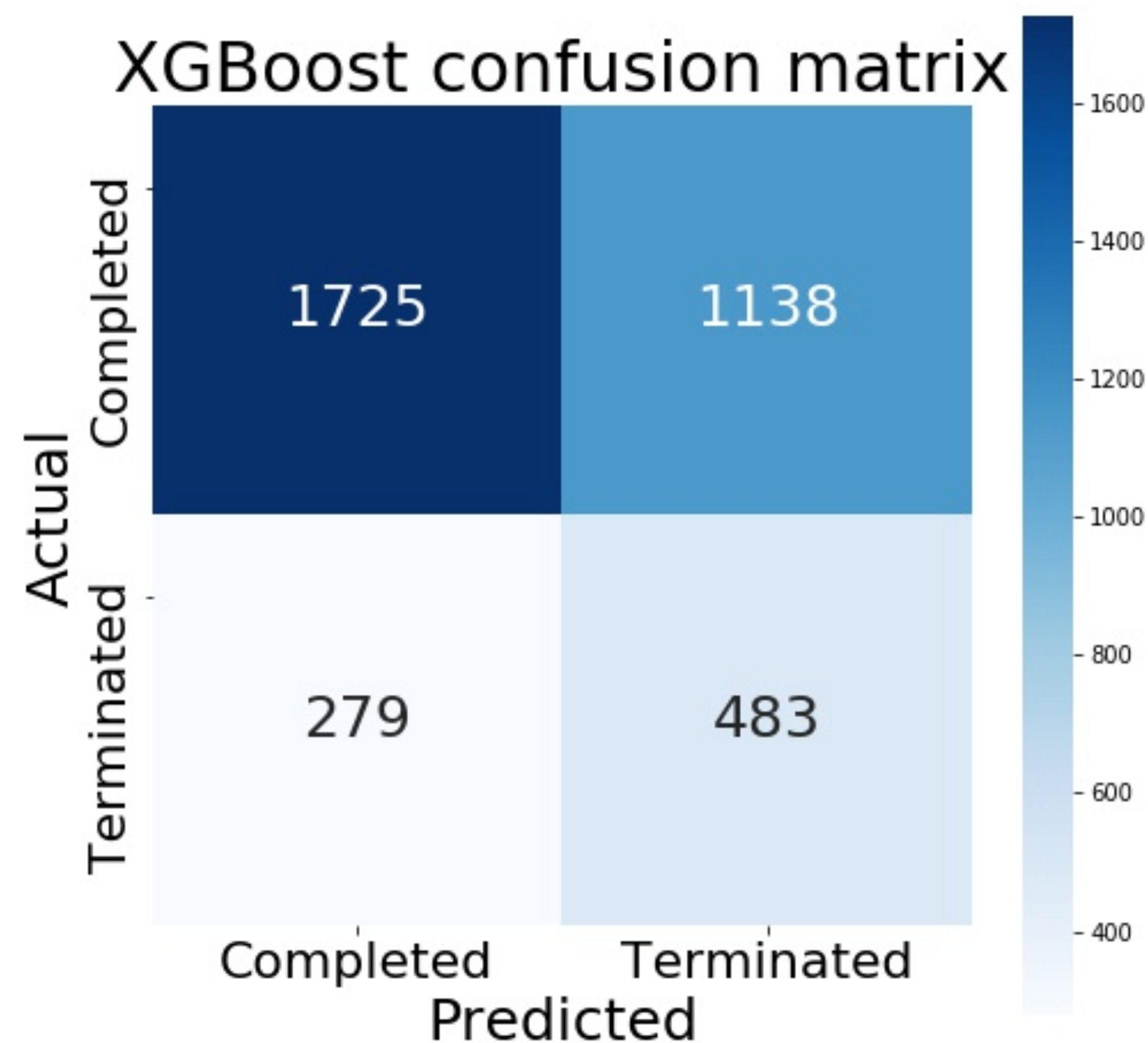
Model Performance

Unseen test data



Model Performance

Unseen test data



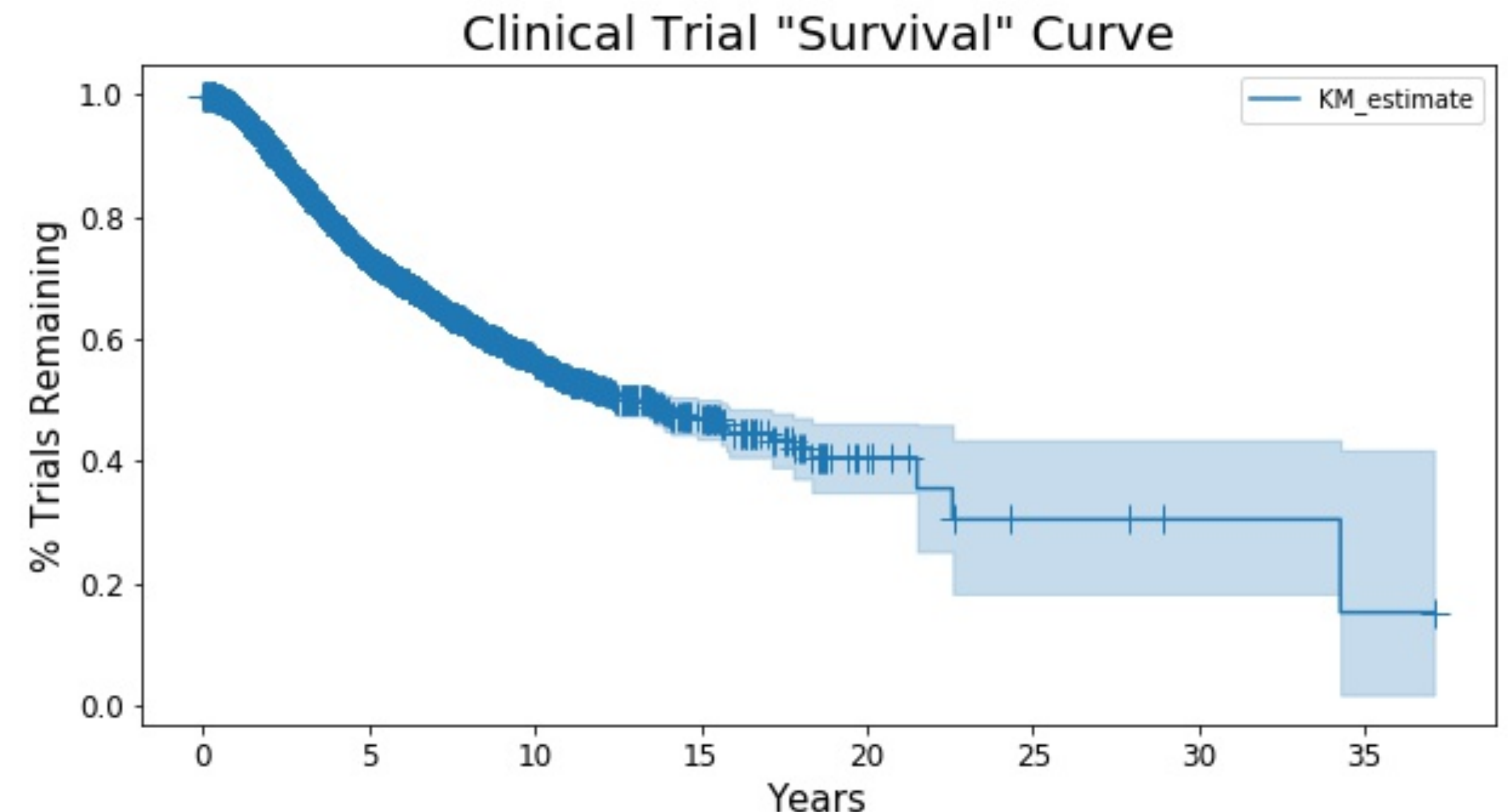


App Demo Time!

Future Directions



- Neural networks - Can F1 score be improved?
- More detailed (but paywalled) data sources - PharmaIntelligence
- Incorporate data on chemical structure and target of drug interventions
- Predict specific reasons for terminated trials
- Use time to failure or cox proportional hazard model coefficients of features as weights in model development





Thank you
Please ask me any questions