

Model Comparison in Credit Card Fraud Detection

Chris Shurtleff
University of Tennessee

Introduction

- Fraud cost the U.S. economy approximately \$16 billion in 2016.
- All major credit card providers and most banks automatically track fraud using data mining techniques.

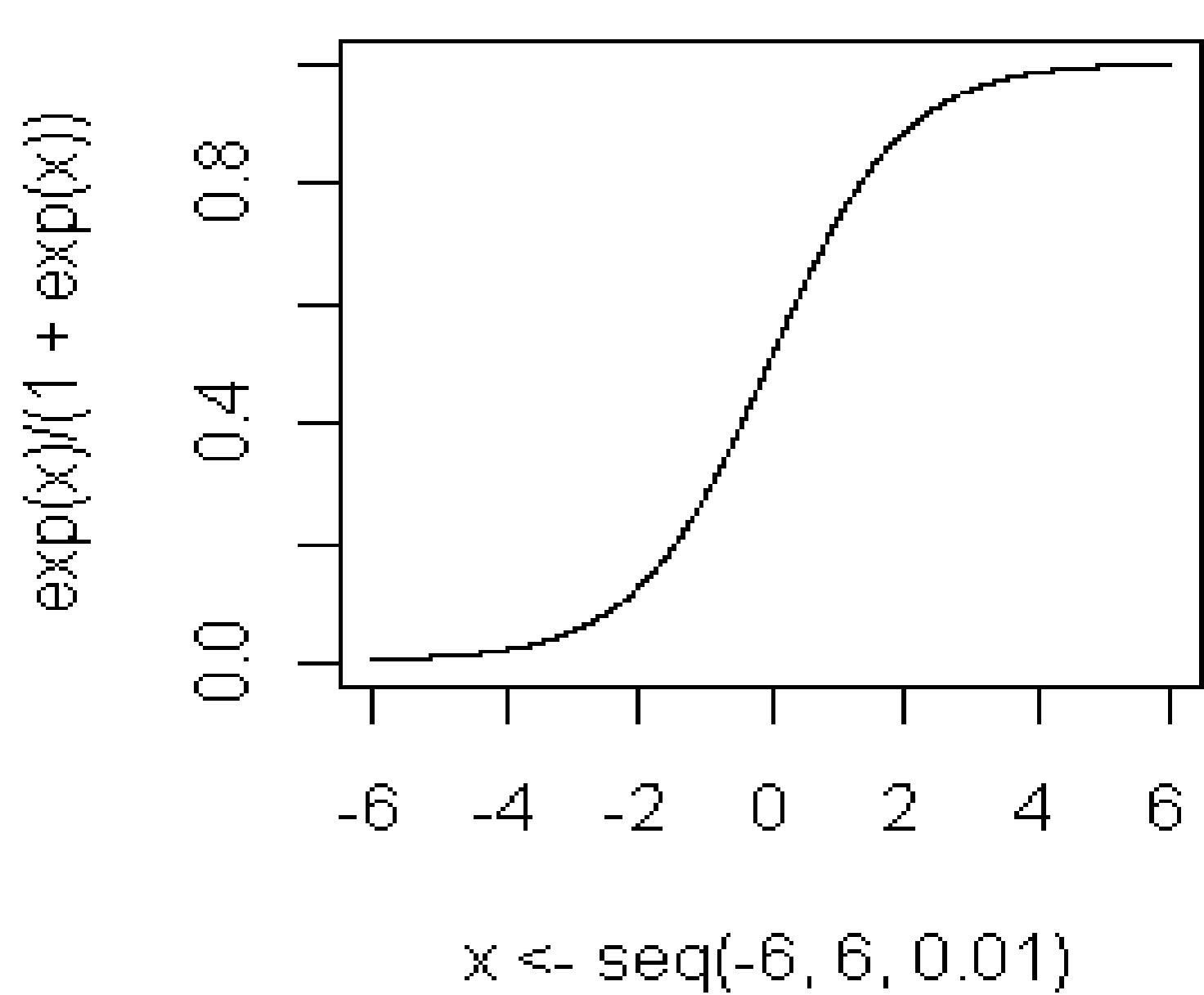


Measurement Statistics

- Accuracy is a poor measure (Assuming no fraud = 99.8% accuracy!).
- Area Under Receiver Operating Curve (AUROC).
- Assuming no fraud AUROC = 0.5, indistinguishable from guessing.
- $AUROC = \int_0^1 \frac{TP}{P_o} d \frac{FP}{N_o}$

Logistic Regression

- AUROC = 0.981.
- Easy implementation.
- Computationally inexpensive.
- Competitive performance.
- Best predictors: V10, V27, V4, V22, and V14.

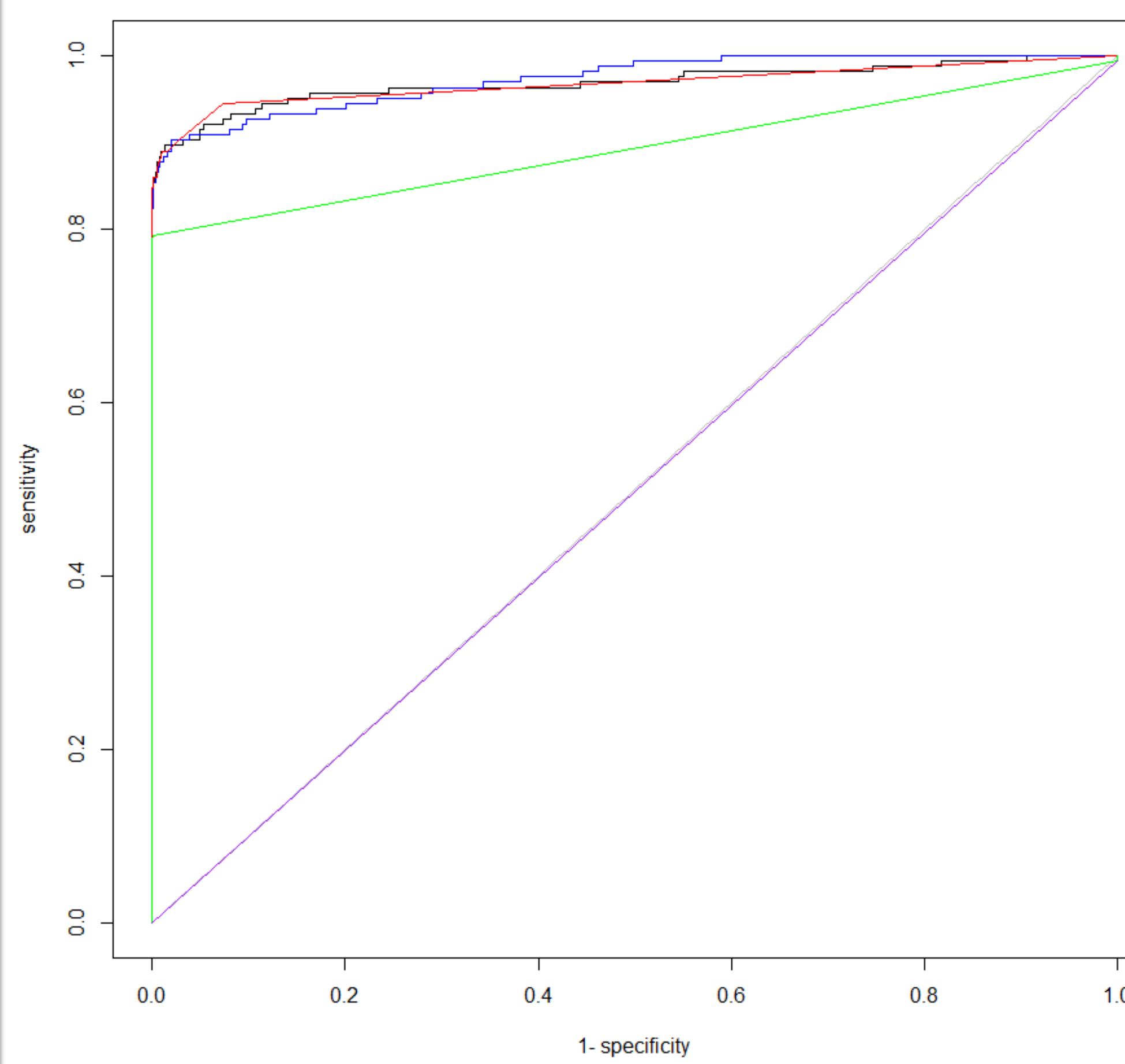


References

- Andrea Dal Pozzolo, Olivier Caelen, Reid A. Johnson & Gianluca Bontempi. Calibrating Probability with Undersampling for Unbalanced Classification. In Symposium on Computational Intelligence and Data Mining (CIDM), IEEE, 2015 .
- <http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/> Source: Nature
- <http://www.zdnet.com/article/identity-theft-credit-card-fraud-cost-us-consumers-16-billion-in-2016/>

Results

- Logistic regression wins! (blue).
- Neural Network (black) and Random Forest (red) are highly competitive.
- Decision tree (green) is much less competitive.
- No fraud model (purple) indistinguishable from guessing.
- Possibly little difference between competitive algorithms

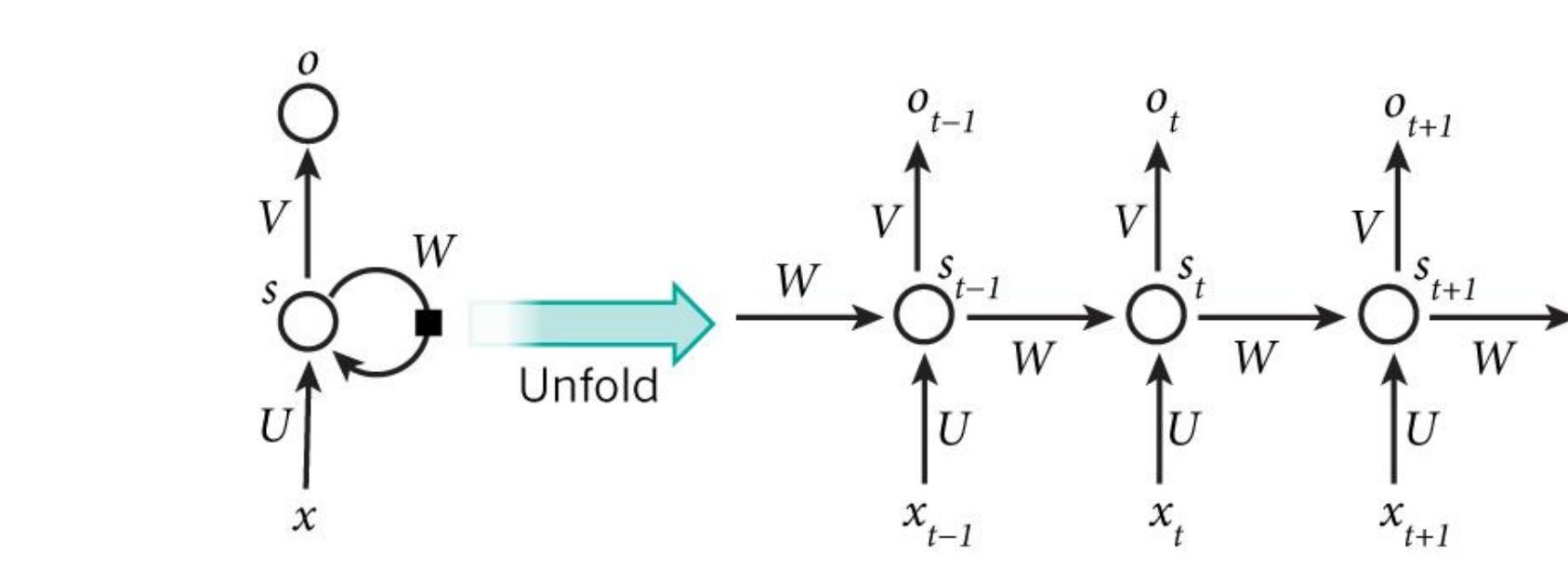


Neural Network

- AUROC = 0.968.
- Computationally expensive.
- Tends to overfit.
- High Performance.
- More hyperparameters to tune.
- Best target for improvement (number of hidden layers, number of nodes per hidden layer, convergence limits)

Future Work

- Precision-recall curves.
 - Possibly better measurement statistic.
- Better models:
 - Deeper neural network.
- Better features:
 - Deep learning feature selection (RNN).
 - Not possible with current dataset (PCA).
- Scalability:
 - Scala or Python for production purposes.

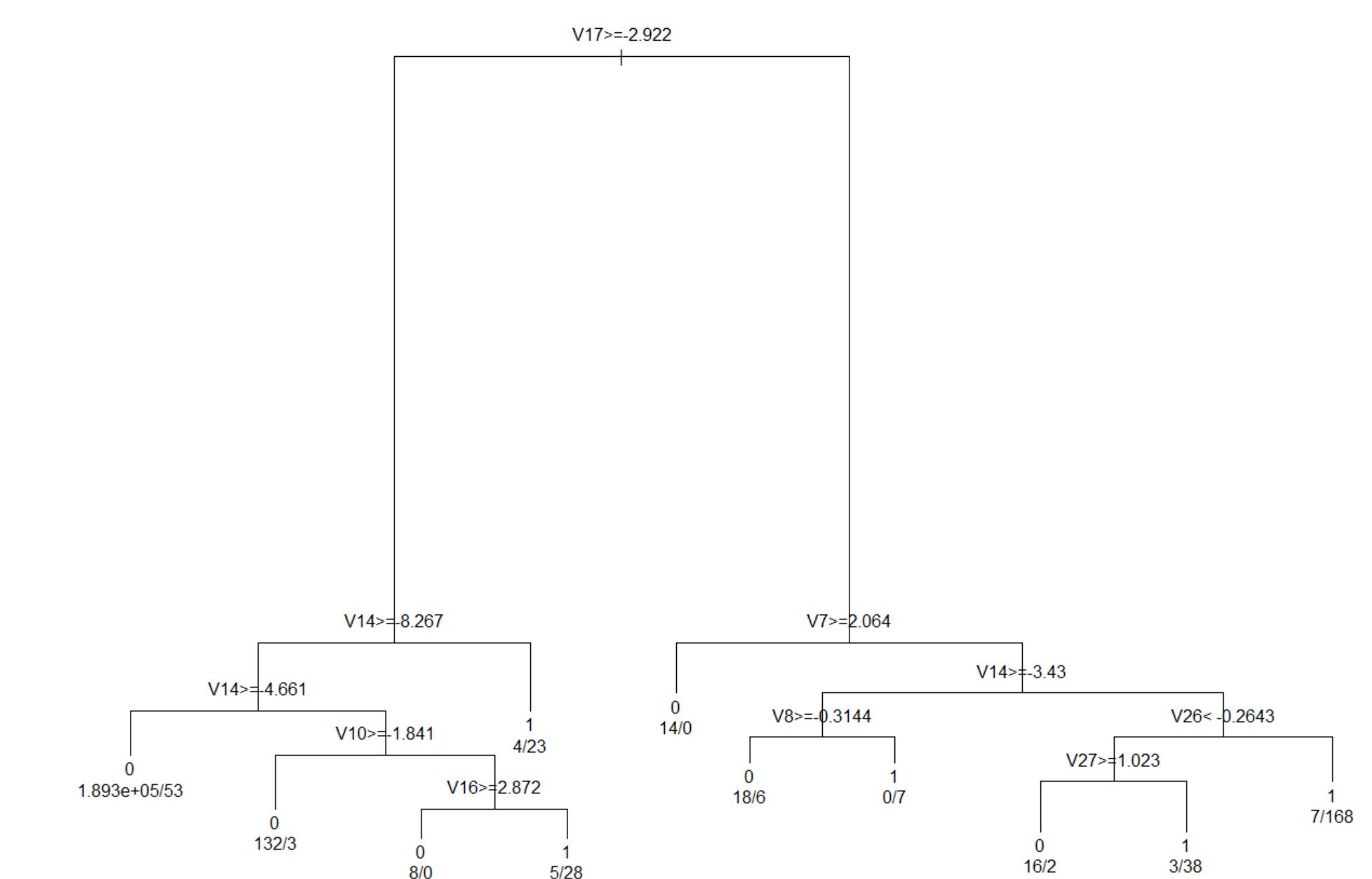


Data

- 284,807 transactions, of which 492 were fraudulent (0.17%).
- Approximately €25,163,000 worth of transactions .
- 28 anonymized features (PCA).
- Transaction value and Time (from unspecified beginning) left unaltered.
- Anonymization precludes interesting analyses (RNN).

Decision Tree

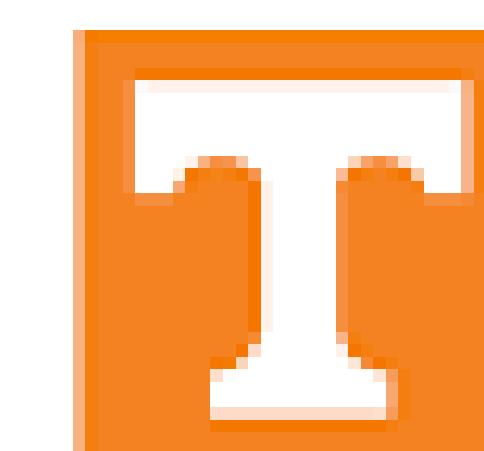
- AUROC = 0.893.
- Easily readable.
- Tends to overfit.
- Relatively poor performance



Random Forest

- AUROC = 0.968.
- Easy implementation.
- Competitive performance.

0.5 cutoff	P _P	N _P
P _O	131	33
N _O	12	94760
0.1 cutoff	P _P	N _P
P _O	139	25
N _O	54	94718



THE UNIVERSITY OF
TENNESSEE
KNOXVILLE