Credit Card - Fraud Detection

Patrick Schneider

Date:15.12.2018

Objectives

1. Introduction
    a. Data set
    b. Background Credit Card Fraud
2. Optimization Metric?
3. Evaluation
    a. SVM Kernel
    b. Different prediction algorithms
    c. Sampling methods on XGB
4. Conclusion
5. Future Work

Data set 1/2 - Source

- Worldline and Machine Learning Group of Universite Libre de Bruxelles
- Paper: Calibrating Probability with Undersampling for Unbalanced Classification*

*https://www3.nd.edu/~dial/publications/dalpozzolo2015calibrating.pdf

Data set 2/2 - Data set

284.807 transactions

492 Fraud cases (0.172% of all transactions)

Features: 30 pca attributes + 'Time' and 'Amount'

Target Class label = 0/1

Characteristics:

- no outlier
- no NAN values

Background Credit Card Fraud

Online Shopping - large and fast growing trend

Online Payment not requires physical card

Risk - credit/debit card detail is known to others

Cardholder comes to know only after the fraud

Background Credit Card Fraud

- 2016 - 6 billion $ lost to bank fraud in US alone

- 20 percent of customers change their banks after experiencing scams

source:
https://www.altexsoft.com/whitepapers/fraud-detection-how-machine-learning-systems-help-reveal-scams-in-fintech-healthcare-and-ecommerce/

Optimization Metric?

What to optimize

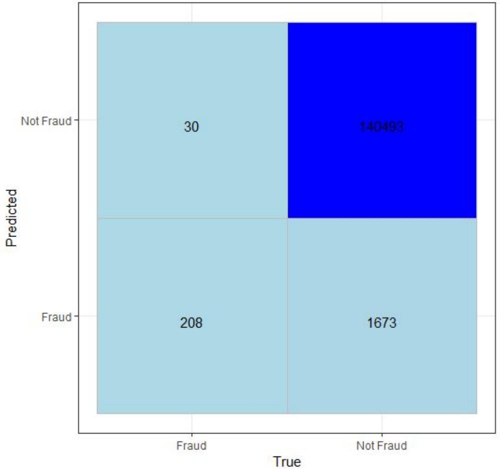
- Fraud capture rate?
- False positive rate?
- Precision recall?
- ROC?
- F1-score?

Assumption:

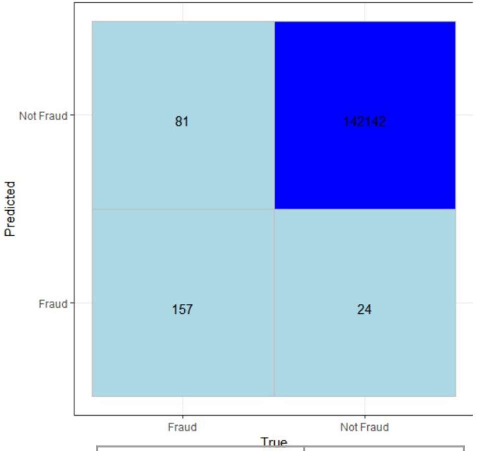
- False Negative create direct financial cost with a fraud
- False Positive creates indirect financial costs with valid payments that are blocked->unsatisfied with service->switch institute

Total Loss COST= 10*SUM(FN) + SUM(FP)

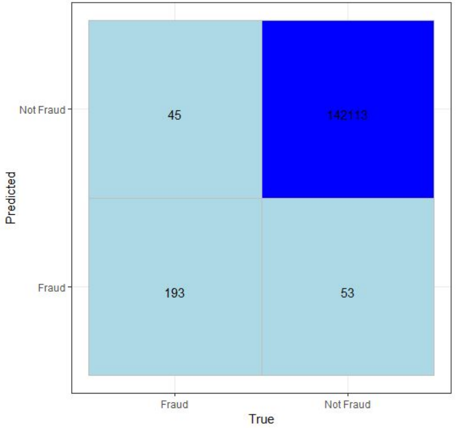
Metric selection: Trade-off



Recall	0.87
Precision	0.988
Cost	1973
F-score	0.92



Recall	0.66
Precision	0.999
Cost	834
F1-score	0.79



Recall	0.81
Precision	0.999
Cost	503
F1-score	0.89

Evaluation: SVM Kernel

Method	Best AUC	Own Cost function AUC (cost)	SMOTE
linear (Weights)	0.86	0.86 (803)	1000
Radial (Weights)	0.93	0.85 (813)	(1641)
poly	0.89	0.90 (521)	0.90 (1038)
radial	0.91	0.88 (581)	942
linear	0.88	0.88 (567)	

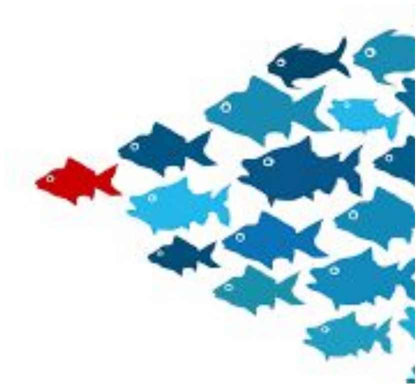
Testset -Prediction result AUC

Method	Imbalanced	SMOTE result	cost based
GLM	0.91	0.87	0.87
Decision Tree	0.89	0.89	0.88
RF	0.92	0.94	0.91
SVM	0.85	0.93	0.91
XGB	0.91	0.94	0.92

train = 20%
val = 40%
test = 40%

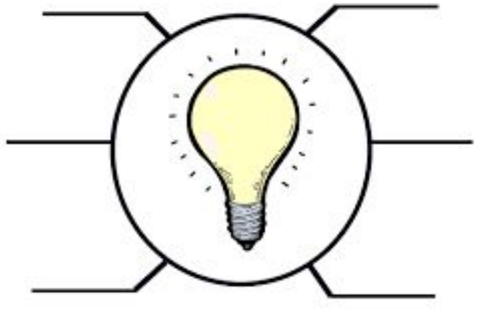
Evaluation: Sampling Methods on XGB

Method	AUC Score	AUC own cost function	(threshold)
unbalanced	0.913	0.913	(0.15)
up	0.93	0.910	(0.6)
down	0.90	0.87	(0.8)
smote	0.94	0.92	(0.95)
rose	0.900	0.900	(0.72)



Conclusion

- Definition of business/cost metric is crucial
  - Trade-off: Higher detected fraud cases =>more unnecessary blocked legit transactions
- Sampling method needs to be considered wisely
  - depending on evaluation metric, results can vary
- Bad classifier not be excluded based on unbalanced data prediction results
- Big differences in training time



Future Work

- Improve feature selection
  - Feature selection
- Class Imbalance
  - Combine methods e.g: SMOTE + ENN
- Evaluation of other prediction algorithm
  - Hidden markov model
  - DNN
- Model ensemble methods
- Tuning of Algorithms

BACKUP: example