RSK-Mind: Machine Learning for Transactions

The Fraud Detection Scenario

RSK-Project

Fintech Crowdhackathon

April 22-23, 2016

The team

Who we are?

- 4 Software Developers BSc (IT):
- 1 Ph.D. candidate in Machine Learning at Ecole Polytechnique, Paris
- 1 MBA holder with 12 years in investment services (derivatives, international markets, portfolio management)
- Competition: 6th place in Data Mining Cup 2013 (biggest student machine learning contest)
- Participation in many Kaggle contests

The problem

Interesting data

- Exponential increase of web transactions
- In 2014, from 1 out of 114 the number of fraudulent transactions rised to 1 out of 86 in 2015
- 30% increase in fraudulent transactions (2014-2015)

Why is it crucial?

- Banks remain the guards of economic system, so they need to adjust and follow the new demands
- Fraud creates not only economical damage but loss in trust of brandname
- In case a bank declares a transaction as fraud, the e-shop is obliged to return the money
- Fraud is an adaptive crime. It evolves

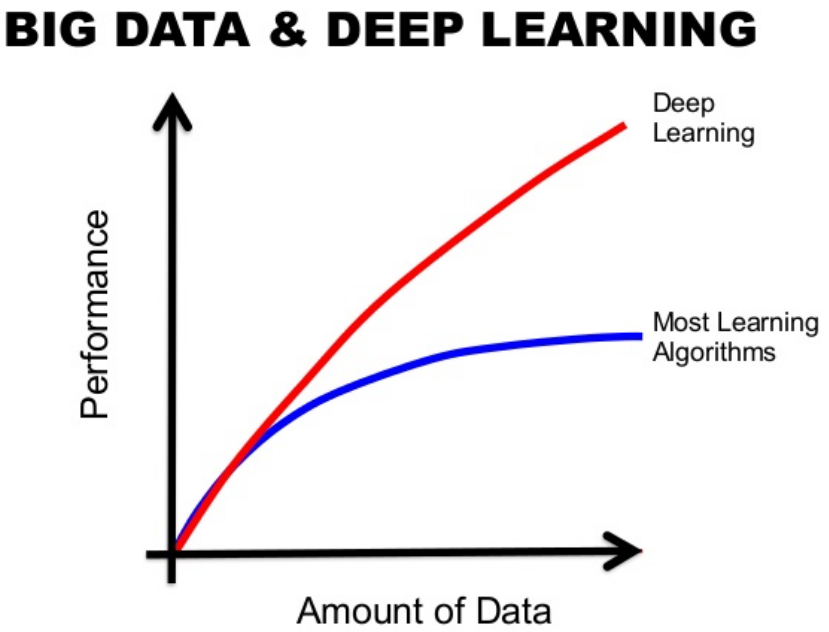
Fraud Detection in Industry

Evolution of Fraud Detection

- Large banks and e-shops use rule-based control of transactions. Based in statistical analysis and manual selection of experts in security.
- The disadvantage of these solutions do not respond to new types of fraud attempts. Success rate: 70-90%
- During the last years, **Machine Learning** (AI) approaches arise, systems that learn themselves on transaction history. Success rate: $>>$ 90%

Deep Learning for Fraud Detection

- Paypal introduced a paper on DL for Fraud Detection
- More data improves results.



The actual cost

Our solution (1/3)

RSK-Mind

The solution our team proposes is a platform based on Deep Learning. It works with open architectures, feeds on data coming from OpenBank API, which are enriched by metadata coming from installed plugins on the payment system of the customer. Our goal is to further enrich the data we want to evaluate, in order to achieve bigger rates of successful prediction of malicious transactions.

Key points of RSK-Mind framework

- A machine learning framework
- Features: transaction data, geolocation data, proxy and Tor detection, browser information (system timezone)
- Biometric data for user behaviour recognition (mouse-move record during the payment process). Leads to better bot detection
- Works with open APIs (OpenBank API) and restricted ones
- Fast and easy integration to customer's platform with our API

Our solution (2/3)

Key points of RSK-Mind framework

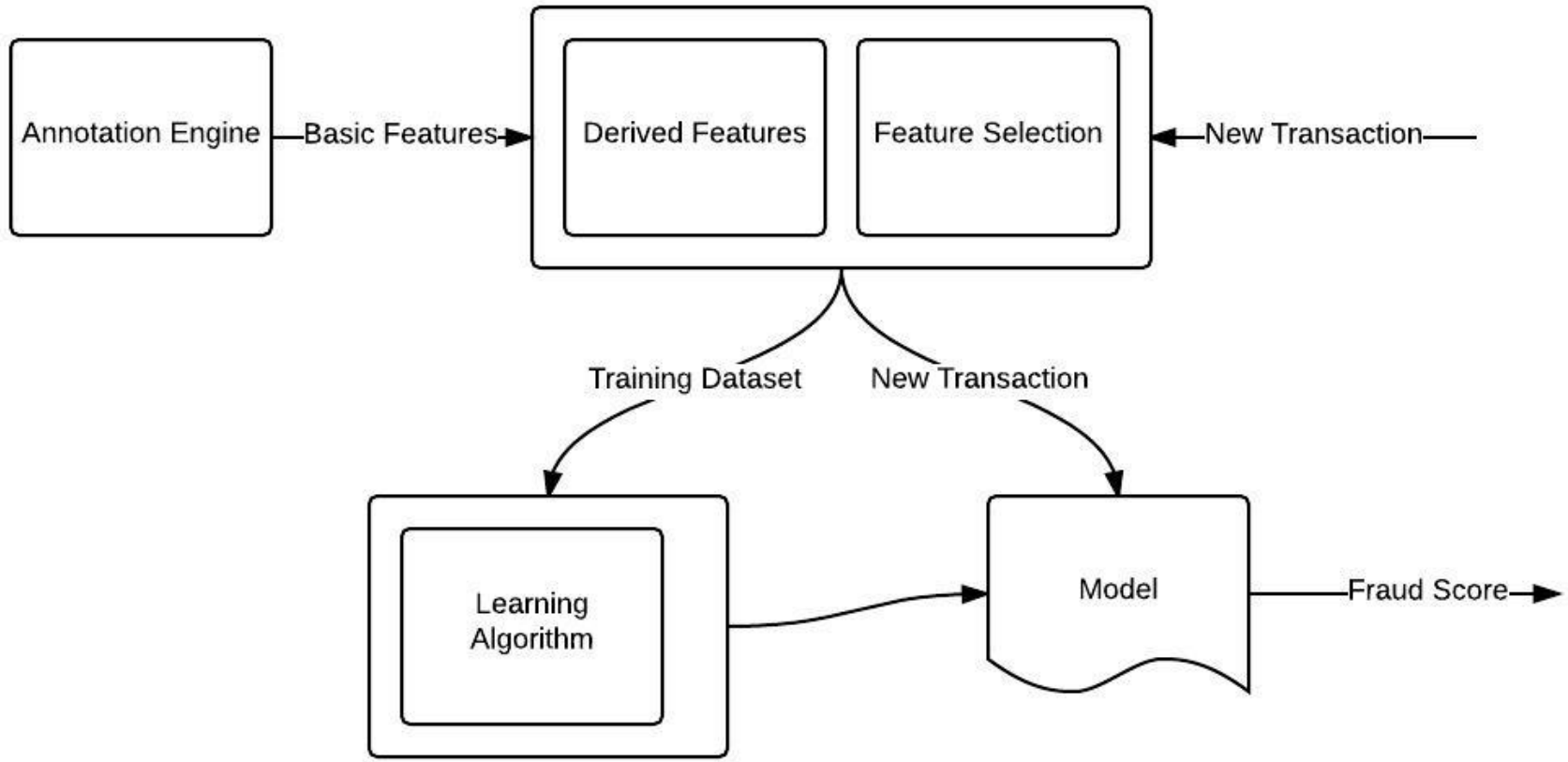
- Real-time prediction of an incoming transaction as fraudulent (fraud-score). For example, returning a fraud score of 30 states a 30% probability that the questioned transaction is fraudulent. The customer will set the actual threshold of accepting a transaction as fraudulent
- More transactions translate to better evaluation and increase in successful prediction
- Dashboard with analytics on given transactions, success rates, history logs, graphs
- Low cost solution based on API usage. Fix cost per transaction lower than 0.5 cents per transaction make it affordable even to small e-shops

Our solution (3/3)

Key points of RSK-Mind framework

- Modular platform: replace or add parts of our framework to meet customer's demands and even extend it
- Our platform can work efficiently on anonymized data. No need in real names of customers' transactions, bank accounts, detailed addresses. With proper transformation by the customers (e-shop or bank), they can be anonymized before the data analysis and learning process
- Plugin solution to major e-commerce frameworks like magento, presta-shop, open cart

The pipeline



Dashboard


Dashboard
Search projects...
Rated by Customer Transactions
2.045k
Rated Transactions
69k
Suspicious Transactions
4.560
Rated by Customer Transactions Overview
Min Fraud Amount: €50
Max Fraud Amount: €1.560
0.90€ Min Transaction Amount process
2.005€ Min Transaction Amount process
Total: €30.000.000
Iran 10%
Russia 14%
China 33%
Iran
Russia
China
Other
Transactions in 2016
Quick view of the total transaction of this year.
Rated by Engine Transactions 69k
69k transactions has rated by our engine
Rated by Customer 2,045k
2,045k transactions rated by customer this month

What we accomplished

During the hackathon

- Generated initial transactions set of data for the actual learning process
- Created the machine learning pipeline for transactions classification by using the open source platform xgboost, which utilizes the gradient boostring trees algorithm
- Tested the open source platform H2O by training our dataset with Deep Neural Network (RNN)
- Tested random scenarios in order to realize the success percentage of our algorithm
- Created javascript plugin that captures the mouse motion during the payment process
- Dashboard demo with graphs and analytics of the transactions (HTML5, AngularJS)
- Created API with Python/Django REST framework. API is asked with a recent transaction instance and returns a fraud score (probability of a transaction being fraudulent)

Thank You !

Questions?