E.T.-RNN: Applying Deep Learning to Credit Loan Applications

Dmitrii Babaev, Maxim Savchenko, Dmitrii Umerenkov *Sberbank AI Lab*. Alexander Tuzhilin *New York University*

Applied Data Science Track Poster N: 31, Poster time: 7:00 pm to 9:30 pm on August 6

Problem: credit scoring (predicting the probability that bank customer will default on a loan) using customers' card transactional data available to the bank.

Current approaches: hand-craft features by aggregating transactional data then use these aggregates as input for logistic regression or decision trees.

Our approach: take raw customers' credit card transactional data, embed each transaction into a latent space and use the sequence of embeddings as input for recurrent neural network.

Performance of our method: it dominates the SoA baselines used in the industry while requiring no feature engineering.

	ROC AUC	N Features
Logistic regression	0.78	~400
LGBM	0.81	~7000
E.T.-RNN	0.83	12

Benefits of our approach:

- Can make credit loan decisions in nearly real-time.
- Even a person without any credit history can be reliably accessed for credit-worthiness using his or her transactional history.
- A fair approach to credit scoring as it does not rely on personal or demographic information of an individual
- Information in the transactional data is exceptionally hard to forge
- No need for feature design