FRAUD DETECTION BASED ON GAT AND DRSA

Thursday 23rd April, 2020

FuXinlan XingYuyan ZhuChaona

Introduction and Research Status of Fraud Detection

Problem introduction

Difficulties in fraud detection:

- Difficult problem definition
- high labeling cost
- few black samples

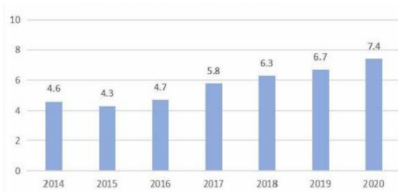- white samples are noisy;

- fraudsters are evolving

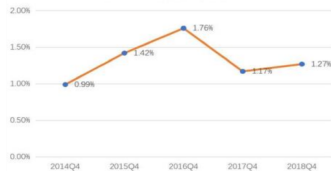Figure 1: Number of credit cards (100 million)

Figure 2: Credit card loss rate

FRAUD DETECTION

Literature review

	Supervised learning Ref[1,2]	unsupervised learning Ref[3]	semi-supervised learning Ref[4,5]
Advantage	The recall rate can reachthe limit of the quality of the existing data samples (Fully adjust the parameters)	Can discover new anomalies	The recall rate is very good,the unknown maliciousness can be found, the resistance is good, and the timeliness is also better than supervised learning
Disadvantage	The timeliness of the model is poor, and the cost of labeling and adjusting is high	Uncontrollable results and lack of interpretation	Modeling and tuning are complex

Innovation

Graph embedding(GAT) is used to combine the network structure of the relationship between users to extract features, and then the extracted features are used to train the deep survival analysis model(DRSA), and the two parts are combined to make an end to end model.

- **graph:** Different from the traditional anti-fraud model mainly based on identifying personal risks, but through the connection between users to mine the characteristics of fraudsters and identify potential group fraud.

  In the graph structure with a small number of label nodes, according to the propagation algorithm, the label categories of unlabeled nodes are predicted.

Innovation

Graph embedding(GAT) is used to combine the network structure of the relationship between users to extract features, and then the extracted features are used to train the deep survival analysis model(DRSA), and the two parts are combined to make an end to end model.

- **graph:** Different from the traditional anti-fraud model mainly based on identifying personal risks, but through the connection between users to mine the characteristics of fraudsters and identify potential group fraud.

  In the graph structure with a small number of label nodes, according to the propagation algorithm, the label categories of unlabeled nodes are predicted.

- **GAT (Graph Attention Network):** Can handle dynamic images, model is timely.

Innovation

Graph embedding(GAT) is used to combine the network structure of the relationship between users to extract features, and then the extracted features are used to train the deep survival analysis model(DRSA), and the two parts are combined to make an end to end model.

- **graph:** Different from the traditional anti-fraud model mainly based on identifying personal risks, but through the connection between users to mine the characteristics of fraudsters and identify potential group fraud.

  In the graph structure with a small number of label nodes, according to the propagation algorithm, the label categories of unlabeled nodes are predicted.

- **GAT (Graph Attention Network):** Can handle dynamic images, model is timely.

- **Deep Survival Analysis (DRSA):** The model has the ability to predict the dynamic point in time, not only can analyze the influencing factors of fraud, but also can establish an appropriate model based on the influencing factors to measure fraud time.**1**

FRAUD DETECTION

References

[1] Kim M J,Kim T S A neural classifier with fraud density map foreffective credit card fraud detection[C]

[2] 《基于 CNN 的信用卡欺诈检测》

[3] 《Unsupervised profiling methods for fraud detection》

[4] 《Analysis of credit card fraud detection methods[J]》

[5] 《基于图特征的欺诈检测方法研究与应用》