Case Study

Intel® Xeon® Processor Family
Intel® Optimization for TensorFlow*
Intel® Distribution for Python*
BigDL
Anti-financial Fraud

Research and Applications of Anti-financial Fraud Models Based on Sandwich-structured Deep Learning Framework

The National Engineering Laboratory for E-commerce and E-payment*, approved by the National Development and Reform Commission in 2013, is the first national engineering lab established by China UnionPay* in the financial industry. The aim is to build a domestically top-ranking and internationally renowned e-commerce and e-payment research base to strengthen technological breakthroughs and stay ahead in the e-commerce, e-money and e-payment technologies. ZhongAn Technology*, another rising star in the industry, is an offshoot of the insurance titan ZhongAn Insurance*. It is oriented toward the R&D of basic technologies, such as cloud computing and AI, and its businesses encompass the upstream, downstream and periphery of the financial and healthcare industries. It leads the charge in exploring innovative business models in the online insurance ecology, which are now outputting comprehensive industry solutions for users.

With rapid business expansions in the financial industry, the risk index has also risen sharply, especially in financial fraud risks. A Nielsen report on global bank cards shows that the loss rate of bank card fraud cases worldwide hit 0.0715% in 2016[2]. Underlying this high fraud rate are challenges in precision and timeliness faced by traditional risk control methods. Therefore, there is a pressing need for financial enterprises to establish newer intelligent risk prevention and control systems. In view of this, the National Engineering Laboratory for E-commerce and E-payment, ZhongAn Technology and Intel joined forces to research deep learning-based anti-fraud technologies. Drawing from their experience in rule-based methods and traditional machine learning, they came up with an innovative, multilayered fraud detection solutions framework based on a sandwich structure. The solutions are now being tested in several real contexts, such as card forgery and cash fraud, and the results have been encouraging. There is a good chance that these solutions will prove to be effective in other risk detection contexts, including transaction fraud, credit fraud and insurance fraud.

Challenges

New challenges in the battle against financial fraud: The arrival of the Internet era has rocked society as financial fraud cases become more frequent, precise and tricky. The traditional ways and models of combating fraud must be improved to handle these new challenges.

Inadequate learning of sophisticated transaction features: When the rule-based and machine learning anti-financial fraud model is learning complex serialized transaction features, the effect is below expectations. At the same time, a standalone deep learning method is also showing limitations in the learning of features within a single transaction.

Solutions

Sandwich-structured fraud detection model: The National Engineering Laboratory for E-commerce and E-payment, ZhongAn Technology and Intel proposed the innovative GBDT→GRU→RF sandwich-structured fraud detection model framework, which is able to overcome the inadequate learning of serialized transaction features and single transaction features via a multilayered learning approach.

Comprehensive support from Intel® technologies: Besides powering the new model with the high-performance Intel® Xeon® Scalable processors family, Intel also provides targeted technologies and tool optimization for the three-layered solutions in the model, thereby enhancing the overall working efficiency of the model.

Results

Innovative anti-fraud model as a benchmark for the industry: The National Engineering Laboratory for E-commerce and E-payment, ZhongAn Technology and Intel made innovative use of a multilayered deep learning method to boost the performance of the anti-financial fraud model. The system has been tested and the results show that the solutions are feasible. It also demonstrates the further applications and innovation of advanced technologies, such as deep learning, in the financial industry.

Applications of Intel technologies in the financial world: The applications and success of the sandwich-structured fraud detection model architecture proved that Intel products and technologies can give AI applications a positive boost in the financial industry. In the meantime, the experience gained from these programs can help Intel improve its own products and technologies.

As we enjoy the convenience in life brought about by the financial industry, we are also vulnerable to the increasing risk of fraud. Besides traditional scams, such as credit fraud, credit card theft, malicious cash advance and insurance fraud, crimes that have emerged with the Internet era, including personal data leaks and hacking, have become more rampant with fraudsters, e-connoisseurs and cyber scammers that strike with higher frequency and precision. Taking credit cards for example, there has been a continuous revival of fraudulent tactics as traditional risks blend with new types of risks. As seen in Figure 1, since 2010 the loss rate of global bank card fraud has been climbing with growing fraud losses. Fraud losses in 2016 amounted to US$24.71 billion, which is a 60%[1] rise from 4.5BP in 2010.

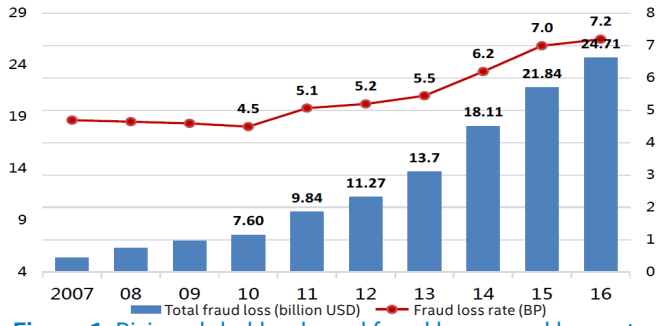

Figure 1 Rising global bank card fraud losses and loss rates

To tackle these problems, the National Engineering Laboratory for E-commerce and E-payment, ZhongAn Technology and a number of titans in the financial industry have been launching various actions against financial fraud. They have teamed up with Intel to introduce cutting-edge AI technologies into the anti-fraud field and construct a series of highly efficient and reliable anti-fraud models.

Innovative Sandwich-structured Fraud Detection Model

Currently, the financial industry primarily uses two methods to detect transaction fraud risks: Rule-based and machine learning-based algorithms. The rule-based method works by continuously establishing and renewing the rule library, which is based on transaction behavior features, and at the time of transaction, it will distinguish the underlying risks in the transaction by querying the rule library. For instance, when there is a big payment made at a convenience store, the rule library will conduct a match to see if this transaction behavior has any abnormal features. The rule-based anti-fraud method is based on the summarization and generalization of previous transaction fraud experiences. It mainly relies on the experience of experts, which will contain subjective elements and inevitably result in omissions. Hence it becomes inadequate in this day where transaction frauds come in all shapes and forms.

By comparison, the machine learning-based method uses a more objective and accurate method to tackle fraud. Some of the good classification algorithms in machine learning, such as LR (Logistic Regression), RF (Random Forest) and GBDT (Gradient Boosting Decision Tree), allow better learning of some underlying fraud models, and have been applied to the anti-fraud model of the National Engineering Laboratory for E-commerce and E-payment, ZhongAn Technology and other organizations. But in an actual transaction context, feature engineering may be more complex. For instance, when a credit card that has long been unused suddenly shows a big sum payment behavior at a convenience store at late night, or when a patient with viral flu often visits a doctor and is being prescribed with expensive medication. In such cases, complex features, such as "late night", "long been unused", "convenience store", "viral flu" and "big sum", will appear which pose a challenge to traditional machine learning methods.

To tackle this challenge, engineers from the National Engineering Laboratory for E-commerce and E-payment and Intel carried out a prophase process modeling based on technologies, including BigDL library and Spark PipeLine*. These technologies are able to streamline complex feature learning work and enhance the final model effects. During the modeling process, the engineers derived hundreds of feature factors from a small number of original fields, and formed feature variables in six dimensions, such as long-term and short-term statistics and credibility, to help the model learn better.

But during actual testing and application, some problems showed up in the model. One problem was that the modeling of feature engineering and computing is reliant on the experience of experts, and the analysis of longitudinal transaction behavior was ineffective as there is no transaction serialization analysis on the machine learning-based methods. Therefore, the National Engineering Laboratory for E-commerce and E-payment, ZhongAn Technology and Intel furthered their cooperation and tried to use deep learning methods to automatically learn the features relating to transaction serialization.

In their solutions, first of all they used the unitary RNN (Recurrent Neural Networks) method and adopted the LSTM (Long Short-Term Memory) method or GRU (Gated Recurrent Unit) method to conduct fraud detection modeling directly on transaction data, but the results were not satisfactory. This is because although RNN is able to learn the feature correlation between transaction sequences, the learning ability for features within a single transaction can only be on par with that of a traditional shallow neural network, which falls short of expectations.

In order to let the new model deepen the learning of the feature correlation between transaction sequences, experts from the three partners proposed an innovative, multilayered hybrid fraud detection model framework. This framework makes use of a sandwich structure in this order: GBDT→GRU→RF. As indicated in Figure 2, the framework will first take aim at the inadequate feature learning ability within a single transaction in the unitary RNN method and perform further feature optimization by adding the GBDT model to the front end of the framework. The optimized features and artificial features will be combined as input for the GRU network to enable learning of inter-sequence features and sequential features in a single transaction.

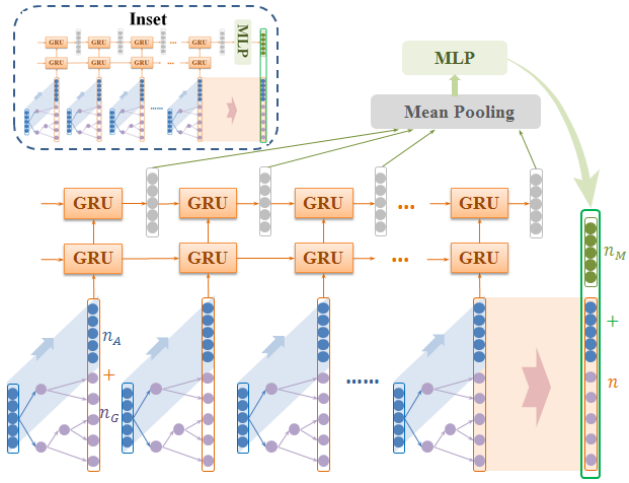

Figure 2 Front part of GBDT→GRU structure

In the middle layer, the framework does not directly use the GRU network output to distinguish fraud detection, but rather as a step in inter-sequence feature learning. The inter-sequence features learned will be combined with the original features in the transaction to form a final transaction feature vector. On this basis, in order to further conduct combined learning with the sequential features, the framework will lastly add a top layer on the RF model as a final classifier to distinguish fraud. The overall structure of this framework is indicated in Figure 3.

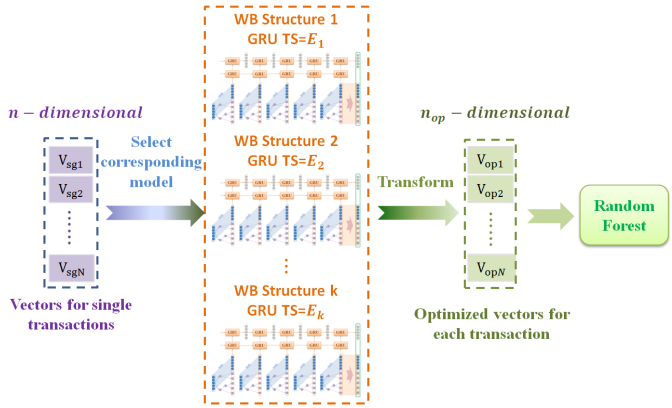

Figure 3 GBDT→GRU→RF sandwich-structured fraud detection model framework

For the new deep learning fraud detection model, engineers from the three companies made use of real data to conduct several simulation verifications based on common application contexts, such as card forgery detection and insurance fraud in banking and insurance businesses. In terms of the recall rate or precision rate, the GBDT→GRU→RF sandwich-structured fraud detection model met the expectations. Compared with the traditional classifier approach or unitary RNN method, the F1 value of this model (a type of weighted average of precision rate and recall rate to measure the performance of a detection model) shows significant growth. Under partial context, it can obtain results optimization of more than 1.5 times.

As an illustration, let's take the example of the National Engineering Laboratory for E-commerce and E-payment adopting the GBDT→GRU→RF sandwich-structured model in the context of fraud detection in credit card forgery. Assuming an account displayed a number of small sum payments late at night, followed by a big sum trans-regional transaction, through the preceding "GBDT→GRU", the sandwich-structured fraud detection model is able to learn the sequential relation similar to that between integrated features 1 ("late night" + "small sum") and integrated features 2 ("late night" + "big sum"). This function can be called "sequencing of integrated features". With the addition of the subsequent "GRU→RF" process, it can further learn abnormal features similar to the further combination of sequential features 1 (testing of several small sums + big sum cash out) and sequential features 2 (current transaction location and previous trans-regional transaction locations). This function that "integrates sequential features" can further deepen the feature learning ability of the model. Figure 4 displays the actual test results in the context of credit card forgery detection. The valid F1 value may exceed 0.4.

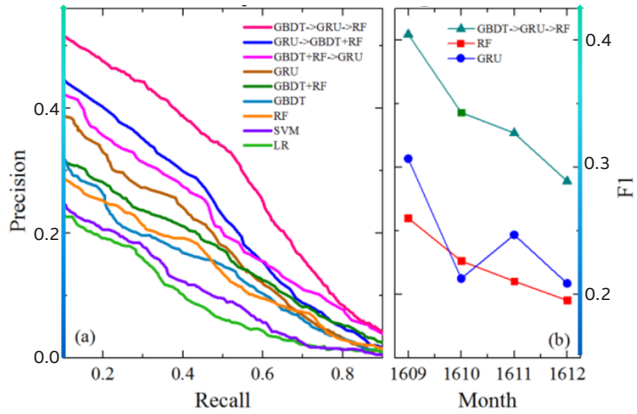

Figure 4 Assessment results of GBD→GRU→RF sandwich-structured fraud detection model

From Model to Application

1. The IAaaS (Intelligent Analysis as a Service) platform of the National Engineering Laboratory for E-commerce and E-payment makes use of the lab's edge and experience in several areas, including big data, AI, machine learning modeling and graph computing, to encapsulate and integrate interface contents of existing business models, algorithms, input and output interfaces. It also provides intelligent analysis services and solutions to external parties. By calling the encapsulated API interface, users can obtain data of the results after model algorithm analysis, so as to better serve the front-end business department and partners to robustly support innovative payment business. At present, the platform has been providing API interfaces such as public security investigation, e-connoisseur query, query for terminal in illegal replacement and evaluation service of marketing results etc. To better utilize the application value of the "sandwich" deep learning framework, as indicated in Figure 5, the National Engineering Laboratory for E-commerce and E-payment deployed this framework to the IAaaS to provide detection services for card forgery, cash advance and other fraudulent behaviors based on this framework. Staff need not study this complex model algorithm framework. All they need is to call API according to the data interface specification to get the service.

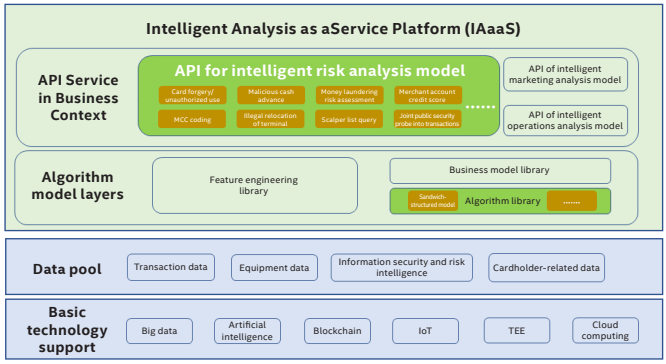

Figure 5 IAaaS of the National Engineering Laboratory for E-commerce and E-payment

2. ZhongAn Technology also adopted the GBDT->GRU->RF sandwich-structured fraud detection model to help the Ministry of Human Resources and Social Security identify user behavior of insurance fraud. This effectively raised the model's distinguishing precision rate and coverage rate, with the final F1 going up to 0.591. The model is able to significantly reduce wastage of medical insurance resources, alleviate strained medical and social insurance funds and direct the money to people who really need it. Besides this, the serialized modeling of health screening reports and consultation/diagnosis information for medical insurance – the information of which can be used for a longer period of time (5-10 years) – can be used to distinguish and forecast users' health conditions more comprehensively. By predicting users' health conditions, the insurance company can measure risks more precisely to bring about a revolution in areas such as risk-based pricing, cost budgeting and fraud distinguishing. On top of that, the insurance company can also provide its clients with more suitable and targeted healthcare recommendations and services, and turn passive health protection into proactive risk control in order to specifically discover and prevent diseases. In short, as data collection technology matures and deep learning models, especially serialized models, gradually optimize, there will be wider room for applications of deep learning in the areas of insurance fraud detection, credit security and anti-e-connoisseur campaigns. Through the offline and online combination, cold boot and man-machine closed loop, and the "sandwich" model, ZhongAn Technology designed a highly efficient anti-insurance fraud framework to provide robust support for business operations, as shown in Figure 6.

Comprehensive Technical Support from Software to Hardware

Besides the unique innovation in algorithm design, the underlying hardware infrastructure of the GBDT→GRU→RF sandwich-structured fraud detection model should also be credited for the model's success in its powerful performance support. Intel provides the model with its high-performance Intel® Xeon® processor family, which not only features excellent core and cache performance but also utilizes massive hardware-enhanced technologies to strengthen the performance of the framework. These include the Intel® Advanced Vector Extensions 2 (Intel® AVX2) and Intel® QuickPath Interconnect (Intel® QPI).

Meanwhile, the newer generation of Intel® Xeon® Scalable processors can bolster the performance of the framework better than their predecessors and integrate the Intel® AVX-512 technology. Its outstanding parallel computing ability is in line with the requirements of AI and can play an important role in the operation process of the framework.

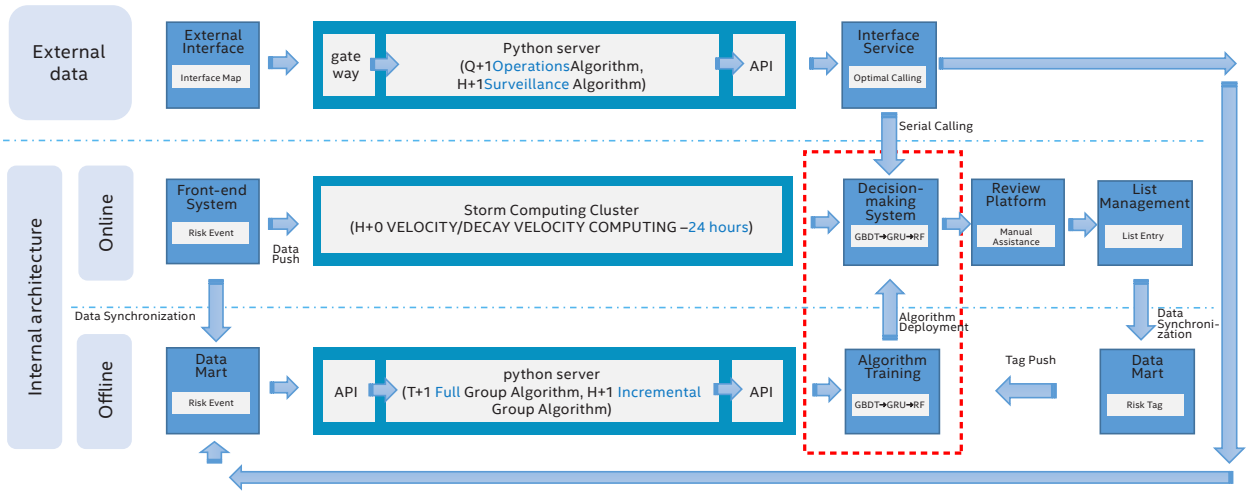

Figure 6 Anti-insurance fraud framework based on "sandwich" structure

Besides the processor products, Intel also offers effective and all-encompassing optimization methods and tools for the GBDT, GRU and RF methods in the model. First of all, for the GBDT method, Intel provides an Apache Spark* computing cluster-based open source deep learning library, BigDL, which allows users to develop their own deep learning application as a standard Spark program. This brings about greater consistency and efficiency for the users. Secondly, in the GRU method, the new model uses the Intel® Optimization for TensorFlow*. Intel provides a variety of effective optimization methods for TensorFlow, such as Intel® MKL-DNN, and with the introduction of the TensorFlow code, the user can make full use of the scalability of the Intel® Architecture processor to reduce the system's overheads brought about by the data format conversion to optimize system load. Finally, in the phase of the RF method application, the Intel® Distribution for Python has a built-in Intel® Data Analytics Acceleration Library (Intel® DAAL), which can provide users with building modules for tasks, including data preprocessing, conversion, modeling and prediction, to effectively improve the working efficiency of the entire model.

So far, the new sandwich-structured fraud detection model has performed up to expectations in various assessments by the National Engineering Laboratory for E-commerce and E-payment and ZhongAn Technology. The R&D and construction of new models jointly carried out by the National Engineering Laboratory for E-commerce and E-payment, ZhongAn Technology and Intel provide useful experience and exploration for AI in applications in the field of anti-financial fraud. This smooths the path for the applications of the various new technologies and algorithms of deep learning in the financial context. In the future, the three parties will continue their technical cooperation, introduce more advanced technologies and products, and accelerate their research in financial fraud prevention to nip financial risks in the bud.

Experience:

The learning ability of traditional machine learning-based anti-financial fraud model for serialized transaction features is inadequate, while the single-method deep learning model has limited learning ability for features of a single transaction. By using a multilayered deep learning model, the inadequacies can be optimally avoided, which enhances the working efficiency and performance of the anti-fraud model.

Not only does Intel power the new anti-fraud model with its high-performance processor, it also provides diversified and comprehensive technical support. For the methods used in each layer of the sandwich-structured fraud detection model, it provides specific optimization means and tools to help the entire anti-fraud model achieve higher working efficiency.

[1] Source: Nielsen Report on Global Bank Card

[2] Source: https://nilsonreport.com/upload/content_promo/The_Nilson_Report_Issue_1118.pdf

0418/SHA/PMG/XX/PDF 336977-001US