Deep Science Publishing
https://doi.org/10.70593/978-93-49910-16-4_6

DeepScience
Open Access Books

Chapter 6: Credit monitoring and proactive risk management with deep learning models

6.1. Introduction

Credit risk monitoring is a long-overdue endeavor in the field of credit management, and this research sheds new light on how to proactively assess risk using deep learning methods on irregular time series. The greatest challenge is data irregularity, in which the timestamp of an event is often missing and its arrival can be delayed for a long time, which can result in silence intervals. Unlike methods developed in the financial domain, the shares of data discontinuity are studied in the field of event sequences to design a new simulation method of irregular time series data. A new architecture is also designed to address the challenge of irregular time series, incorporating specific embeddings, graph attention networks, and several attention modules based on multi-domain event information. This research ends with discussions on applications to other fields, the interpretability of the model's decisions, and risk monitoring from a financial-metric-centered perspective (Chen, 2019; Elngar et al., 2022; Javaid et al., 2022).

The proper prediction of credit risk is crucial to individual consumers, firms, and industries as a whole, especially with the development of peer-to-peer lending platforms and other new types of credit consumption. Overall, P2P is necessary and promising in China, which is supported by both individuals' credit scores and social networks. Additionally, since P2P companies lack supervised information, there's potential to utilize unstructured information in the financial arena or account social networks learned from overseers. P2P companies need new analytical and managerial capabilities regarding honoring anti-money laundering regulations, law enforcement sharing standards, and investigating high-risk customers with comprehensive territory analysis. Finally, consumers need to simplify their credit management processes, expectations of

loan owing, and repayment terms.Credit risk prediction involves predicting the probability of default for a certain applicant on an ad-hoc loan. Either a deterministic number of seconds or an ordered series of seconds can be treated as a time series. Time series data from real-world FinTech companies can either be transformed to a non-sequential format after data preprocessing or work with the block box of models without clear causal relationships. During the past decade, multi-stage deep learning models with different mechanisms, illustrations and neural architectures including recurrent neural networks, long short-term memory, attention network and transformers have been devised to deal with the sequential and temporal nature of the data. Nonetheless, the overwhelming majority of existing works rely on widely adopted time series databases or simulated sequential datasets. Our work differed fundamentally from these existing works in financial credit risk prediction with sequential features in three aspects: unique modeling of features as linguistic text; multi-head design of dependency models (Nauta et al., 2019; Shrier & Pentland, 2022).

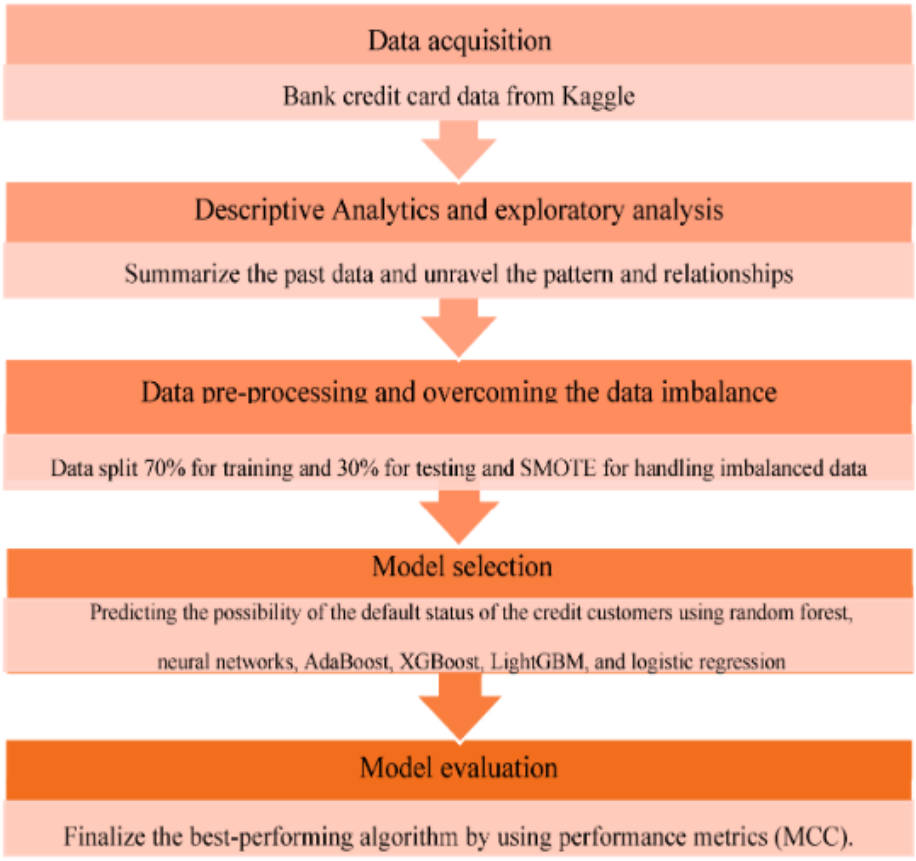

Fig 6.1: Credit Risk Prediction Using Deep Learning

6.2. Background

Mackenzie summarized the evolution of credit scoring modeling in four stages: heuristic, rule-based, statistical, and scoring-based. In the heuristic stage (early 1900s), the scoring system was designed manually. Various rules based on the experts' knowledge and experiences were applied to segment good and bad borrowers. In the rule-based stage (1950s), the experts' rules worked as the basis to build automation models. The models were called expert systems. In the statistical stage (1980s), fast growing statistics and data mining learners were utilized to build credit scoring models. Regression, Discrete Choice and Tree-based models became popular. However, all these models could not stand nonlinear and complex data yet. In the scoring-based stage (2000s-present), emerging and advanced machine learning models attempted to construct sophisticated scoring models. Recently, Deep Learning models have initiated a premium evolution different from previous machine learning approaches. Deep Learning accounts for multiple transformations modeling, which can automatically learn feature representation hierarchy without domain knowledge. Considering the automatic feature extraction ability, Deep Learning models are very likely to outperform previous machine learning techniques in credit risk modeling when sufficient data is available. However, in practice when analyzing financial data, obtaining labeling (status in default, no default) information for deep learning directly is usually infeasible. Probabilistic Survival Analysis is a good approach to partially remedy this issue. It models the occurrence time of events with the use of survival distribution that provides both likelihood and timing information.

6.2.1. Overview of Credit Monitoring

Credit monitoring is defined as the continual evaluation of the potential risk associated with clients. It includes not only those who have already received a credit but also those looking to take out a loan at some point in the future. Many credit card services can, for example, approve or deny credit almost instantaneously. Credit monitoring is designed to predict when risk levels set precedent will be violated, indicating that the current assessment of the risk associated with a client is outdated. The ability to do this in a timely manner can be seen as crucial to any lending institution. For instance, early detection of a borrower who has worsened economically (s)he can lead to prevention of defaults, i.e., encourage the user to pay back early. On the other side, an institution’s ability to discover an evolution toward lower risk can be critical to offer better terms to clients.

The objective of credit monitoring models is to give a score per borrower at regular intervals in order to make predictions of the future two values of the client-risk probability terms. A number of general comments are made about the desirable features

of the credit monitoring module. A score produced by a credit monitoring model is generally not actionable on its own. It is typically proposed to compare the scores produced by the monitoring model to a threshold. This threshold is, for example, two-fold: to assign clients to a 'risk-not' or 'risk' set. In this scope, the concept of 'failure' includes both defaults and a set of states associated with warning signs that need scrutiny. Model calibration is an important area of inquiry when addressing the concept of credit monitoring.

6.3. Deep Learning Models in Finance

Credit risk prediction has long been a standard topic in the field of finance and is also of great importance in real-world applications including credit issuance. There exist extensive research efforts in both academia and industry to address this problem. Traditionally, credit risk data is managed in a structured way, and the features are mainly numeric, categorical or binary. Various statistical or machine learning methods could be utilized including logistic regression, support vector machines, random forests and gradient boosting decision trees. However, recent years have witnessed the booming of financial tech companies which consider different type of data including unstructured data such as textual data, image data or audio data. Credit risk prediction has evolved in a more complex fashion to deal with these new features as well as new models. In addition to tree-based models, deep learning techniques inspired by biological neural networks allow the automatic hierarchical feature abstraction through establishing and training multi-layer neural networks and have found their way into managing and providing financial services.

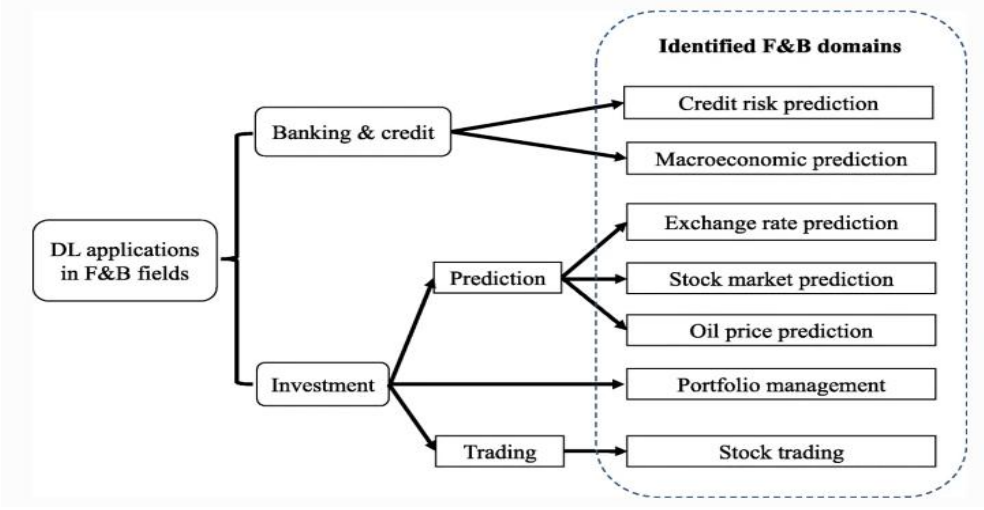

Fig 6.2: Deep learning in finance

6.3.1. Introduction to Deep Learning

Deep learning (DL) has been attracting much attention recently due to its good performance in many real-world applications, including computer vision, natural language processing, and finance. With the booming growth of information and data, many financial institutions have invested lots of resources in exploring online information in the hope of identifying trading opportunities and helping to improve the decision-making process. For example, retail banks are eager to adopt the so-called "big data" analysis strategies with the aim of automating the trading processes and enriching the customer relationship management systems. Despite the recent trend in using DL models, traditional models based on statistical learning techniques remain the most effective techniques on hand for understanding consumer behavior in the finance industry. Compared with the massive amounts of information and data on consumer credit provided by many different E-commerce platforms, online stock trading is yet to be fully explored due to the unavailability of such data.

Lending money is risky for financial institutions because there is a chance that a borrower will fail to pay back. Here, credit risk is the risk of loan default or loan delinquency when a borrower fails to repay on time. For financial institutions, a default means financial loss and capital chain break. Therefore, credit risk prediction is an analytical problem of utmost significance in the fields of credit risk management and quantitative finance for banks and other financial institutions, especially for the consumer credit increasing at an explosive pace in recent years. With credit risk prediction techniques, financial institutions can automatically evaluate the credit risk level of an applicant based on the applicant's information and facilitate decision-making by follow-up credit policies. For example, the predicted applicant is classified as healthy or bad regarding credit risk and different credit limits can be assigned accordingly. Accurate credit risk prediction is vitally important for financial institutions when they are formulating lending strategies for loan applications. The main purpose is to keep the bad debts at a low level. In this way, the direct substantial loss of the multi-billion dollar credit loan industry can be saved. Better credit risk prediction also improves the risk management capacity of banks and fintech companies.

6.4. Data Sources for Credit Monitoring

The assessment of credit quality is based on the consideration of ten fundamental reference attributes evaluated in credit risk modeling systems (CRMS). The most important attributes of creditworthiness in short-term consumer credit monitoring are income, loan, probability of default (PD), observation date, amount of loan defaulted, first observation date, installment payment, count of installment payment, total no of installment periods, and quick assessment guarantee funds. Credit scoring models are

designed to estimate the probability of default based on various observed customer attributes. The consistency of the assessment of monitored data must be maintained across all observation time intervals. There are two hypotheses – priori and posteriori regarding how early to start monitoring after a loan is issued. The extensive consideration of customer credit indicators is essential for observing credit score changes. Inconsistency is estimated and evaluated by monitoring customer credit attributes over time.

As mentioned earlier, CRMS needs to be further developed with the consumption of new data sources. The PD should be initially estimated using economic attributes (EA), meanwhile the monitoring of credit scoring should not increase the complexity of the adapted model. Even though the simple logistic regression could still produce good performance, the predictive accuracy for rare events might be overwhelmingly affected by non-monitored attributes. For one of the EAs – inflation rate, a random walk model could be used for a better adaptation to new observations. The score adaptation should be reassessed after changes of more recent information, in-order to verify whether and how credit history changes its impact on credit scoring. Furthermore, it has been proven that updates of naive risk factors with changes in more recent data positively affect risk dynamics projections. The tune threshold could be indirectly learnt from odds score updates and used for cohort level tests on the level of currency of predictions. Additional examples from other domains such as systematic approach for cascade architecture to capture the spatial correlations between pixels provide insights for risk modeling architectures.

6.4.1. Financial Data

Financial data is of great importance in the financial industry, which is composed of a wide variety of information. Depending on the business scenario and volume, basic user information (e.g., gender, age, job), three-party information (e.g., annual income, credit history, total credit), internal behavioral buried information of the product (e.g., number of times to borrow, number of loans), etc., can be obtained. These information contain time-series features, numerical features, category features, etc. A basic loan data, which is a typical example of credit risk prediction, is provided to illustrate the data scenario. The developed production model can be directly used for credit risk control on basic loan data. In this case study, they present the prevalent case for credit risk prediction.

Due to the fierce competition in the finance industry, effective risk control, credit risk planning, and compliance analysis are of great importance. These tasks involve comprehensively analyzing user data, transaction data, marketing strategy data, third-party data, and other features of various data types and dimensions, which contain important signals for proactive risk management. Credit risk prediction is one of the

essential abilities for financial institutions to better formulate credit policies and risk control measures on user applications. Institutions with investment in deeper modeling methods can better catch the latest risk trends and avoid investment losses. Generally, credit risk prediction can be posed as a binary classification problem, where the good users and bad users are classified based on their features. On one hand, better credit risk prediction indicates a lower portion of bad debt, leading to more stable revenue in lending policies. On the other hand, better credit risk prediction can improve the risk management capacity in loan issuing, collecting, and other approaches.

6.5. Methodology

This work concerns a research problem with significant application impact in the industry, involving the detection of bad credit loans with substantial financial loss. The goal is to determine whether the applicant would be in default on the loan repayment from a financial perspective by building a scoring model based on the applicant's loan history and demographic information. The work assumes a cross-sectional dataset, a standard setting in credit risk evaluation, and works for binary classification that outperforms a production scoring model.

This work represents an industrial case study about credit risk prediction considering one cross-sectional dataset. Deep learning models have achieved superior performance in many fields, but are not widely adopted in the financial industry. Financial data are often of much lower quality than those of other industries. Financial data is of high dimensionality, with a relatively small sample size. Moreover, there are many extreme values, abused features, and missing entries. Thus, as a black box and end-to-end solution, deep learning models are often regarded as a dangerous approach by prominent industry leaders due to their poor interpretability and over-fitting risk. Non-DE models, such as XGBoost and tree models, are still the best and most commonly taken solution for credit risk evaluation.

This work uses a real-world financial dataset that is common practice in the financial industry to tackle these problems, since it consists of many more lost entries and extreme value features, as well as much sparser and higher dimensional layout than e-commerce data. Moreover, it provides a successful deep learning case study of broader interest, along with observation and suggestions for its further adoptions. This work proposes to monitor a potential re-default for an already-allocated loan using the information from new candidate customers and learn the function similar to a credit scoring model, which has not been addressed in prior work.

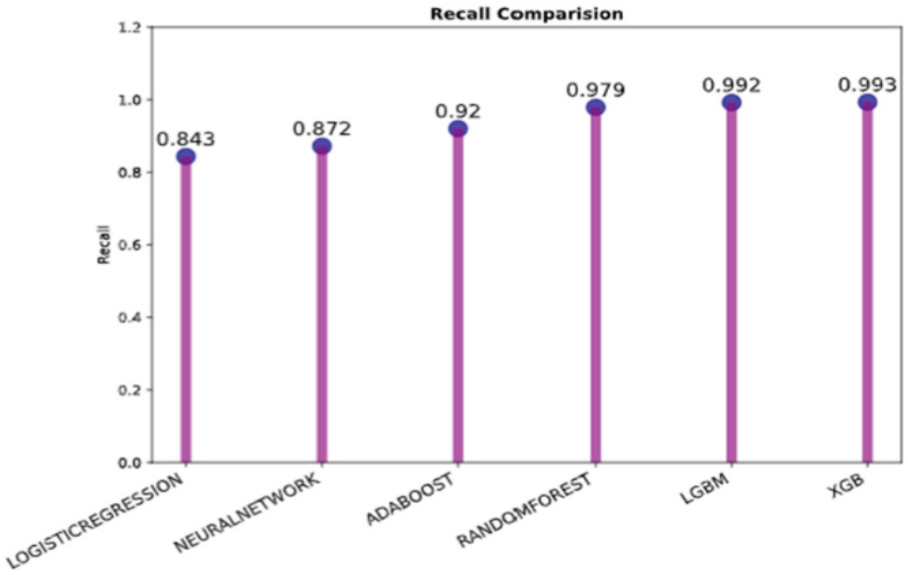

Fig : Credit Risk Prediction Using Machine Learning and Deep Learning

6.5.1. Data Preprocessing

A multi-stage careful data processing procedure is applied to turn noisy and irregular input features from a real-world financial institution into a neatly structured format. The noisy non-sequential real-value features from internal logs are extensive and variable. As a result, only a small amount of features are neat and useful, while a huge proportion of them are useless noise and problematic for training deep networks. There are two main types of records in the data set: application records, which will be processed before the test time, and daily logs, which include transaction records and will be updated every day. For each applicant, a snapshot of fixed detailed attributes is collected (basically non-sequential features), while all application records and logs are sequential time series or even irregular recordings. For the static features, all the fixed attribute features from the first time are collected as non-sequential features. It is natural to directly select and discard useless features. However, for the sequential features, although there are many methods proposed to process sequential data, they usually assume a neatly structured input but do not involve any data processing.

To build the year-based data set, firstly a time window to select the snapshot day has to be determined. The snapshot date is required to be chosen such that there are enough numbers of applications before this day. There are two principles to take into consideration: a window with a pretty long size would lose time relevance since early applications may be less informative for predicting current credit risk; while on the contrary, a window that is too small would make the training process easy but may suffer

from a relatively weak generalization ability. All training samples collected from the application records before the snapshot date filter by the application day monotonically forward.

Then, as for the non-sequential features, the training samples and output labels are extracted from the filtered data. There are two main tasks in non-sequential data preprocessing to obtain high-quality features: selecting effective features and dealing with problematic values. Proper data pre-preprocessing can be significantly beneficial for the optimization of DL models. The problematic values of a feature are denoted as 0 value or missing values, which will be considered accordingly. All the non-sequential numerical features have a strikingly huge proportion of 0 values. One reason is that many applicants may not have any transaction records seized by the logs, leading to a large number of 0s that are systematic and meaningless labels in threat segmentation.

6.6. Conclusion

Due to the increasing demand for credit by individuals and businesses, lenders are facing increasing challenges in finding ways to ensure credibility in credit risk management. Involving hundreds of millions of loan records along with complex raw information, the large-scale, high-dimensional, heterogeneous, and noisy financial data presents additional challenges for serving as critical reference information in credit risk assessment. Conventional risk management models based on statistical methods, while mature in academics, are difficult to use in practice due to their simplified assumptions on data distributions, which could lead to serious lending failures and business losses. Recently, with the rapid deep learning development over the last decade, there are growing interests in evolving financial tasks with complex and large-scale financial data from statistical to deep learning-based modeling due to the better capacity of deep learning models in modeling the coarse and sufficient representation of complicated data distributions.

In this study, an effective deep learning-based framework for credit risk assessment with complex and noisy financial data is presented. The presented framework integrates a wide collection of costs, data types, and architectures of deep learning. With successfully evaluating a bank's existing systems with large-scale testing over random projections of millions of loan records, a better joint risk quantile model for a new credit score business against the existing decision-tree model is developed. With joint learning of multi-stage financial data pre-processing and model architecture fine-tuning, a large proportion of effectively selected inputs were demonstrated to significantly improve the prediction accuracy through transferring low-dimensional features for accelerating the later training stages. After its joint implementation in practice, the testing results over the similar datasets collected during one year currently indicate that this new joint DSPEC

framework can largely guarantee the estimated quality and be actively used for proactive risk monitoring without re-selection of features and model architectures.

References

Elngar, A., Kayed, M., & Abo Emira, H. H. (2022). The Role of Blockchain in Financial Applications. In Artificial Intelligence and Big Data for Financial Risk Management (pp. 140–159). https://doi.org/10.4324/9781003144410-9IJISAE

Chen, E. (2019). Implementing Blockchain Technology in the Financial Services Industry. In Essentials of Blockchain Technology (pp. 257–272). https://doi.org/10.1201/9780429674457-12IJISAE

Shrier, D. L., & Pentland, A. (2022). Fintech Foundations: Convergence, Blockchain, Big Data, and AI. In Global Fintech (pp. 7–32). https://doi.org/10.7551/mitpress/13673.003.0004 IJISAE

Javaid, M., Haleem, A., Singh, R. P., Suman, R., & Khan, S. (2022). A Review of Blockchain Technology Applications for Financial Services. BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 2(3), 100073. https://doi.org/10.1016/j.tbench.2022.100073 ResearchGate

Nauta, M., Bucur, D., & Seifert, C. (2019). Causal Discovery with Attention-Based Convolutional Neural Networks. Machine Learning and Knowledge Extraction, 1(1), 312–340. https://doi.org/10.3390/make1010019