██████████-driven credit risk: a systemic review

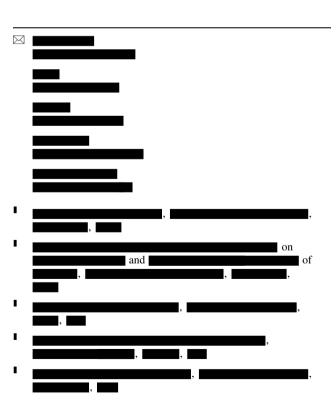████[1] · Rita Tse[1,2] · Wuman Luo[1] · Stefano D'Addona[1] · ████████████[1,2]

Abstract

Credit risk assessment is at the core of modern economies. Traditionally, it is measured by statistical methods and manual auditing. Recent advances in financial ███████████ stemmed from a new wave of ██████████ (ML)-driven credit risk models that █ined tremendous attention from both industry and academia. In this paper, we systematically review a series of major research contributions (██ papers) over the past eight years using statistical, ██████████ and ████ learning techniques to address the problems of credit risk. ██████████, we propose a novel classification methodology for ML-driven credit risk algorithms and their performance ranking using public datasets. We further discuss the challenges including data imbalance, dataset inconsistency, model transparency, and inadequate utilization of ███ learning models. The results of our review show that: █) most ████ learning models outperform classic ██████████ and statistical algorithms in credit risk estimation, and █) ensemble methods provide higher accuracy compared with single models. Finally, we present summary tables in terms of datasets and proposed models.

✉ ██████████
██████████████
████
██████████████
████
█████████████
██████████
██████████████
██████████████

▪ ██████████████, ██████████████,
██████, ███

▪ ██████████████████████ on
██████ and ████████████ of
██████, ██████████████, ██████,
███

▪ ██████████████████, ██████████,
████, ██

▪ ██████████████████████,
██████████, ████, ██

▪ ██████████████████, ██████████,
██████, ██

█ Introduction

██████████████ advances heavily affected industry and academia in the past decades, ultimately transforming people's daily life. ██████████████ (AI) has been applied to almost every human activity, including pattern recognition, image classification, business, agriculture, transportation, and finance. This paper focuses on ██████ ███████ applied to finance and credit risk estimation. Modern financial systems rely on credit and trust. Credit risk is a fundamental parameter that measures and predicts the default probabilities of a debtor. The correct estimation of credit risk is paramount for the entire system. Failing in the credit risk estimation can ████ to systemic failures such as the sub-prime crisis of ████. Consequently, lenders devote large amounts of resources to predict the credit-worthiness of consumers and companies to develop appropriate lending strategies that minimize their risks. Historically, credit risk approaches use statistical methods such as Linear Discriminant Analysis [█] and Logistic Regression [█]. These methods, however, do not easily handle large datasets.

Advances in computing power and availability of large credit datasets paved the way to AI-Driven credit risk estimation algorithms such as traditional ██████████ and ████ learning. Conventional ██████ ███████

techniques, e.g., k-Nearest Neighbor [█], ████████████ [█] and ████████████████████ [█], are more effective and flexible than statistical methods. In particular, the vital branch of ████████████-███ learning techniques [█] applied to large credit risk data lake outperform their predecessors both in accuracy and efficiency.

This paper presents a systemic review of credit risk estimation algorithms. It analyzes both the major statistical approaches and AI-based techniques with a critical spirit. The aim is to provide a comprehensive overview of the current ████ing credit risk estimation technology, providing justification and connections between past and present works. This work proposes a novel taxonomy combining finance with ████████████ techniques. In addition, this work ranks their performance in terms of accuracy and costs. This paper also discusses the challenges and possible solutions in terms of four aspects: data imbalance, dataset inconsistency, model transparency, and inadequate utilization of ████ learning methods.

The remainder of the paper is or█nized as follows: the survey methodology will be discussed in ███. █. ███ion █ introduces the principles of statistical learning, ████████ ████████ and ███ learning. ███ion █ analyzes credit risk-related applications in detail. In ███. █, presented algorithms are discussed and ranked by their performance a█inst public datasets. Finally, results and current challenges are summarized in ███. █; while ███. █ concludes this work.

█ Survey methodology

█.█ Methodology

We applied ████████ (Preferred Reporting Items for Systematic Reviews and Meta-Analyses ███. █) reviewing methodology in our paper. First, we adopted ███ searching platforms for our investi█tion: ████████████, ████, ████████, ████████, and ████████████. We used the keywords "████████████" or "███ learning" combined with "credit risk" while searching. We got ███ articles in total. Then, we applied a filtering algorithm considering the trade-off between publication year and citations to proceed. After removing ████ duplicate records, ███ ineligible records, and ██ incomplete articles, we obtained ██ screened records. Based on the relevance, we excluded ██ articles less related to the topic. After manually checking whether the paper has clear evaluation metrics, we further excluded another ██ papers. Finally, we kept ██ studies in terms of the relevancy to the research topic, precision of evaluation metrics, publication time, and number of citations as our source of reviewing.

█ure█ depicts the ████████ flow diagram

████████

█.█ Inclusion and exclusion criteria

In this paper, we select three inclusion criteria: (█) the relevance of research topic, (█) the precision of evaluation metrics, (█) the publication year and citations. Moreover, the papers will be excluded if they are duplicated, incomplete, too early, low-related with the topic, having no clear metrics or comparatively low citations.

We show the whole workflow of the selection process in ██. █.

█.█ The datasets and approaches of the reviewed articles

The mainly used datasets by the papers under review are German and Australian public credit data from the ███ ████████████ repository [█–█]. In addition, there exist some researches that discover and mine their own data. For example, ████████████ (████) uses data from a firm in ████████ [█]. Authors in [█–█] all employ their unique dataset. Those articles mainly emphasize the significance and the veracity of the original data.

We discuss the principles and application of the overall ████████████ approaches. The traditional ████████ ████████ models for credit risk contain ████████████ ████████ (SVMs) [█], k-Nearest Neighbor (k-NN) [█], ████████████ (RFs) [█], Decision Trees (███) [█–█], AdaBoost [█], Extreme Gradient Boost (XGBoost) [█], Stochastic Gradient Boosting (████) [█], Bagging [█], ████████████████ (████) [█] and ██ (Genetic Algorithm) [█]. Neural network models generally belong to ███ learning methods. Most of them include Convolutional ████████████ (CNNs) [█], ████ Belief ████ ████████ (DBNs) [█], ████████████████████ (████) [█], LSTM (Long Short-Term Memory) [█], Restricted Boltzmann Machines (RBMs) [█], ████████████ Perceptron (DMLP) [█], and ████████████████ (RNNs) [█].

Summary tables and bar charts re█rding all the methods of the reviewed papers are provided.

█.█ Taxonomy

The taxonomy is shown as ██. █. We can divide it into two parts: the first is re█rding computing technology and the second is credit risk application domain. The two parts are further categorized into sub████ions. These two parts are connected and fused with each other. All the right-side subdomains include the left-side techniques, and all the techniques can be applied in the financial domains.