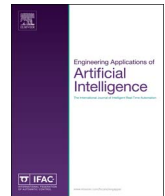




Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

A deep learning approach for credit scoring using credit default swaps

Cuicui Luo*, Desheng Wu, Dexiang Wu

Stockholm Business School, Stockholm University, Stockholm, Sweden

ARTICLE INFO

Keywords:

Deep learning
CDS
Credit scoring
Machine learning

ABSTRACT

After 2007–2008 crisis, it is clear that corporate credit scoring is becoming a key role in credit risk management. In this paper, we investigate the performances of credit scoring models applied to CDS data sets. The classification performance of deep learning algorithm such as deep belief networks with Restricted Boltzmann Machines are evaluated and compared with some popular credit scoring models such as logistic regression, multi-layer perceptron and support vector machine. The performance is assessed using the classification accuracy and the area under the receiver operating characteristic curve. It is found that DBN yields the best performance.

1. Introduction

During US subprime mortgage crisis and the European sovereign debt crisis, established financial institutions in the USA and Europe suffered catastrophic losses. The crisis has raised concerns regarding the use of credit default swaps (CDS). Accordingly, credit risk management is becoming an increasingly important factor and attracted significant attention from researchers and market participants. In order to effectively manage the credit risk exposures and optimize profits, it has become one major focus for financial institutions to develop an accurate credit scoring model. A range of different statistical and machine learning techniques have been developed to build credit rating models.

After commercial scorecard was introduced, many statistical methods have been used for credit risk assessment. Though with wide application, these models have difficulty in modeling complex financial systems due to the use of fixed functions and statistical assumptions. Related studies have shown that machine learning techniques are superior to that of statistical techniques in dealing with credit scoring problems (Saber *et al.*, 2013). Logistic regression is one of the most frequently used statistical model used in credit scoring. Meanwhile, some shallow architectures such as support vector machines (SVMs) and multi-layer perceptron (MLPs) with a single hidden layer, have been widely applied to credit scoring.

Shallow architectures have been shown effective in solving many simple or well-constrained problems. However, these methods mainly focus on the outputs of classifiers at the abstract level, while neglecting the rich information hidden in the confidence degree (Hinton and Salakhutdinov, 2006). Their limited modeling and representational power can cause difficulties when dealing with more complicated real-

world applications. To tackle these drawbacks, Hinton and Salakhutdinov (2006) first successfully introduced training algorithms for deep architectures. The deep belief networks (DBN) with sufficient hidden layers are developed as a powerful ensemble technique to capture the rich information in the confidence degree. Since then, deep networks have been applied with success in classification task, e.g., computer vision (Krizhevsky *et al.*, 2012), health state classification (Tamilselvan and Wang, 2013; Abdel-Zaher and Eldeib, 2016), speech and language processing (Mohamed *et al.*, 2012; Ling *et al.*, 2013) and emotion recognition (Le and Provost, 2013). Recently, the deep belief networks (DBN) has also been applied in financial prediction (Ribeiro and Lopes, 2011) Whether such deep architectures have theoretical advantages compared to shallow architectures in credit risk assessment remains an open question. To the best of our knowledge, there were few studies on credit risk assessment by using the DBN.

Many empirical studies have investigated the performance of these credit scoring models in credit risk assessment. Bellotti and Crook (2009) test support vector machines against several other well-known algorithms on a large credit card database and find that SVMs are successful in comparison to established approaches to classifying credit card customers who default. Li *et al.* (2006) find SVMs outperform MLP in credit scoring by applying consumer credit data. It is also found that SVMs perform slightly better than LR by applying SVM to a database of applicants for building and loan credit (Schebesch and Stecking, 2005). According to Bellotti *et al.* (2011), support vector machines can produce notably better predictions of international bank ratings than the standard method. Han *et al.* (2013) find orthogonal support vector machine achieves better performance in achieves better performance. However, another study conducted by Van Gestel *et al.* (2006) finds no significant difference between SVM, LR in terms of

* Corresponding author.

<http://dx.doi.org/10.1016/j.engappai.2016.12.002>

Received 19 September 2016; Received in revised form 20 October 2016; Accepted 1 December 2016
0952-1976/ © 2016 Elsevier Ltd. All rights reserved.

proportion of test cases correctly classified. The performance of four learning algorithms for corporate credit ratings are compared over a data set consisting of real financial data (Zhong et al., 2014). Lessmann et al. (2015) compare 41 classifiers in terms of six performance measures across eight real-world credit scoring data sets. They suggest that several classifiers predict credit risk significantly more accurately than the industry standard LR across eight real-world credit scoring data sets. The credit scoring accuracy of five neural network models for both the German and Australian credit data sets is investigated by West (2000). He finds that LG is the most accurate of the traditional methods and also suggests that neural network credit scoring models can achieve fractional improvements in credit scoring accuracy. Bhattacharyya et al. (2011) evaluate the performance of two advanced data mining techniques, random forests, support vector machines and logistic regression, for credit card fraud detection. Despite a bewildering array of models, relatively little research compares the performance of these models with DBN in terms of their classification accuracy in credit scoring. Meanwhile, the above studies only consider two-class classification problems.

To our knowledge, this is the first comprehensive study of DBN model in corporate credit rating based on CDS data. Therefore, this paper fills in such a literature gap by introducing DBN as the algorithm for credit rating to generate fast and accurate individual classification results. The goal of the paper is to provide a set of descriptive results and tests that lay a foundation for future theoretical and empirical work on DBN in credit scoring in CDS markets. In this paper, we investigate the performances of different credit scoring models by conducting experiments on a collection of CDS data. The data set contains 661 companies with eleven input attributes and three classification categories. The 6-month, 1-year, 2-year, 3-year, 4-year, 5-year, 7-year and 10-year spreads, recovery rate, sector and region will be used as input variables for the learning algorithms. The output variable contains three rating categories: A, B and C. In our experiments, we compared the results of MLR, MLP, and SVM with the Deep Belief Networks (DBN) with the Restricted Boltzmann Machine by applying 10-fold cross-validation. Our findings demonstrate that the deep learning algorithm significantly outperforms the baselines. Our paper contributes to this literature by investigating the performance of DBN in corporate credit scoring.

The remainder of the paper is organized as follows. Section 2 describes the credit scoring models examined in the paper. Section 3 describes CDS data set. Section 4 presents the empirical results from comparing the models. Section 5 summarizes the paper and makes concluding remarks.

2. Models

In this section, we present four popular machine learning algorithm used for credit scoring.

2.1. Multinomial logistic regression

Logistic regression is one of the most frequently used statistical model in credit scoring. The basic setup of Multinomial Logistic Regression (MLR) is the same as in logistic regression and the only difference is that the dependent variables are categorical rather than binary.

Suppose we have a set of training data with n observations $\{x_i, y_i\}$, where $x_i \in \mathbf{R}^m$ and $y_i \in \{1, \dots, K\}$ for $i = 1, \dots, n$. The probability that the i -th observation belongs to the j -th category with the exception of the last class is

$$\pi_{ij} = P(y_i = j) = \frac{e^{\beta_j \cdot x_i}}{1 + \sum_{k=1}^{K-1} e^{\beta_k \cdot x_i}} \quad (1)$$

Because the sum of all the probabilities equals one, the last class has

probability

$$\pi_{iK} = 1 - \sum_{j=1}^{K-1} \pi_{ij} = \frac{1}{1 + \sum_{k=1}^{K-1} e^{\beta_k \cdot x_i}} \quad (2)$$

where β_j is the set of regression coefficients associated with outcome j and β is an $m \times (K-1)$ matrix.

The negative multinomial log-likelihood is:

$$L = - \sum_{i=1}^n \left(\sum_{j=1}^{K-1} Y_{ij} \cdot \ln \pi_{ij} + (1 - \sum_{j=1}^{K-1} Y_{ij}) \cdot \ln \left(1 - \sum_{j=1}^{K-1} \pi_{ij} \right) \right) \quad (3)$$

where Y_{ij} denotes the output variable that the i -th observation belongs to the j -th category.

The above log-likelihood function is solved by a Quasi-Newton maximization Method which is described in details by Weka (Hall et al., 2009). A ridge estimator is used in order to prevent over-fitting by penalizing large coefficients (Le Cessie and Van Houwelingen, 1992).

2.2. Support vector machine

Support Vector Machine (SVM) was first introduced in 1992 by (Boser et al., 1992). It is a classification and regression tool that applies machine learning technique to maximize predictive accuracy while automatically avoiding over-fit to the data. It can learn both simple and highly complex classification models and employs sophisticated mathematical principles to avoid over-fitting. We describe SVMs for two-class classification here and it can be extended to K class classification easily by constructing K two-class classifiers (Boser et al., 1992).

Given a set of training data $\{v_i, y_i\}$, where $v_i \in \mathbf{R}^m$, and $y_i \in \{-1, 1\}$ for $i = 1, \dots, n$. The Convex quadratic programming problem (QP) for SVM classification is

$$\text{Maximize} \quad \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(v_i, v_j) \quad (4)$$

with constraints

$$0 \leq \alpha_i \leq C \quad (5)$$

$$\sum_{i=1}^n y_i \alpha_i = 0 \quad (6)$$

where C is an upper bound and α_i is a Lagrange multiplier for each sample i . $K(v_i, v_j)$ denotes the value of the SVM kernel function for i -th and j -th inputs.

In this paper, we consider Gaussian radial basis function (RKF) with parameter σ , which is defined by

$$K(v_i, v_j) = e^{-\frac{\|v_i - v_j\|^2}{2\sigma^2}} \quad (7)$$

For any given input v , the output prediction of a non-linear SVM is explicitly:

$$u = \text{sign} \left(\sum_{i=1}^n \alpha_i y_i K(v_i, v) + b \right) \quad (8)$$

where bias b and vector of α_i are the variables determined by the above QP optimization problem.

The Karush-Kuhn-Tucker (KKT) conditions are necessary and sufficient conditions for an optimal solution of a positive definite QP problem. The KKT conditions for the QP problem (1) are

$$\alpha_i = 0 \implies y_i u_i \geq 1 \quad (9)$$

$$\alpha_i = C \implies y_i u_i \leq 1 \quad (10)$$

$$0 < \alpha_i < C \implies y_i u_i = 1 \quad (11)$$

Download English Version:

<https://daneshyari.com/en/article/4942617>

Download Persian Version:

<https://daneshyari.com/article/4942617>

[Daneshyari.com](https://daneshyari.com)