Credit Risk Theoretical Model on the Base of DCC-GARCH in Time-Varying Parameters Framework

Nikita Moiseev 1, Aleksander Sorokin 1, Natalya Zvezdina 2, Alexey Mikhaylov 3,* 4, Lyubov Khomyakova 4 and Mir Sayed Shah Danish 5

1 Department of Mathematical Methods in Economics, Plekhanov Russian University of Economics, 117997 Moscow, Russia; moiseev.na@rea.ru (N.M.); sorokin.as777@rea.ru (A.S.)
2 Department of Statistics and Data Analysis, Faculty of Economic Sciences, National Research University Higher School of Economics, 101000 Moscow, Russia; zvezdina@hse.ru
3 Financial Research Institute of Ministry of Finance of the Russian Federation, 127006 Moscow, Russia
4 Institute for Research of International Economic Relations, Financial University under the Government of Russian Federation, 124167 Moscow, Russia; aalmakul@ya.ru
5 Strategic Research Projects Center, University of the Ryukyus, Nishihara, Okinawa 903-0213, Japan; mdanish@lab.u-ryukyu.ac.jp
* Correspondence: ayumihajlov@fa.ru

Abstract: The research paper is devoted to developing a mathematical approach for dealing with time-varying parameters in rolling window logit models for credit risk assessment. Forecasting coefficients yields a better model accuracy than a trivial approach of using computed past statistics parameters for the next time period. In this paper, a new method of dealing with time-varying parameters of scoring models is proposed, which is aimed at computing the default probability of a borrower. It was empirically shown that in a continuously changing economic environment factors’ influence on a target variable is also changing. Therefore, forecasting coefficients yields a better financial result than simply applying parameters obtained by accumulated statistics over past time periods. The paper develops a new theoretical approach, incorporating a combination of the ARIMA class model, the DCC-GARCH model and the state–space model, which is more accurate, than using only the ARIMA model. Rigorous simulation testing is provided to confirm the efficiency of the proposed method.

Keywords: default probability; scoring model; logistic regression; time-varying parameters; time series forecasting; ARIMA; DCC-GARCH; Kalman filter; state–space model

1. Introduction

Since the loan lending procedure is one of the primary operational activities for most banks around the world it is necessary to properly classify their borrowers by the risk of default occurrence. If a bank’s scoring system fails to reliably select solvent clients it threatens the bank’s financial stability, damages the banking sector and, as a consequence, can lead to an economic crisis. Thus, a good scoring model can significantly improve a bank’s profit and prevent unnecessary losses. However, to elaborate a reliable and accurate scoring model in the context of a dynamic and volatile economic environment becomes a challenging task. Since the end of the last century, researchers have tested a vast number of methods for predicting the probability of default (PD) by incorporating various factors associated with each particular borrower. A list of developed methods and their extensions includes such popular approaches as: linear discriminant analysis (LDA), logistic regression, support vector machine (SVM), neural networks, naïve Bayes classifier, classification trees, random forest, gradient boosting, ensembles of different models and others. However, it is not so easy to understand which method provides the best model. Conventionally it is performed by the following procedure: obtained statistics are split into
training and testing datasets. Then parameter fitting is performed for the analysed methods on a training dataset and the model’s accuracy on the testing dataset is consequentially verified. Such an approach, nevertheless, has a flaw, since in this case the models are tested on already filtered borrowers, which was carried out by the initial scoring system. Thus, tested methods fail to take into account the so-called “grey zone” (clients, that either did not apply for a loan or were filtered out by the initial scoring system. Therefore, obtained results can be regarded as the efficiency of an additional filter of a scoring system. The downside of this method of upbuilding filters is the significant shrinkage of a pool of borrowers, which jeopardises a bank’s profit. In order to objectively compare scoring models, one would require a non-filtered pool of borrowers to be independently assessed by tested models. Besides that, due to the evolution of the economic environment, the true parameters of any model are subject to change. Some factors become insignificant with time, some—increase the extent of their influence.

The paper fills the gap in methods incorporating adjustments of obtained coefficients, performed by Kalman filter since we get two independent estimates: First—by forecasting a parameters’ vector, second—by fitting parameters to accumulated statistics. Traditionally the procedure of implementing a scoring model into a credit organisation implies fitting parameters to obtained statistics and applying the model to coming borrowers afterward. However, if a true model’s parameters are dynamically changing, such an approach leads to applying already obsolete parameters’ values, causing significant inefficiency of the scoring model. The main significance is in the development of a mathematical approach, that is aimed at accounting for time-varying parameters and, thus, will help improve the quality of the produced scoring model and consequently contribute to increasing the stability of the world financial system.

This paper has the following structure. Section 2 presents a quick literature review on existing scoring techniques. Section 3 presents a mathematical approach to adjusting parameters’ estimates, taking into account the volatile economic environment. In Section 4 an empirical reasoning for the proposed approach is presented. Section 5 is devoted to simulation testing of the proposed method. Section 6 sums up the key points of the paper.

2. Literature Review

Credit scoring is an essential part of a bank’s operational activity and guarantees the stability of a country’s financial system. There are numerous statistical methods used in credit scoring such as linear discriminant analysis (LDA) [1], decision trees [2], Markov chain analysis [3,4], profit analysis [5] and logistic regression [6].

Popular machine learning methods include artificial neural networks [7–9], genetic algorithms [10] and artificial immune systems [11].

Traditionally, LDA and logistic regression are the most widespread scoring techniques, which, as of late, are actively challenged by machine learning methods and models’ ensembles. The earliest application of LDA for credit risk analysis, fraud detection and anti-money laundering system dates back to 1941 [12]. The use of logistic regression models, nonparametric models, such as classification trees, neural networks and k-nearest neighbours was thoroughly investigated [13–15]. At the very beginning of the 21st century, it became very popular to apply machine learning to credit risk analysis. Hurley et al. consider credit risk analysis in the era of Big data [16]. The most prominent papers on decision trees include [17–19], regarding neural networks [20–22].

Ensemble methods have become extremely popular in credit risk analysis. The core of this method is the integration of a set of individual models of different nature to capture the best features of each learner. Already at the end of the 20th century, it was shown that ensemble methods are more efficient than individual learners [23]. Huang et al. demonstrated a significant difference between statistical methods and machine learning approaches and along with Fu et al. and Opitz and Maclin stated that ensemble methods are preferable to individual models in terms of their prediction accuracy [24–26].
However, in scoring literature the topic of time-varying parameters is not highlighted, though it is very popular in time-series forecasting, see for example Bitto and Frühwirth-Schnatter, Chan and Eisenstat, and Kalli and Griffin [27–29]. Since most economic time series are subject to the volatility of parameters, such models usually provide better accuracy. In credit scoring, this approach, as will be shown further in the paper, can also increase the efficiency of previously constructed models [30–32].

3. Materials and Methods

Let \( \{ y_t, X_t : t = 1, \ldots, T \} \) be a considered real-valued sample, where \( y_t = ( y_{t1}, y_{t2}, \ldots, y_{tn} )^T \) is a binary target variable, which equals one in case a loan was defaulted and zero otherwise, and \( X_t = (1, x_{t1}, x_{t2}, \ldots, x_{tn} ) \) is a countable longitudinal dataset of possible explanatory variables, where \( x_{it} = ( x_{1it}, x_{2it}, \ldots, x_{nit} )^T \). Suppose the probability of default can be modelled by logistic function as follows:

\[
p(1|X_t) = \frac{1}{1 + e^{-X_t\beta_t}}
\]

where \( \beta_t \)—time-varying parameters vector.

We can model the dynamics of each \( \beta_{it} \) either by seasonal trend model or by some ARIMA class model, classical representation of which is presented below:

\[
\hat{\beta}_{it} = a_0 + a_1\hat{\beta}_{i(t-1)} + \cdots + a_p\hat{\beta}_{i(t-p)} + \gamma_1\hat{\epsilon}_{i(t-1)} + \cdots + \gamma_q\hat{\epsilon}_{i(t-q)} + \hat{\epsilon}_{it}
\]

where the hat sign corresponds to estimators obtained from the analysed sample.

Thus, it is stated that if forecasted parameters are used instead of obtained ones, see Equation (2), the model will be more efficient.

\[
p(1|X_t) = \frac{1}{1 + e^{-X_t\beta_t}}
\]

Moreover, at each time period two independent estimates of true parameters’ vector \( \beta_t \) are obtained: one from the forecasting Equation (3) and the other one—from model fitting according to Equation (1). Thus, these two estimates by Kalman filter can be combined to obtain a more precise parameter vector. The probability density function for such parameters can be computed as shown below.

\[
f(\beta_t|\hat{\beta}_{t-1}, \ldots, \hat{\beta}_{t-p}, X_t) = \frac{f_1(\beta_t|\hat{\beta}_{t-1}, \ldots, \hat{\beta}_{t-p}) f_2(\beta_t|X_t)}{\int f_1(\beta_t|\hat{\beta}_{t-1}, \ldots, \hat{\beta}_{t-p}) f_2(\beta_t|X_t) d\beta_t}
\]

where:

\[
f_1(\beta_t|\hat{\beta}_{t-1}, \ldots, \hat{\beta}_{t-p}) = (2\pi)^{-\frac{m+1}{2}} \det(\Sigma_t)^{-\frac{1}{2}} e^{-\frac{1}{2}(\beta_t-\hat{\beta}_t)^T \Sigma_t^{-1}(\beta_t-\hat{\beta}_t)},
\]

\[
f_2(\beta_t|X_t) = (2\pi)^{-\frac{m+1}{2}} \det(\Omega_t)^{-\frac{1}{2}} e^{-\frac{1}{2}(\beta_t-\tilde{\beta}_t)^T \Omega_t^{-1}(\beta_t-\tilde{\beta}_t)}
\]

The covariance matrix of parameter vector \( \Omega_t \) from Equation (6) is usually estimated by Fisher information matrix as follows:

\[
\Omega_t = X_t^T \tilde{W} X_t
\]

where:

\[
\tilde{W} = \text{diag}\left( \frac{\sum_{j=0}^{m} \hat{\beta}_{jX_{1j}}}{(1 + \sum_{j=0}^{m} \hat{\beta}_{jX_{1j}})^2}, \ldots, \frac{\sum_{j=0}^{m} \hat{\beta}_{jX_{nj}}}{(1 + \sum_{j=0}^{m} \hat{\beta}_{jX_{nj}})^2} \right)
\]
Covariance matrix of parameter vector $\Sigma_t$ from Equation (5) can be obtained by a slightly modified DCC-GARCH model.

$$\Sigma_t = D_t R_t D_t$$  \hspace{1cm} (9)

$$D_t = \begin{bmatrix} \sqrt{h_{1t}} & 0 & \cdots & 0 \\ 0 & \sqrt{h_{2t}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{h_{mt}} \end{bmatrix}$$ \hspace{1cm} (10)

where each $h_{it}$ is a variance of the numerically generated probability density function ($pdf_{it}(\beta_{it})$), which is the average of $N$ probability density functions for the forecasted value of $\beta_{it}$, obtained from model (2). The reason for numerical generation, in this case, is the fact, that we never have the true values of $\beta_{it}$, but rather their estimates with some uncertainty. That is why we generate results of $\beta_{t-1}, \ldots, \beta_{t-p}$ based on their pdf, obtained at previous steps in order to get $pdf_{it}(\beta_{it})$ as follows:

$$pdf_{it}(\beta_{it}) = \frac{1}{N} \sum_{j=1}^{N} pdf_{ijt}(\beta_{ijt})$$ \hspace{1cm} (11)

Since $R_t$ represents conditional correlation matrix for $\beta_t$ and looks as follows:

$$R_t = \begin{bmatrix} 1 & \rho_{12t} & \cdots & \rho_{1nt} \\ \rho_{21t} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \rho_{nt} & \rho_{n2t} & \cdots & 1 \end{bmatrix}$$ \hspace{1cm} (12)

Then from Equations (10) and (12) each element of $\Sigma_t$ can be presented as below:

$$[\Sigma_t]_{ij} = \sqrt{h_{it}h_{jt}\rho_{ijt}}$$ \hspace{1cm} (13)

The evolution of the correlation matrix $R_t$ can be modelled as follows:

$$R_t = Q_t^{-1} \bar{Q}_t Q_t^{-1}$$ \hspace{1cm} (14)

$$Q_t = (1 - a) \bar{\Theta} + a \Theta_{t-1}$$ \hspace{1cm} (15)

where:

$$\bar{\Theta} = \frac{1}{T} \sum_{t=1}^{T} \Theta_t$$ \hspace{1cm} (16)

$$Q_t^* = \begin{bmatrix} \sqrt{q_{11t}} & 0 & \cdots & 0 \\ 0 & \sqrt{q_{22t}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{q_{mm}} \end{bmatrix}$$ \hspace{1cm} (17)

Here $R_t$ is decomposed into $Q_t^{-1}$ and $Q_t$ to ensure that absolute values of all entries are less or equal to one. In order to estimate parameters of the DCC-GARCH model, the paper uses the Maximum Likelihood Estimation method.

Having obtained the pdf for $\beta_t$, the corrected estimates are computed by maximising the likelihood of Equation (4):

$$\hat{\beta}_t = \arg\max_{\beta_t} f(\beta_t | \tilde{\beta}_{t-1}, \ldots, \tilde{\beta}_{t-m}, X_t)$$ \hspace{1cm} (18)
These corrected estimates and pdf are then used in the forecasting Equation (2) to produce predictions for the next values of true parameters $\beta_{t+1}$. It is worth noticing that if each $\beta_{it}$ is subject to or can be approximated by the normal distribution and $\Sigma_t$, $\Omega_t$ are diagonal, the corrected probability density is also normal and Equation (4) can be explicitly written. In order to show that the study introduces the following notation for the sake of simplifying further calculations: $E(\beta_t|\hat{\beta}_{t-1},\ldots,\hat{\beta}_{t-m}) = \mu_1$, $E(\beta_{it} | X_t) = \mu_2$, $\text{var}(\beta_{it} | \hat{\beta}_{i,t-1},\ldots,\hat{\beta}_{i,t-m}) = \sigma_i^2$, $\text{var}(\beta_{it} | X_t) = \sigma_i^2$. Then pdf for $\beta_{it}$ from Equation (4) can be rewritten as below:

$$pdf(\beta_{it}) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(\beta_{it}-\mu_1)^2}{2\sigma_1^2} - \frac{(\beta_{it}-\mu_2)^2}{2\sigma_2^2}}$$

(19)

In order to show this pdf is normal, it is critical to first consider only the numerator of this fraction. Simplifying this expression will yield the result, shown below:

$$\frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(\beta_{it}-\mu_1)^2}{2\sigma_1^2} - \frac{(\beta_{it}-\mu_2)^2}{2\sigma_2^2}} = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(\beta_{it}-\mu_1)^2}{2\sigma_1^2} \frac{(\beta_{it}-\mu_2)^2}{2\sigma_2^2}}$$

(20)

After integrating the denominator of the fraction in Equation (19) and cancelling out the term $\frac{\mu_1^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2} - \left(\frac{\mu_1\mu_2}{\sigma_1^2} + \frac{\mu_2\mu_1}{\sigma_2^2}\right)^2$, the following expression for the probability density function is obtained:

$$pdf(\beta_{it}) = \frac{\sqrt{\sigma_1^2 + \sigma_2^2}}{2\pi\sigma_1\sigma_2} e^{-\frac{(\beta_{it}-\mu_1)^2}{2\sigma_1^2} + \frac{(\beta_{it}-\mu_2)^2}{2\sigma_2^2}}$$

(21)

What proves that the corrected pdf for $\beta_{it}$ is also normal with mean, shown in Equation (22), and variance, shown in Equation (23).

$$\mu = \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_1^2 + \sigma_2^2}$$

(22)

$$\sigma^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

(23)

From Equation (23) it is straightforward that $\sigma^2 < \min(\sigma_1^2, \sigma_2^2)$, which clearly states that, by applying this method, a better estimate of $\beta_{it}$ is obtained. Thus, the article shows that the expression Equation (4) can be substituted by a direct and more simple-to-compute form, given that the above-mentioned assumptions hold [33–35].

Such an approach of forecasting parameters, as will be shown in Section 5, will provide a better financial result for a financial institution compared to the traditional way of fitting a model on a dataset, available for period $t$-1, and simply applying obtained parameters for period $t$.

4. Empirical Reasoning for the Method

To instantiate the above-mentioned statement concerning time-dependent parameters, when dealing with scoring models in an economic environment, we provide the results of an empirical investigation for a microfinancial organisation.

The study uses data from 2015 to 2019 from Thomson Reuters. For the experiment only two factors were used: “Closed credit sum” and “Number of payments to loan term.
ratio” to construct ten models for 5 consecutive years, each of them was fitted on a dataset covering only 6 months. All in-sample results (2015–2019) are shown in Table 1.

Table 1. Coefficients’ dynamics for a two-factor logit model.

| Subsample Number | Time Range          | Closed Credit Sum | Number of Payments to Loan Term Ratio |
|-------------------|---------------------|-------------------|---------------------------------------|
|                   | From    | To       | Coef     | p-Value | Coef | p-Value |
| 1                 | 12.01.2015 | 29.06.2015 | -0.767   | 0.039   | -0.147 | 0.718 |
| 2                 | 01.07.2015 | 30.12.2015 | -0.449   | 0.061   | -0.432 | 0.126 |
| 3                 | 11.01.2016 | 30.06.2016 | -0.200   | 0.479   | -0.375 | 0.225 |
| 4                 | 01.07.2016 | 31.12.2016 | -0.569   | 0.002   | -0.505 | 0.011 |
| 5                 | 01.01.2017 | 30.06.2017 | -0.522   | 0.003   | -0.403 | 0.038 |
| 6                 | 01.07.2017 | 31.12.2017 | -0.729   | 0.000   | -0.521 | 0.001 |
| 7                 | 01.01.2018 | 30.06.2018 | -0.687   | 0.000   | -0.462 | 0.000 |
| 8                 | 01.07.2018 | 31.12.2018 | -0.427   | 0.000   | -0.386 | 0.000 |
| 9                 | 01.01.2019 | 30.06.2019 | -0.306   | 0.000   | -0.304 | 0.000 |
| 10                | 01.07.2019 | 31.12.2019 | -0.254   | 0.000   | -0.188 | 0.000 |

As it can be seen from Table 1, the subsample size is growing since the considered organisation is expanding. Obtained coefficients are mostly significant except for a couple of subsamples where the number of observations was relatively small.

In Figures 1 and 2 the dynamics of coefficients for the two above-mentioned factors and their polynomial approximations are shown. It is easy to track the cyclicity of the tested parameters, which proves that true parameters are time-dependent.

![Figure 1. Coefficient value dynamics for closed credit sum.](image)

Such a simple empirical experiment shows that the methods, developed in the paper, can substantially boost the performance of logit models when developing a credit risk analysis system in a real-life economy.
5. Simulation Experiment

In our simulation experiment we consider a three-factor logit model:

\[ p(1|X_t) = \frac{1}{1 + e^{-\beta_{t0} - \beta_{t1}x_{t1} - \beta_{t2}x_{t2} - \beta_{t3}x_{t3}}} \]  

(24)

For each parameter, an autoregressive cyclical process with different periodicity is generated. Then we generate samples for each modelling period according to model Equation (24). For testing the proposed methods we use generic MLE (maximum likelihood estimation) to fit parameters’ values to a generated sample and auto.arima function in R environment to build ARIMA class models for parameters’ predictions. Figures 3–6 illustrate estimated by MLE parameters (circles) and their ARIMA prediction (solid line).

Figure 2. Coefficient value dynamics for number of payments to loan term.

Figure 3. Coefficient value dynamics for \( \beta_{01} \).
Figure 3. Coefficient value dynamics for $\beta_t^0$.

Figure 4. Coefficient value dynamics for $\beta_t^1$.

Figure 5. Coefficient value dynamics for $\beta_t^2$.
To prove the efficiency of the proposed method we compare it with several other popular classification techniques. Comparisons are performed on the out-of-sample performance, by analogy with [36], of (1) a generic logit model, where obtained coefficients were applied at period $t-1$ to borrowers, coming at period $t$, (2) random forest algorithm from library “randomForest” in R package (default settings), (3) gradient boosting algorithm from library “gbm” in R package (default settings), (4) forecasted coefficients model, where parameters’ values are forecasted for period $t$ and applied them to borrowers, coming at period $t$, (5) forecasted coefficients by a combination of state space approach and DCC-GARCH model. The interest rate is set at the level of 10% which will also be the cut-off rate for approving a loan. Considered methods are tested on a time range of 30 periods, each of them including 20,000 potential borrowers, each of them borrowing 1 money unit. Received profit is then calculated as follows:

$$profit = (nl - nd) \times r - nd$$

(25)

where $nl$ is the number of issued loans, $nd$ is the number of defaulted loans, $r$ is the interest rate. The summary of such investigation is presented in Table 2 and Figure 7.

Table 2. The numerical forecasting results of proposed models and other ML approaches.

| Method        | Profit Received | Number of Issued Loans | Number of Defaulted Loans | Percent of Defaulted Loans | Mean Squared Forecast Error | Mean Absolute Error |
|---------------|-----------------|------------------------|---------------------------|----------------------------|-----------------------------|--------------------|
| Generic logit | 19,414          | 301,291                | 9741                      | 3.23%                      | 0.030                       | 0.035              |
| Random Forest | 19,627          | 301,232                | 9542                      | 3.17%                      | 0.030                       | 0.035              |
| Gradient boosting | 19,735       | 301,168                | 9438                      | 3.13%                      | 0.030                       | 0.035              |
| ARIMA         | 24,537.6        | 300,904                | 5048                      | 1.68%                      | 0.024                       | 0.029              |
| State space   | 28,531.8        | 300,831                | 1405                      | 0.47%                      | 0.016                       | 0.021              |
| DCC-GARCH     |                 |                        |                           |                            |                             |                    |
As can be seen from Table 2, the number of issued loans is almost the same for all considered methods (for the proposed methods we issued slightly fewer loans than for the logit model and two machine learning methods), however the difference in bad loans is dramatic: around nine and a half thousand for generic logit, random forest and gradient boosting against five thousand for forecasted by ARIMA coefficients and only one and a half thousand for state–space DCC-GARCH model.

Consequently, the percent of defaulted loans for proposed methods is significantly lower, which results in the fact, that profit, received according to the proposed methods by 26% exceeds the profit of generic logit for forecasted by ARIMA coefficients and by 47% for state–space DCC-GARCH model.

Figure 7 clearly illustrates the absolute supremacy of the proposed methods over generic logit almost for all 30 considered time periods. Profit, received by the trivial method has a distinct seasonality with sharp drops, whereas profit, received by proposed methods remains stable over time and, thus, can be considered as a dominant strategy for credit risk management systems.

Figure 8 shows the dynamics of machine learning methods efficiency against the proposed state–space DCC-GARCH model. Machine learning approaches display a similar pattern to the generic logit model, although slightly better efficiency. Gradient boosting method is by a small margin outperforms random forest. The proposed method is almost for all considered periods significantly outperforms tested machine learning methods and thus, can be considered as a preferable alternative.

Summarising the results of performed experiments it can be stated, that in the economic environment, when the assumption of constant true parameters is violated, it is more rational to apply forecasted coefficients’ estimates instead of ones, simply computed by the accumulated dataset that reflects already past interconnections within the data.
6. Conclusions

In this paper, a new method of dealing with time-varying parameters of scoring models is proposed, which is aimed at computing the default probability of a borrower. It was empirically shown that in a continuously changing economic environment factors’ influence on a target variable is also changing. Therefore, forecasting coefficients yields a better financial result, than simply applying parameters, obtained by accumulated statistics over past time periods. The paper fills the gap in methods incorporating adjustments of obtained coefficients, performed by Kalman filter since we get two independent estimates: first—by forecasting a parameters’ vector, second—by fitting parameters to accumulated statistics. It was shown by conducting an extensive simulation experiment, that proposed method is significantly better than the traditional approach and, thus, can be considered as a preferable alternative when dealing with credit risk assessment models in a volatile economic environment.

Author Contributions: Conceptualisation, A.S.; methodology, N.Z. and L.K.; validation, M.S.S.D.; Writing—Original draft preparation, N.M. and A.M.; writing—review and editing, A.M. All authors have read and agreed to the published version of the manuscript.

Funding: N.M. and A.S. performed the study in the framework of the state task in the field of scientific activity of the Ministry of Science and Higher Education of the Russian Federation, project “Development of the methodology and a software platform for the construction of digital twins, intellectual analysis and forecast of complex economic systems”, grant no. FSSW-2020-0008.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of Financial University under the Government of Russian Federation.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bansal, G.; Sinha, A.P.; Zhao, H. Tuning data mining methods for cost-sensitive regression: A study in loan charge-off forecasting. *J. Manag. Inf. Syst.* 2008, 25, 315–336. [CrossRef]

2. Zhang, H.; Legro, R.S.; Zhang, J.; Zhang, L.; Chen, X.; Huang, H.; Casson, P.R.; Schlaff, W.D.; Diamond, M.P.; Krawetz, S.A.; et al. Decision trees for identifying predictors of treatment effectiveness in clinical trials and its application to ovulation in a study of women with polycystic ovary syndrome. *Hum. Reprod.* 2010, 25, 2612–2621. [CrossRef]
3. Smith, L.D.; Lawrence, E.C. Forecasting losses on a liquidating long-term loan portfolio. *J. Bank. Financ.* 1995, 19, 959–985. [CrossRef]
4. Greenidge, K.; Grosvenor, T. Forecasting non-performing loans in Barbados. *J. Bus. Financ. Econ. Emerg. Econ.* 2010, 5, 80–108.
5. Abdou, H.A.; Pointon, J. Credit scoring, statistical techniques and evaluation criteria: A review of the literature. *Intell. Syst. Acc. Financ. Manag.* 2011, 18, 59–88. [CrossRef]
6. Darroch, J.N.; Ratcliff, D. Generalized iterative scaling for log-linear models. *Ann. Math. Stat.* 1972, 43, 1470–1480. [CrossRef]
7. Eddy, Y.; Engku Abu Bakar, E.M.N. Credit scoring models: Techniques and issues. *J. Adv. Res. Bus. Manag. Stud.* 2017, 7, 29–41.
8. Eletter, S.F.; Yaseen, S.G.; Elrefaei, G.A. Neuro-based artificial intelligence model for loan decisions. *Am. J. Econ. Bus. Adm.* 2010, 2, 27. [CrossRef]
9. Khashman, A. Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. *Expert Syst. Appl.* 2010, 37, 6233–6239. [CrossRef]
10. Finlay, S. Are we modelling the right thing? The impact of incorrect problem specification in credit scoring. *Expert Syst. Appl.* 2009, 36, 9065–9071. [CrossRef]
11. Kamaloo, E.; Saniee Abadeh, M. Credit risk prediction using fuzzy immune learning. *Adv. Fuzzy Syst.* 2014, 2014, 7. [CrossRef]
12. Durand, D. *Risk Elements in Consumer Installment Financing*; National Bureau of Economic Research: New York, NY, USA, 1941.
13. Makowski, P. Credit scoring branches out. *Credit World* 1985, 75, 30–37.
14. Angelini, E.; Di Tollo, G.; Roh, A. A neural network approach for credit risk evaluation. *Q. Rev. Econ. Financ.* 2008, 48, 733–755. [CrossRef]
15. Henley, W.; Hand, D.J. AK-Nearest-Neighbour Classifier for Assessing Consumer Credit Risk. *J. R. Stat. Soc. Ser. D Stat.* 1996, 45, 77–95.
16. Hurley, M.; Adebayo, J. Credit scoring in the era of big data. *Yale J. Law Tech.* 2016, 18, 148.
17. DAVIS, R.H.; Edelman, D.B.; Gammerman, A.J. Machine-learning algorithms for credit-card applications. *IMA J. Manag. Math.* 1992, 4, 43–51. [CrossRef]
18. Frydman, H.; Altman, E.I.; Kao, D.L. Introducing recursive partitioning for financial classification: The case of financial distress. *J. Financ.* 1985, 40, 269–291. [CrossRef]
19. Zhou, S.R.; Zhang, D.Y. A nearly neutral model of biodiversity. *Ecology* 2008, 89, 248–258. [CrossRef]
20. Jensen, H.L. Using neural networks for credit scoring. *Manag. Financ.* 1992, 18, 15–26. [CrossRef]
21. West, D. Neural network credit scoring models. *Comput. Oper. Res.* 2000, 27, 1131–1152. [CrossRef]
22. West, D.; Dellana, S.; Qian, J. Neural network ensemble strategies for financial decision applications. *Comput. Oper. Res.* 2005, 32, 2543–2559. [CrossRef]
23. Dietterich, T.G. Machine-learning research. *AI Mag.* 1997, 18, 97.
24. Huang, Z.; Chen, H.; Hsu, C.J.; Chen, W.H.; Wu, S. Credit rating analysis with support vector machines and neural networks: A market comparative study. *Decis. Support Syst.* 2004, 37, 543–558. [CrossRef]
25. Zhu, Y.; Xie, C.; Wang, G.J.; Yan, X.G. Comparison of individual, ensemble and integrated ensemble machine learning methods to predict China’s SME credit risk in supply chain finance. *Neural Comput. Appl.* 2017, 28, 41–50. [CrossRef]
26. Opitz, D.; Maclin, R. Popular ensemble methods: An empirical study. *J. Artif. Intell. Res.* 1999, 11, 169–198. [CrossRef]
27. Bitto, A.; Frühwirth-Schnatter, S. Achieving shrinkage in a time-varying parameter model framework. *J. Econ.* 2019, 210, 75–97. [CrossRef]
28. Chan, J.C.; Eisenstat, E. Bayesian model comparison for time-varying parameter VARs with stochastic volatility. *J. Appl. Econ.* 2018, 33, 509–532. [CrossRef]
29. Kalli, M.; Griffin, J.E. Time-varying sparsity in dynamic regression models. *J. Econ.* 2014, 178, 779–793. [CrossRef]
30. Orlando, G.; Pelosi, R. Non-performing loans for Italian companies: When time matters: An empirical research on estimating probability to default and loss given default. *Int. J. Financ. Stud.* 2020, 8, 68. [CrossRef]
31. Aslan, A.; Poppe, L.; Posch, P. Are sustainable companies more likely to default? Evidence from the dynamics between credit and ESG ratings. *Sustainability* 2021, 13, 8568. [CrossRef]
32. Orlova, E.V. Methodology and models for individuals’ creditworthiness management using digital footprint data and machine learning methods. *Mathematics* 2021, 9, 1820. [CrossRef]
33. An, J.; Mikhaylov, A.; Jung, S.-U. A Linear Programming Approach for Robust Network Revenue Management in the Airline Industry. *J. Air Transp. Manag.* 2021, 91, 101979. [CrossRef]
34. Mikhaylov, A. Development of Friedrich von Hayek’s theory of private money and economic implications for digital currencies. *Terra Econ.* 2021, 19, 53–62. [CrossRef]
35. An, J.; Mikhaylov, A.; Richter, U.H. Trade War Effects: Evidence from Sectors of Energy and Resources in Africa. *Helijon* 2020, 6, e05693. [CrossRef] [PubMed]
36. Zhu, Y.; Zhou, L.; Xie, C.; Wang, G.-J.; Nguyen, T.V. Forecasting SMEs’ credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach. *Int. J. Prod. Econ.* 2019, 211, 22–33. [CrossRef]