CHAPTER 8

Application of Industry 4.0 Concept in Financial Services. Using Artificial Neural Network for Credit Scoring Model in Rating Agency

Jacek Krawiec[1], *Paweł Śliwiński*[2]

[1] Poznań University of Economics and Business, Poland; ORCID 0000-0003-1399-8714
[2] Poznań University of Economics and Business, Poland; ORCID 0000-0001-8479-3252

Summary. The chapter is a case study showing how new technologies can have a huge impact on financial services. The idea of Industry 4.0 affects various spheres of business. The use of digital technologies has contributed also to the development of the financial sphere. The objectives of this study were threefold. Firstly, to show how the concept of Industry 4.0 can influence the financial services on the example of the rating agency; secondly to illustrate business – science cooperation in developing and commercializing the scoring model; and thirdly to contribute to the literature on bankruptcy prediction by building a credit risk assessment model in accordance with the guidelines of the European Securities and Markets Authority (ESMA). The presented model was based on publicly available financial statements of small and medium-sized enterprises listed on the Warsaw Stock Exchange and was part of a rating agency's credit scoring system registered, was reported to and approved by ESMA. The model is used nowadays in business practice and in science. The study can be used in the discussion of how organizational development and new technologies affect financial markets by providing risk-related information, hitherto not available to ordinary investors.

Keywords: scoring model, rating agency, Industry 4.0, financial innovations

1. Introduction

Throughout their history, organizations have undergone constant changes that directly translated into their development. Periodically, the pace of this development increased so much that these periods are commonly called industrial revolutions. We are currently dealing with the fourth such period. The concept of Industry 4.0 was introduced in Germany in 2011 during the Hanover trade fair and was adopted in 2013 by the German government in the "High-Tech Strategy 2020 Action Plan" (Wang et al., 2016). This initiative was to symbolize the new industrial revolution. Among many definitions of Industry 4.0 concept which were collected, e.g., by Tay et al. (2018) special attention should be given to the

following. Kagermann et al. (2013) note that Industry 4.0 utilizes the power of communication technology and innovative inventions to increase the development of the manufacturing industry. Wang et al. (2016) point out that Industry 4.0 makes full use of emerging technologies (e.g., Internet of Things, big data, cloud computing. wireless sensor networks, embedded system) and together with the rapid development of machines and tools it copes with global changes to improve industry level. According to Rüssmann et al. (2015) new digital industrial technology is characterized by nine technologies: (i) autonomous robots, (ii) big data and analytics, (iii) augmented reality, (iv) simulation, (v) horizontal and vertical system integration, (vi) the industrial internet of things, (vii) cybersecurity, (viii) cloud-based services and (ix) additive manufacturing (3D printing). New technologies are changing the economy. For example, Artificial Intelligence (AI), one of the driving force of Industry 4.0, alone can influence global economic activity by adding 1,2% extra GDP growth per year by 2030 (McKinsey Global Institute, 2016).

Machoń and Brodziak (2018) indicate that in Poland – excluding the financial sector – only 10% of enterprises used new technologies. However, more and more companies notice and implement their potential. For example, half of Polish AI companies introduced the technology over the last two-three years. The main obstacles for their development are lack of understanding of potential benefits arising from AI and access to insufficient data. According to Digital Poland Report on AI (Digital Poland Foundation, 2019) two services that are most frequently provided by Polish AI companies are analytics, big data and business intelligence (43%) and sales, marketing and advertisement (37%). 28% of companies using AI provide services in financial services and insurance and 27% of them in the area of internet of things and Industry 4.0.

Mrugalska and Wyrwicka (2017) notice that Industry 4.0 is an advantage to stay competitive in any industry but digitalization and virtualization affect various spheres of business, not just industry. The use of digital technologies has contributed also to the development of the financial sphere. New technologies are affecting financial services in many ways. Among the most important of them, PwC (2020) lists digitalization, blockchain technologies, big data, advances in robotics and AI, and public clouding. The aim is to make financial services more competitive and more accessible to customers. This rapid change creates a chance for new companies, start-ups in many cases, to compete with large, well-established major financial institutions which were often beneficiaries of their size, history, and scale of operations. The technological challenge is created by fintech companies in almost all financial services segments. The number of digital banks, crowdfunding, and peer-to-peer lending platforms or other digital platforms on which customers have access to insurance and banking products offered by different market players is increasing as many financial services are transferred to the virtual world. Customers got used and liked mobile and online banking, virtual financial consultations or automated matching platforms.

New technology sector both globally and in Poland is highly supported by the university community. Digital Poland Report shows that 77% of companies in AI sector collaborate with scientists to engage researchers and students in the day-to-day operations of the company. Some statistics seem particularly interesting. Nearly half of AI companies engage academic researchers in the development of their AI solutions. Also, half of them employ scientists with a minimum PhD degree and based on the work and research in 39% of companies papers were published.

The objectives of this study are threefold. Firstly, to show how the concept of Industry 4.0 can influence the financial services on the example of the rating agency, secondly to illustrate the cooperation between science and business developing and commercializing the scoring model, and thirdly to contribute to the literature on bankruptcy prediction by building a credit risk assessment model in accordance with the guidelines of the European Securities and Markets Authority (ESMA) regarding the construction of rating models and their validation.

The chapter is organized in five sections. Following the introduction, the next section describes general idea of application of Industry 4.0 concept in financial services on the example of the rating agency. This part gives an overview of the problem to be solved by the R&D team consisting of scientists and practitioners. Section 3 deals with the problem how to use Artificial Intelligence in credit risk assessment on the example of INC Rating scoring model. In the literature review, we show growing popularity of more advanced data analysis methods in models used for bankruptcy prediction. Next, the theoretical framework for using neural networks in building the credit risk assessment model is presented, followed by proposition of the methodology of building the credit risk assessment model based on artificial neural networks. The fourth section is of application character, it is a practical implementation of the earlier considerations. The bankruptcy forecasting model based on the neural network was built and then tested against alternative credit risk assessment models. At the beginning of this part, we show some proposal of alternative models of forecasting bankruptcy, and next we validate our scoring model. In the last section, we draw the main conclusions.

2. Application of Industry 4.0 guidelines in case of rating agency

Fitch Polska S.A., EuroRating Sp. z o.o. and INC Rating Sp. z o.o. are registered by ESMA (the European Securities and Market Authority) as rating agencies located in Poland[1]. None of them has developed universal financial models for providing scoring services for local government units, municipal companies and commercial companies adapted to the conditions of the Polish economy and the Polish legal system.

Identifying the market gap, INC Rating[2] has taken up the technological challenge creating original concepts for information processing and innovations in the form of developing and implementing new intelligent optimization algorithms in the field of financial models – credit risk assessment model (scoring model). The project was submitted to the Innovation Development Program under Operational Objective 2: "Increasing pro-innovative attitudes in enterprises" as part of Intelligent Specialization 3 "Innovative Industry" (ICT industry). The project was part of the operational objective 3.3. "Supporting activities in the field of product and technology modernization". Under the project, INC Rating established cooperation with

[1] These agencies are directly registered in Poland. Large US rating agencies: S&P Moody's and Fitch also provide activities in Poland but on a smaller scale (Śliwiński et al., 2017).

[2] INC Rating was registered as a credit rating agency in the EU by ESMA in 2015. It is a private company being a part of INC group which is listed on the Warsaw Stock Exchange. The company was established by INC and financed by "Interior" VC fund.

the Poznań University of Economics strengthening knowledge transfer mechanisms between science and business to commercialize research results[3].

Credit risk assessment model was based on algorithms created specifically for the needs of the project using IT methods, Big Data technology and LOD software[4]. As a result a specialized ICT tool was created which automatically generates credit scores of local government units, municipal companies, commercial companies. The service is provided using a dedicated web platform. The platform exchanges information between INC Rating the client. The client can send the rating agency the data needed for analysis (if the data of a particular entity is not present in the databases owned by INC Rating). This minimizes the client's involvement in the service provision process. The information provided by the client is automatically uploaded to the program containing the appropriate financial model depending on the client's legal form (local government unit/municipal company /commercial law company). Dedicated system performs credit risk assessment, and the customer receives a response within several days of placing the order via the Internet (compliance check of the data sent by the customer is needed, because of ESMA requirements).

After completing the project, improvement of efficiency and effectiveness of processes in the rating agency was achieved. The rating agency provides scoring services via the web platform in an automated manner. This translates into the low cost of providing services. The company assumes the cost of performing a single scoring service would be 10,000 PLN without the use of automated credit risk assessment model due to high personnel costs. INC Rating can offer automatic scoring services at prices ranging from 800 to 3000 PLN. There are no competitive solutions for scoring services on the Polish market for local government units, municipal and commercial companies. The only potential alternative is rating. However, it is an expensive tool that only the largest and richest companies and local government units can afford (the approximate fee for the rating of debt securities issued by LGU ranges from PLN 30,000 to 135,000). Low price and high speed of service should provide the rapid popularization and sprawl of the new market of scoring services.

The Polish Financial Supervisory Authority (KNF – Komisja Nadzoru Finansowego), the Warsaw Stock Exchange and investing intuitions draw attention to the lack of credit and solvency assessment tool especially for debt issuers. The automated scoring model can be used among others by borrowers to improve the negotiating position during talks with banks on the cost of credit or to assess the rationality of decisions in terms of debt costs. In turn, investors should obtain a universal and effective tool for making investment decisions. In addition, scoring can be used by EU funds disposers to determine the financial capacity of applicants to implement projects and maintain their results. Currently, the financial assessment is based mainly on the applicants' forecasts, which are often unrealistic.

---

[3] The authors took part in the project as the academic director (project manager) and analyst, three Ph.D. students were also engaged in the project which was then consulted with INC Rating Committee employing Ph.Ds from Poznan University of Business and Economics. Based on the project and R&D results, one Ph.D. dissertation was defended and the other one is being prepared.

[4] Linked Open Data (LOD) technology describes the method of publishing structured data, so that connections can be made between them, making them more useful and machine-readable and usable.

3. Using Artificial Intelligence in credit risk assessment on example of INC Rating scoring model

3.1. Bankruptcy prediction in literature

The problem of bankruptcy prediction is the subject of many studies and research works[5]. The dynamics of using models in bankruptcy prediction is shown in Figure 1.

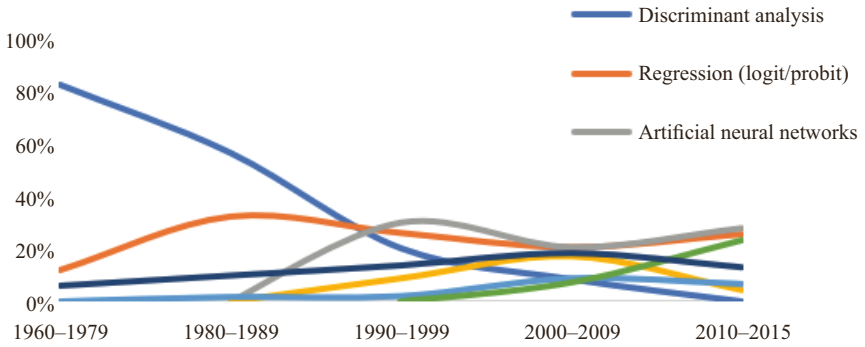

Fig. 1. Models used for bankruptcy prediction in reviewed literature in 1960–2015

Initially, after the publication of Altman (1968), most of the work reproduced the method of discriminant analysis he used. In the following years, the popularity of discriminant analysis declined significantly. At the same time, a significant increase in interest in regression methods can be observed. The largest percentage of publications containing logistic or probit models was recorded in the 1980s. A metanalysis of the listed publications[6] indicates that with the development of computing power, the popularity of more advanced data analysis methods has increased. The most popular models used since 1990 are artificial neural networks, logistic regression and support-vector machines which replaced discriminatory models and models based on market data.

Artificial neural networks are mathematical models derived from neurobiological inspiration[7]. One of the first attempts to use artificial neural networks in modelling bankruptcy was made by Odom and Sharda (1990). They used 5 variables previously proposed by Altman (1968). The model they created contained one hidden layer and one neuron in the output layer. The network was trained using a set consisting of 74 cases (38 bankrupt and 36 healthy companies). The testing set included 55 enterprises (27 of them went bankrupt). For comparison, the authors also created a discriminant analysis model. Ultimately, the artificial neural

[5] The detailed literature review is delivered by e.g. Bellovary et al. (2007), Aziz and Dar (2006) and Balcean and Ooghe (2006).

[6] Literature reviews made by Bellovary et al. (2007), Balcean and Ooghe (2006), Aziz and Dar (2006) have been supplemented with an original review of publications from 2005–2015. Ultimately, the list included 330 items, including 90 from period 2005–2015.

[7] Suggesting modeling on the structure and operation of the brain. The neural network is however only its mathematical simplification.

network correctly classified 81.81% of cases. The outcome by 7.53 pp exceeded the results obtained by the discriminatory model. The above research initiated the rapid development of artificial neural network models.

3.2. Use of neural networks in building the credit risk assessment model – theoretical framework

The basic module of the model is the neuron. The main task of the neuron is to sum up the input data and pass them to neurons in the next layer, through the activation function, which can take various forms. The logistic function will be used in the process of building the scoring model due to binary character of the dependent variable. A single neuron (O_i) can be thus depicted with the Formula (1):

$$O_i = f\left(w_{i1}x_1 + \ldots + w_{in}x_n + b_i\right) = \frac{1}{1+e^{-\left(w_{i1}x_1+\ldots+w_{in}x_n+b_i\right)}} \tag{1}$$

where w_{in} stands for the weight assigned to neuron's input, x_n represents n input to the neuron and b_i represents the so-called bias. Neurons are grouped in a layer.

The computational units using input data (data from the input layer) form the first hidden layer. The next layer neurons use data from the previous layer. The task of the last layer (output layer) is to present the final network result. A diagram of a four-layer artificial neural network with 2 hidden layers and 5 input data is shown in Figure 2.

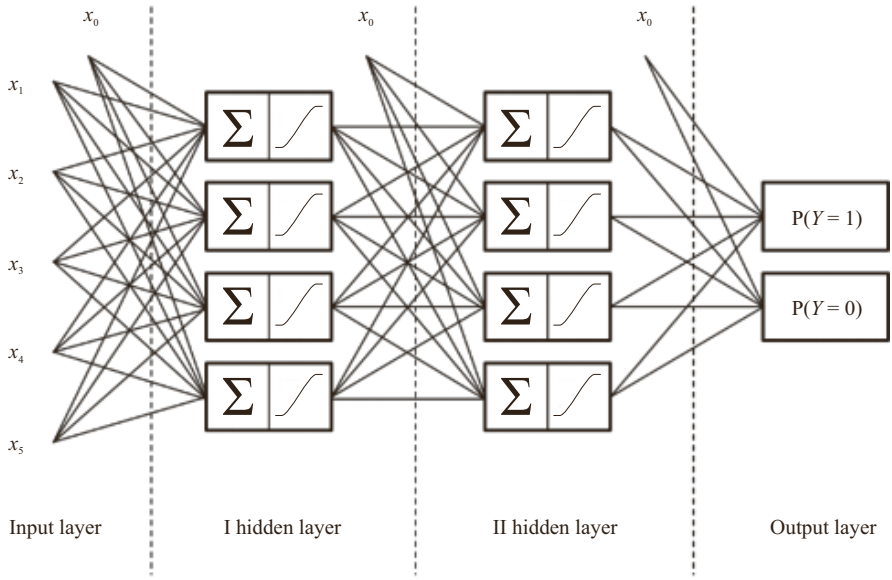

Fig. 2. Diagram of an artificial neural network with two hidden layers

The single layer of the neural network can be described using the Formula (2):

$$o_i = f\left(W_i \cdot o_{i-1} + b_i\right) \tag{2}$$

where o_i represents the column vector of results of neurons placed in the i layer of the artificial neural network. $f(x)$ is the activation function. In the case of the network described in this chapter, it is a logistic function. W_i is the matrix of coefficients w_{jn} of the i network layer, where j is the number of the neuron in the layer and n is the number of the input data to which the coefficient is assigned. The matrix W thus has the following form:

$$W = \begin{bmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{j1} & \cdots & w_{jn} \end{bmatrix} \tag{3}$$

The o_{i-1} vector is a column vector of input data, being the output of the previous layer. Vector b_i is a column vector of bias coefficients assigned to each layer. The two-layer neural network can therefore be described (except the output layer) with the following formula:

$$o_2 = f\left(W_2 \cdot f\left(W_1 \cdot o_0 + b_1\right) + b_2\right) \tag{4}$$

In the case of the network shown in Figure 2, vector o_2 is a four-element column vector. The network is a multi-class classifier. In the process of building the scoring model, it classifies cases into one of two classes (bankrupt or non-bankrupt). The activation function assigned to the output layer neurons is the softmax function. According to (Statsoft, 2018) this function is most appropriate for use in artificial neural networks for classification tasks. The function has the following formula:

$$Y_{ci} = \frac{e^{w_{ci}^T o_{i-1} + b_i}}{\sum_{c=1}^{k} e^{w_{ci}^T o_{i-1} + b_i}} \tag{5}$$

where Y_{ci} is the result of calculations of the neuron c of the output layer (with the number i). This value corresponds to the probability of belonging of the modelled observation to class c. The vector w_{ci}^T is a transposed vector of coefficients located in row c of the W_i matrix. The W_i matrix is of key importance for the functioning of the artificial neural network. This matrix describes the functioning of individual neurons and the entire layers of the network. The quality of a model depends on the proper selection of matrix coefficients. Their selection is automatic but in line with the adopted strategy. The strategy of selecting coefficients is also called the network learning strategy, which is in fact an optimization task. It consists of the selection of the W_i matrix coefficients to minimize the selected error function[8]. Due to the strongly non-linear nature of the models, error functions can have many local extremes. The optimal learning method will achieve a global minimum avoiding local extremes.

[8] There are many strategies for optimizing the W_i matrix, but their description is beyond the scope of this chapter.

3.3. Methodology of building credit risk assessment model based on artificial neural networks

Preliminary selection of variables and data set

As the intended effect is the construction of a scoring model based solely on publicly available data, it was assumed that all data would come from the financial statements published by the assessed enterprises[9]. The preliminary selection of variables was based on a bibliographic query (252 references). 38 indicators were found which appear in at least 5 models[10].

In the credit risk assessment model quarterly consolidated reports were used except for those companies which publish only unconsolidated financial statements. The database contains data from 904 companies whose shares were listed on the main market of the Warsaw Stock Exchange and on the NewConnect alternative trading system run by the Warsaw Stock Exchange on December 31, 2016. In addition, the database contains reports of 335 companies whose shares have ceased to be listed before the indicated date. As the result, the total financial statements of 1239 companies were analysed. The original database size was limited to micro, small and medium enterprises defined according to the European Commission in Regulation No. 651/2014 of 17 June 2014 (*Commission Regulation (EU) No. 651/2014...*, 2014) However, the criterion related to the employment in the company was abandoned, the criteria regarding the revenues and the total value of assets were applied.

Ultimately, 561 companies remained in the database, for which 38 indicators most commonly found in bankruptcy prediction models were calculated. The database also includes 125 cases of bankruptcy. As the above information indicates, the dataset is highly unbalanced. Due to the inability to obtain new data, it was necessary to use other database balancing techniques. In the analysed case it was decided to use the SMOTE (Synthetic Minority Over-sampling Technique) (Chawla et al., 2002) which allows balancing cases of different classes in a data set[11].

Before building the model, data transformation was applied by removing outliers and normalizing the data, which is suggested by (Bishop, 1995; Lai et al., 2006). All transformations were carried out in the database in the manner proposed by Fuertes and Kalotychou (2007)[12]. Data transformation into the range <0; 1> is necessary due to the need to unify data occurring at different scales. The use of data at various scales can lead to unstable operation of the model (Gershenfeld & Weigend, 1994).

---

[9] The financial statements come from databases made available by Notoria S.A.

[10] Starting from return-to-equity ratio (ROE) which appears in 115 out of 252 indications and ending with EBITDA-to-sales, EBITDA-to-total assets and share capital-to-total assets ratios which occurred in 5 models.

[11] As indicated by Blagus and Lusa (2013) this procedure brings the expected results only in sets with a relatively small number of dimensions (up to 300 variables).

[12] However, the winsorization procedure proposed by Fuertes and Kalotychou was modified. All observations whose value exceeded the value of the 5th or 95th percentile were replaced by the values of the corresponding range border.

Selection of variables for model

The goal of building each model should be to find the simplest explanation of the facts, using the smallest number of variables (Lai et al., 2006). The key stage in the construction of an econometric model is the selection of variables for the model. Too many variables can interfere with the model's performance (Bishop, 1995). In addition, with the growing number of potential explanatory variables, it is necessary to provide a sufficiently large training set. This relationship is called the *curse of dimensionality* (Bellman, 1961).

The first step in the process of reducing the number of variables is the initial assessment of their discriminatory power. Two sets of data were determined based on the value of the explained variable (bankrupt/non-bankrupt companies). The Mann–Whitney U (Aczel, 2000, pp. 716–717) test was used to verify the hypothesis about the lack of difference in variable distributions in two samples. Rejection of the null hypothesis allows us to state that examined variable takes different values depending on the group to which it was qualified (1 for bankrupt and 0 for healthy companies). Except for five variables, there is a statistically significant difference in distributions. These variables were further analysed. The next step in the selection of variables was the rejection of variables introducing collinearity to the model[13]. A commonly used way to assess collinearity is the variance inflation factor (VIF) (Marquaridt, 1970).

After initial checks, 28 potential explanatory variables remained. The final selection was carried out using two methods. The first procedure involved iterative rejection of variables whose significance in the model is the lowest. *The backward stepwise elimination procedure* was applied in this process. Maximizing the Gini coefficient was adopted as the decision rule (Witzany, 2017, pp. 44–48). First, a logistic regression model was built, using all the explanatory variables remaining. The database was divided randomly into two groups – training and test set. Then, synthetic observations were generated in the training set to reduce the disparity in the number of observations assigned to both analysed classes. The Gini coefficient was calculated next for both sets. Its value became the base value for the elimination of variables in the next step. This step consists of iteratively rejecting each variable individually from the model and assessing how the rejection of the variable affects the value of the Gini coefficient (in relation to the base value). Finally, the variable was removed from the model, without which the model achieved the highest Gini coefficient relative to the base value. The first (basic set – M1) and the second (reduced set – M2) contain ten and four variables of the highest rank, respectively.

The second procedure used for variables selection was the *random forest algorithm*. It consists in the random creation of decision trees. The number of variables used in each model is drawn. One of the most interesting properties of the algorithm is the ability to directly present the relative importance of each variable. In this way, a list of all potential explanatory variables was obtained. In this case, two sets of variables were recommended, basic and reduced. As in the case of backward stepwise elimination, they had 10 (basic model M3) and 4 variables (reduced model M4) respectively. All the variables which were part of the four sets which would be subject to further analysis are presented in Table 1.

[13] Rejection of these variables is necessary due to comparing the results of various models with the results of the logistic regression model, the application of which requires a lack of collinearity between independent variables (Aczel, 2000, pp. 590–598).

Table 1. Ranking of significance of explanatory variables

Rank	Variable (stepwise elimination)	Variable (random forest)
1	current liabilities-to-total assets	equity-to-liabilities
2	total assets-to-liabilities	current assets-to-current liabilities
3	working capital-to-sales	EBITDA-to-interests
4	long-term liabilities-to-equity	EBIT-to-interests
5	current assets-to-current liabilities	operational cash flow-to-liabilities
6	long term liabilities-to-assets	cash-to-current assets
7	equity-to-liabilities	EBITDA-to-total assets
8	return-to-equity	current assets-to-total assets
9	return-to-assets	sales-to-total assets
10	current assets-to-total assets	sales-to-fixed assets

Construction of scoring model

In the model building process, the Keras artificial neural network library was used, based on the TensorFlow environment. The library was made available based on the open-source software (Chollet, 2015). After choosing the optimal number of training epochs, the parameters of neural network architecture were selected. Finally, the neural network consisted of 3 hidden layers, each containing 15 neurons, and its learning takes place in 25 epochs. The specificity of artificial neural networks does not allow direct determination of the significance of individual variables. For this reason, Receiver Operating Characteristics (ROC) curves and the areas below them (AUROC) were used to recommend one of the analysed models[14]. The curves are presented in Figure 3. The differences in AUROC between all the curves were insignificant. Finally, **M3** and **M2** models were recommended for further analysis.

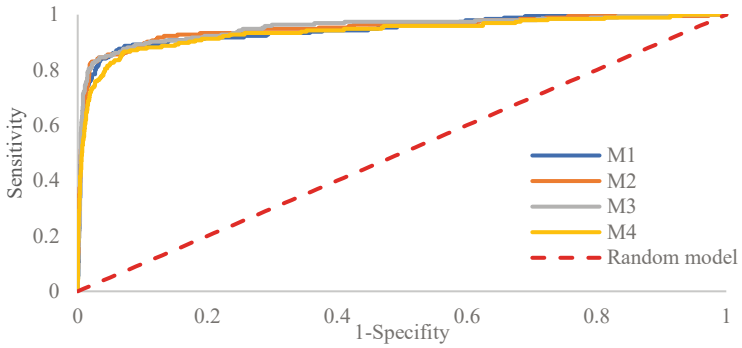

Fig. 3. ROC curves of artificial neural network models

[14] For definition and modelling framework of ROC analysis (see, e.g., Gonçalves et al., 2014).

4. Testing neural network bankruptcy forecasting model against alternative credit risk assessment models

4.1. Alternative models of forecasting bankruptcy of enterprises

As indicated in Figure 1, in addition to neural networks, currently the most frequently chosen methods for forecasting bankruptcy are logistic regressions and support-vector machines.

Unlike standard linear regression, logistic regression can be used when modelling binary variables. For this reason, this method is widely used in forecasting the bankruptcy of enterprises. The variable explained usually is binary. Its values are 1, respectively, in the case of a modelled event (bankruptcy), and 0, if this event does not occur. Another advantage of logistic regression is the standardization of model results in the range (0; 1). This means that they can be read as the probability of occurrence of the modelled phenomenon (Hosmer & Lemeshow, 2000).

Support Vector Machines (SVMs) were first proposed by Vapnik (1995). The model training algorithm focuses on searching for a hyperplane best separating the observations of the training set belonging to two different classes. A hyperplane that best separates observations is understood to be a hyperplane that is located at a maximum distance from observations qualified for different classes. Although the hyperplane indicated by the model is linear, the algorithm is also applicable for non-linear relationships between variables. The solution to this problem is to apply a non-linear mapping of the original set of explanatory variables to spaces with more dimensions and then to find a linear hyperplane separating new observations (Horta & Camanho, 2013).

Building of alternative scoring models was carried out in accordance with the process described earlier: identification of variables, exclusion of variables that lack discriminatory power and introduce colinearity, final selection of variables. Finally, four models were further analysed: logistic regression models (M5 and M6), and support vector machines models (M7 and M8). The results from the regressions M5 and M6 are presented in Tables 2 and 3. Table 4 shows variable weights in support vector machines models (M7 and M8).

Table 2. Logistic basic regression model (M5)

Variable	β	Statistical error	*z*	*p*-value
CD/TA	4.8194	0.1286	37.4712	0.0000
LD/TA	5.7882	0.1915	30.2217	0.0000
LD/E	−3.8377	0.1577	−24.3284	0.0000
ln(TA)	−1.5086	0.1349	−11.1856	0.0000
TA/D	−9.1797	0.5100	−18.0006	0.0000
WC/S	−3.2671	0.2266	−14.4166	0.0000
ROA	−3.4890	0.1649	−21.1524	0.0000
QR	−2.0325	0.6742	−3.0147	0.0026
CF/TA	−0.9663	0.3435	−2.8129	0.0049
CF/S	1.0509	0.3716	2.8280	0.0047

Table 3. Logistic reduced regression model (M6)

Variable	β	Statistical error	*z*	*p*-value
CD/TA	3.0174	0.0658	45.8484	0.0000
LD/TA	3.1363	0.1030	30.4497	0.0000
LD/E	−2.8056	0.1092	−25.7021	0.0000
ln(TA)	−3.7263	0.0852	−43.7139	0.0000

Table 4. Variable weights in support vector machines models (M7 and M8)

SVM model (M7)		SVM model (M8)	
variable	**weight**	**variable**	**weight**
CD/TA	8.66	CD/TA	33.77
LD/TA	10.56	LD/TA	21.16
LD/E	4.23	LD/E	4.27
ln(TA)	1.26	ln(TA)	0.17
TA/D	25.37	–	–
WC/S	8.11	–	–
ROA	4.05	–	–
QR	1.31	–	–
CF/TA	0.38	–	–
CF/S	–	–	–

4.2. Scoring model validation

Assessment of discrimination power of models

The basic way to evaluate the accuracy and credibility of the classification model is to use the confusion matrix. This matrix compares the empirical values of the explained variable with the values predicted by the model. Based on the values entered into the confusion matrix, it is possible to estimate indicators that facilitate the interpretation of the matrix (Fawcett, 2005). The results of the confusion matrix analysis are presented in Table 5.

Table 5. Confusion matrix analysis

	Artificial neural network		Logistic regression		Support vector machines	
	M2	**M3**	**M5**	**M6**	**M7**	**M8**
Sensitivity	0.85	0.85	0.86	0.81	0.87	0.85
Specificity	0.97	0.97	0.94	0.84	0.93	0.95
Precision	0.50	0.50	0.34	0.16	0.31	0.41
Accuracy (ACC)	0.96	0.96	0.93	0.84	0.92	0.95
F1 score	0.63	0.63	0.49	0.27	0.46	0.55
MCC	0.63	0.63	0.51	0.31	0.49	0.57
Odds ratio	162.23	161.24	87.79	21.28	81.54	112.69

The best results were achieved by the reduced artificial neural network model M2 in 6 of the 7 analysed categories. In one (sensitivity), it came in third. Further down are the basic model of artificial neural networks (M3) and the reduced model of support vector machines (M8). The only model that showed the values of the respondents definitely lower than the others was the reduced logistic regression model (M6).

The next stage of the analysis was a comparison of ROC curves, the area under these curve (AUROC), assessment of Gini coefficients and J Youden statistics[15]. The values of all measures are shown in Table 6. Artificial neural networks achieved the best results. The basic model (M2) showed a slightly higher Gini coefficient.

Table 6. Statistics of ROC curves

	Artificial neural network		Logistic regression		Support vector machines	
	M2	**M3**	**M5**	**M6**	**M7**	**M8**
AUC	0.95	0.95	0.94	0.94	0.94	0.94
Gini	0.90	0.91	0.88	0.88	0.88	0.88
J – Youden	0.81	0.81	0.79	0.64	0.79	0.80

The subsequent stages of the analysis were based on the classification of observations into appropriate risk classes. The division of observations into risk classes is presented in Figure 4.

[15] More detailed description of these measures can be found in Hosemer and Lameshow (2000); ECB (2019) or Youden (1950).

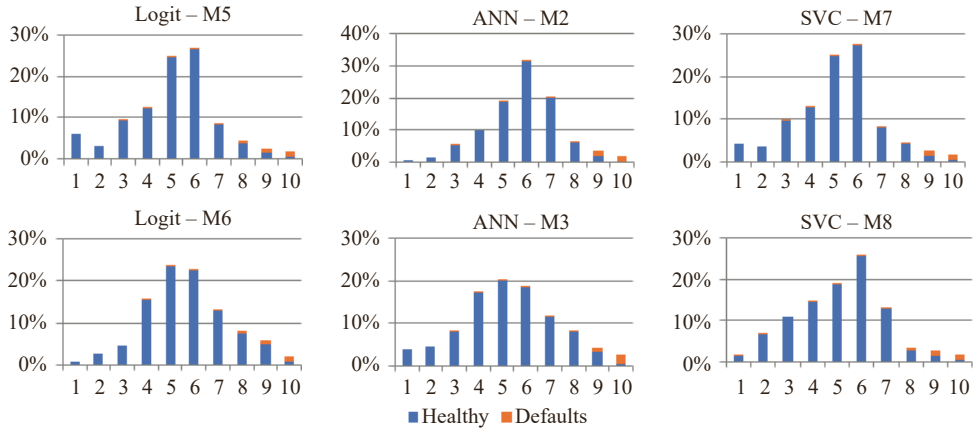

Fig. 4. Division of observations into risk classes

Given the high concentration of bankruptcies in the highest risk classes, this ability can be assessed highly. The quality of the scales developed, despite the identified defects, also seems high. The share of bankruptcy companies continues to increase as the risk class increases. The basic function of each of the scales has therefore been preserved. For most models, over 90% of bankruptcies were in the top four risk classes. Both artificial neural network models achieved the best results in this respect.

Based on the presented scales, χ^2 tests were also carried out. Rejection of the null hypothesis was necessary for all tested models. This is undoubtedly related to the occurrence of bankruptcies in low-risk classes. The construction of χ^2 statistics causes that the occurrence of even single disorders of this type significantly overstates the value of the statistics itself. A comparison of χ^2 statistics shows, however, that the model that was closest to adopting the null hypothesis of the test was the reduced artificial neural network model. This confirms previous observations related to the relatively high quality of this model.

Model calibration

Model calibration assessment was made based on the error function analysis. On average, the lowest error was achieved by the basic model of artificial neural networks. The next place was taken by the basic logistic regression model. All models achieved very low error function values, while the differences between the models in the calibration area were small. The results of the calculations are presented in Table 7.

Table 7. Error function analysis

	Artificial neural network		**Logistic regression**		**Support vector machines**	
	M2	**M3**	**M5**	**M2**	**M7**	**M8**
Mutual entropy	357.9	386.6	421.3	569.7	465.7	453.1
Brier score	0.017	0.016	0.019	0.027	0.019	0.022
MSLE	0.008	0.008	0.009	0.013	0.009	0.010
MAE	0.032	0.026	0.034	0.064	0.035	0.039

Stability assessment

The stability of the models was assessed based on the analysis of migration matrices and coefficients of variation. In all the cases, a strong concentration of observation on the diagonal of the matrix can be seen. As an example migration matrices developed for M2 is presented in Table 8. The reduced artificial neural network model showed a coefficient of variation of 1.25. For this model, the only significant deviation from the diagonal was in the case of class change from 9 to 6. The highest risk classes show a large percentage of enterprises going bankrupt (91% of observations in class 10).

Table 8. Migration matrix of artificial neural networks model M2

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	33%	59%	7%	0%	0%	0%	0%	0%	0%	0%	0%
2	8%	35%	42%	8%	4%	0%	0%	0%	2%	0%	0%
3	0%	8%	64%	24%	1%	1%	0%	0%	0%	0%	1%
4	4%	2%	11%	58%	23%	2%	1%	0%	0%	0%	0%
5	0%	0%	1%	11%	68%	17%	2%	0%	0%	0%	1%
6	0%	0%	0%	1%	8%	76%	11%	1%	1%	0%	1%
7	0%	0%	0%	0%	1%	16%	70%	8%	1%	0%	3%
8	0%	0%	1%	1%	1%	5%	23%	54%	8%	1%	6%
9	0%	0%	0%	0%	1%	6%	8%	7%	11%	0%	67%
10	0%	0%	0%	0%	0%	0%	1%	2%	2%	4%	91%

5. Conclusions

Analysis of presented models indicates their high discriminatory power. Each of them achieved satisfactory values of the indicators taken into account. However, the artificial neural network models were most often indicated as the best. It should be noted that the very high quality of the models means that it is possible to develop scoring models based solely on publicly available data that will meet all the expectations of the relevant institutions regulating the credit assessment market.

The scoring model described in this chapter has been already used to quantify credit risk both in business and in science. In business, it has been tested at INC rating agency and served as the basis for creating a rating methodology for commercial law companies which was approved by the European Securities and Markets Authority. In science, the credit risk assessment model was used to verify scientific hypothesis which were based on the study of the impact of credit risk on equity returns and their volatility in the group of listed SMSs on the Warsaw Stock Exchange. This relationship turns out to be positive, which means that

as credit risk increases, average risk-adjusted return on assets increases. At the same time, their volatility is growing, which directly indicates that credit risk is one of the elements of risk borne by investors. The existence of this relationship means that even if share prices include information on credit risk, they are not obvious to investors who react with a delay. Although the primary data are publicly available, credit risk data are only obtainable after using appropriate models that are not publicly available. However, according to the semi-strong market efficiency hypotheses, share prices should immediately reflect all publicly available information. The above observation could imply the need to reject the hypothesis about the semi-strong market efficiency of the SME market in Poland.

The chapter acts as a case study showing how new technologies can and do have a great impact on financial services. Application of e.g. Artificial Intelligence and big data enabled automatic generation of credit scores of local government units, municipal companies, commercial companies and friendly financial service for clients through a dedicated web platform. This minimizes the client's involvement in the service provision process, improves the speed of service and reduces its costs. This is in line with European Banking Authority (EBA) which in the latest report on big data and advanced analytics (EBA, 2020, pp. 20–21) illustrates their current application with the practical use case – automated credit scoring.

As noticed by PwC (2020), FinTech would drive the new business model which is demanded by customers. Financial customers have had their expectations set by other industries. They demand better services and customer friendly solutions. In a digitalized world it means implementation in finance technologies used in Industry 4.0 concept. This leads to transform the finance industry into Finance 4.0 where such technologies as blockchain, Artificial Intelligence, Internet of Things or Robotics Process Automation can improve traditional financial activities in many fields.

References

Aczel A., 2000. *Statystyka w zarządzaniu*. Wydawnictwo Naukowe PWN, Warszawa.

Altman E., 1968. *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy*. "Journal of Finance", vol. 23(4), pp. 589–609.

Aziz M., Dar H., 2006. *Predicting Corporate Bankruptcy: Where We Stand?*. "Corporate Governance: The International Journal of Business in Society", vol. 6(1), pp. 18–33.

Balcean S., Ooghe H., 2006. *35 Years of Studies on Business Failure: an Overview of the Classic Statistical Methodologies and their Related Problems*. "The British Accounting Review", vol. 38(1), pp. 63–93.

Bellman R., 1961. *Adaptive Control Processes: A Guided Tour*. Princeton University, Princeton, NJ.

Bellovary J., Giacomino D., Akers M., 2007. *A Review of bankruptcy prediction studies: 1930 to present*. "Journal of Financial Education", vol. 33, pp. 3–41.

Bishop C., 1995. *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford.

Blagus R., Lusa L., 2013. *SMOTE for high-dimensional class-imbalanced data*. "BMC Bioinformatics", vol. 14(106), pp. 1–16.

Chawla N., Bowyer K., Hall L., Kegelmeyer P., 2002. *SMOTE: Synthetic Minority Over-sampling Technique*. "Journal of Artificial Intelligence Research", vol. 16, pp. 321–357.

Chollet F., 2015, *Keras*; https://github.com/keras-team/keras [2.11.2019].
Commission Regulation (EU) No. 651/2014 of June 2014 declaring certain categories of aid compatible with the internal market in application of Articles 107 and 108 of the Treaty Text with EEA relevance; https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2014.187.01.0001.01.ENG [2.01.2020].
Digital Poland Foundation, 2019. *Map of the Polish AI*; https://www.digitalpoland.org/assets/reports/map-of-the-polish-ai---2019-edition-i.pdf [20.12.2019].
EBA, 2020. *EBA Report on Big Data and Advanced Analytics, EBA/REP/2020/01*; https://eba.europa.eu/file/609786/download?token = Mwkt_BzI [22.01.2020].
European Central Bank (ECB), 2019. *Instructions for reporting the validation results of internal models. IRB Pillar I models for credit risk*; https://www.bankingsupervision.europa.eu/banking/tasks/internal_models/shared/pdf/instructions_validation_reporting_credit_risk.en.pdf [2.01.2020].
Fawcett T., 2005. *An introduction to ROC analysis.* "Pattern Recognition Letters", vol. 27(8), pp. 861–874.
Fuertes A., Kalotychou E., 2007. *Optimal design of early warning systems for sovereign debt crises*. "International Journal of Forecasting", vol. 23(1), pp. 85–100.
Gershenfeld N., Weigend A., 1994. *Time Series Prediction: Forecasting the Future and Understanding the Past*. "International Journal of Forecasting", vol. 10(1), pp. 161–163.
Gonçalves L., Subtil A., Rosário Oliveira M., de Zea Bermudez P., 2014. *ROC Curve Estimation: an Overview*. "REVSTAT – Statistical Journal", vol. 12(1), pp. 1–20.
Horta I., Camanho A., 2013. *Company Failure Prediction in the Construction Industry*. "Expert Systems with Applications", vol. 40(16), pp. 6253–6257.
Hosmer D., Lemeshow S., 2000. *Applied Logistic Regression*. John Wiley & Sons, New York.
Kagermann H., Wahlster W., Helbig J., 2013. *Securing the Future of German Manufacturing Industry. Recommendations for Implementing the Strategic Initiative Industrie 4.0. Final Report of the Industrie 4.0 Working Group, Forschungsunion*; https://www.din.de/blob/76902/e8cac883f42bf28536e7e8165993f1fd/recommendations-for-implementing-industry-4-0-data.pdf [2.01.2020].
Lai K., Wang S., Yu L., 2006. *An integrated data preparation scheme for neural network data analysis*. "IEEE Transactions on Knowledge and Data Engineering", vol. 3, pp. 217–230.
Machoń A., Brodziak S., 2018. *FinTech in the CEE Region: Artificial Intelligence*. [In:] Fintech in the CEE Region, CEE Capital Market Forum, pp. 50–73; https://paga.org.pl/wp-content/uploads/2019/12/fintech-in-the-cee-region.pdf [27.12.2019].
Marquaridt D.W., 1970. *Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation*. "Technometrics", vol. 12(3), pp. 591–612.
McKinsey Global Institute, 2016. *Digital Finance for All: Powering Inclusive Growth in Emerging Economies*; https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Employment%20and%20Growth/How%20digital%20finance%20could%20boost%20growth%20in%20emerging%20economies/MGI-Digital-Finance-For-All-Executive-summary-September-2016.ashx [22.12.2019].
Mrugalska B., Wyrwicka M.K., 2017. *Towards Lean Production in Industry 4.0*. "Procedia Engineering", vol. 182, pp. 466–473.

Odom M., Sharda R., 1990. *A neural networks model for bankruptcy prediction*. “Proceedings of the IEEE International Conference on Neural Networks”, vol. 2, pp. 163–168.

PwC, 2020. *Financial Services Technology 2020 and Beyond: Embracing disruption*; https://www.pwc.com/gx/en/financial-services/assets/pdf/technology2020-and-beyond.pdf [20.12.2019].

Rüssmann M., Lorenz M., Gerbert Ph., Waldner M., Justus J., Engel P., Harnisch M., 2015. *Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries*; https://www.bcg.com/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries [20.12.2019].

Statsoft, 2018; http://www.statsoft.pl/textbook [4.01.2019].

Śliwiński P., Huczek S., Zygmanowski P., 2017. *Ratingi jednostek samorządu terytorialnego – podstawy metodologii pracy agencji ratingowych*. “Finanse Komunalne”, vol. 6, pp. 18–30.

Tay S.I., Chuan L.T., Aziati A.H.N., Ahmad A.N.A., 2018. *An Overview of Industry 4.0: Definition, Components, and Government Initiatives*. “Journal of Advanced Research in Dynamical and Control Systems”, vol. 10(14), pp. 1379–1387.

Vapnik V., 1995. *The nature of statistical learning theory*. Springer, New York.

Wang S., Wan J., Li D., Zhang C., 2016. *Implementing Smart Factory of Industrie 4.0: An Outlook*. “International Journal of Distributed Sensor Networks”, vol. 6(2), pp. 1–10.

Witzany J., 2017. *Credit Risk Management. Pricing, Measurement, and Modelling*. Springer, New York.

Youden W., 1950, *Index for rating diagnostic tests*. “Cancer”, vol. 3(1), pp. 32–35.