STATISTICAL MODELS FOR CORPORATE CREDIT RISK ASSESSMENT – RATING MODELS

ANETA PTAK-CHMIELEWSKA

Introduction

- Multivariate models for forecasting bankruptcy of the enterprises were introduced by Altman in 1968 but the works on those types of models have much shorter history in Poland.
- The implementation of the western models to the market of enterprises functioning in the conditions of transitional economy as we had in Poland was not successful. It appeared that those models are not working in conditions of political and economic changes. Insufficient effects of adopting foreign models to Polish conditions contributed to the development of research concerning local models. The biggest popularity, similarly to the situation abroad, was gained by the models based on **discriminant analysis**.
- In the 90s the activities were started to build and implement the models adjusted to the specifics of the Polish economy (papers i.e. by Hadasik, Mączyńska, Gajdka i Stos, Pogodzińska & Sojak). The multivariate analysis the regression models and neural networks models were used. Amongst the authors of the bankruptcy models who published their papers after 1990 we can enumerate: Appenzeller, Wędzki, Pogorzelski, Hołda, Michaluk, Gruszczyński, Mączyńska and Zawadzki, and many others.
- The traditional models do not take into consideration changes in time which can be significant. Such changes in time are reflected in the survival models (so called Event History Analysis), the application of which is more and more present in scientific papers (Ptak-Chmielewska & Schab).

From the historical perspective

- **Gajdka & Stos** developed models to predict bankruptcy presented in 1996.
- **Hadasik** used multivariate discriminant function in her models (1998).
- **Hołda** estimated model based on the sample of 80 enterprises (40 bad and 40 good) using multivariate discriminant analysis method.
- Models created by **Prusak** were built based on the sample used for financial ratio analysis of entities.
- Logit models were used by **Stępień & Strąk.** The sample of bankruptices consisted of 39 companies with legal bankruptcy applications submitted in the Court of Szczecin in years 1996-1998.
- In her works **Appenzeller** included dynamic changes in the multivariate discriminant model.
- Research conducted by **Mączyńska** also used discriminant functions. The sample was based on 80 entities and FS data from years 1997-2002.
- **Korol** revealed the advantage of neural networks comparing to discriminant function based on 180 enterprises and FS from period 1998-2001. **Strąk** used decision trees convincing that it is worth moving analysis beyond the traditional discriminant approach.
- **Dębkowska** compared the discriminant analysis method with the logistic regression and decision trees based on the sample of 68 enterprises and FS from 2009.
- **Ptak-Chmielewska** compared the survival analysis with the logistic regression and the discriminant analysis.

Methods & Models

- Discriminant analysis (Fisher multivariate linear function)
- Logistic regression
- Survival models (semiparametric Cox regression model)
- Decision trees
- Neural netoworks

Discriminant analysis

- **The discriminant analysis** is used for classification purposes into two groups: good and bad. This classification is based on the function. In this method the set of variables (interval) is used to construct a rule that distinguishes between good and bad in the best possible way.
- The main purpose is to correct classification into groups. The function maximizes the distance between subpopulations.
- There are some limitations of this method. The variables must be normally distributed which is quite often violated. The next assumption is the equality of variances between groups. Those assumptions must be verified before the final model is estimated.
- A balanced sample is required to obtain good classification and to correct errors estimation.

Discriminant analysis

The discriminant functions, on which the multivariate models detection systems warn against the bankruptcy can assume various forms - they can be linear or square functions, etc. The linear discriminant function usually takes the form:

$$\mathbf{Z = a_0 + a_1X_1 + a_2X_2 + \ldots + a_nX_n,}$$

where:

Z – dependent variable,

a_0 – intercept,

a_i, i = 1, 2, …, n – discriminative coefficients (weights),

X_1, X_2, …, X_n – explanatory variables (financial ratios).

The presented form of discriminant functions called Fisher discriminative function parameters *a*, called discriminative coefficients/ratios (weights).

After determining the form of the discriminative function the cut off value is determined which allows for classifying the entity as financially risky or not risky in a definite manner.

The average value of the discriminant function in specific groups and the cut-off value half-way between the average values are most frequently determined. If 2-value for a given enterprise is lower than Z cut-off, then this entity is classified as susceptible to bankruptcy risk, and if it is higher, then the entity is considered healthy.

Advantages and disadvanages of the linear discriminative analysis

ADVANTAGES	DISADVANTAGES
• easy to understand and easy to apply, • simultaneously taking into consideration many variables due to weights applied, • getting one dimension from multivariable space through transformations in order to evaluate he situation on the basis of selected measure, • possibility to determine the impact of particular explanatory variables on the dependent variable (though not always), • classifications are very precise in the area of analysing the bankruptcy risk of the enterprises, • less costly than e.g. neutral networks, • possibility to apply for dynamic analysis • availability of this method in many popular programs.	• the models can get outdated due to changing economic and spatial conditions – small stability, • difficulties to apply qualitative variables which have significant impact on the enterprise situation e.g. human factor, • quickly outdated because of time criteria, • assumption on normality of explanatory variables distribution, • assumption on equality of the matrix variance – covariance groups of individuals, • the necessity to know opinion probability of the population and costs of error of 1st and 2nd types, • the necessity to have couple and independent date, missings make the classification impossible, • linear dependency between the value of the ratio and the financial status of the entity, although in reality it is usually non-linear, • the lack of direct possibility to determine the probability of bankruptcy (if it is used to build classification models)

Logistic regression analysis

Logit models are very popular method applicable nowdays to forecast bankruptcy cases of enterprises. The logit function in binomial models assumes the form:

$$P(Y=1)=\frac{1}{1+\exp^{-(\beta_0+\beta_1 x_1+\ldots+\beta_k x_k)}}$$

where:

P(Y=1) – dependent-variable, usually determines the probability of bankruptcy

β_0 – intercept

β_i, i = 1, 2, …, k – weights

x_i, i = 1, 2, …, k – explanatory variables – financial ratios.

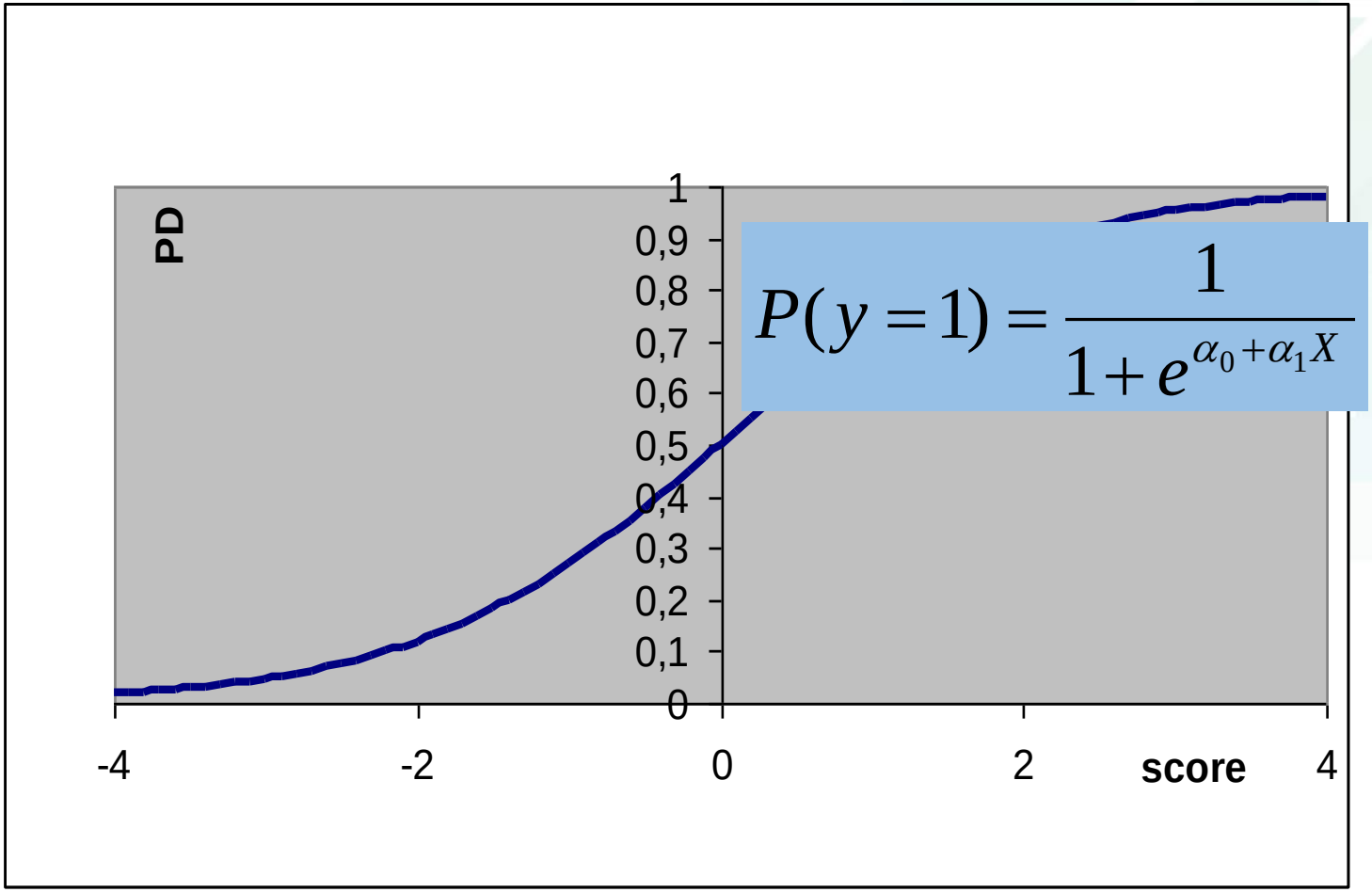
P(Y=1) assumes the value from <0;1>, where 0 – „good" enterprise 1 „bad" enterprise

A significant issue while estimating the binomial models is do properly determine „cut off" pint. In the case of models estimated on the basis of the balanced sample the value of this point is usually equal to 0.5. the structure of the group has the influence on the value of this point (the case of good and bad enterprises)

The so called odds ratio plays the big role in interpreting the results obtained from logit analysis. This ratio is calculated as the relation of the possibility that the event happens to the probability that it will not happen.

The logit model requires that many assumptions are met. The most important are: random character of the sample, big sample, no collinearities of variable, independency of observation

Logistic regression model



$$P(y=1)=\frac{1}{1+e^{\alpha_0+\alpha_1 X}}$$

Advantages and disadvantages of logistic regression

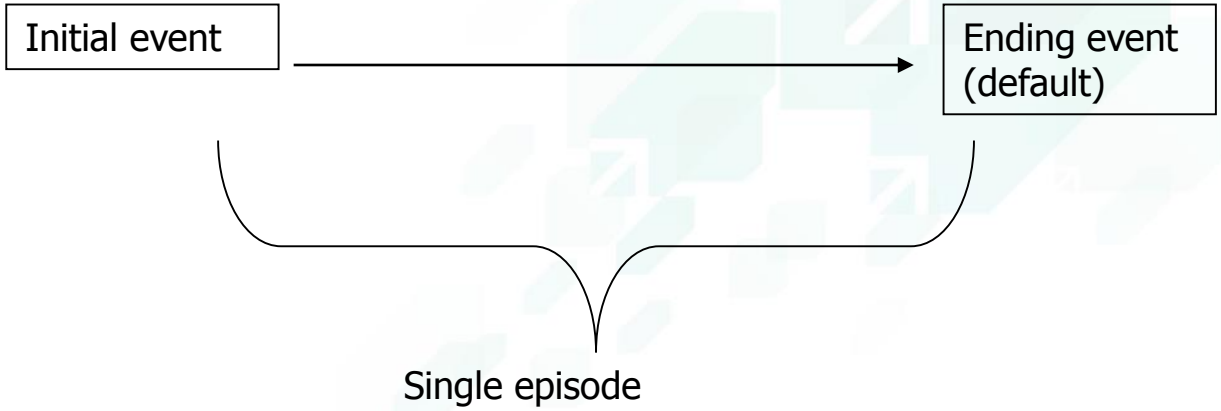
ADVANTAGES	DISADVANTAGES
• no assumptions of normality of distributions of explanatory variables, • no assumption of equality of the matrix of variance – covariance groups, • the result from <0;1> received on forms of the probability of occurence on analysed event, • easy to interpret and easy to understand, • taking into consideration at the same time many variables thanks to the application of weights, • high accuracy of classification comparing to other methods also higher than in the case of linear multivariable discriminant analysis, • possibility to determine the influence of the dependent variable (though in a linear scope), • explanatory variables can be nominal variables, • availability of this method in many statistical programs.	• if the assumption of normality is met, higher accuracy of the classification is obtained by means of linear multivariable discriminant method, • high volatility of the model in relation to correlation of explanatory variables, • high volatility of the model in relation to significant deviations of the explanatory variables distributions from the normal distribution, • worse results of the classification than in the case of artificial neural networks, • necessity to have complete data, missings make the classification impossible, • good results of the classifications are obtained for large samples with high share of bad companies, which is hard to obtain.

Event history analysis – Cox regression model

- **EHA – survival analysis** is described as the set of statistical techniques aiming at the description and studies of the life cycle of an individual, i.e. the frequency of some events, their sequence, distribution, the time the individual spends in different states
- Due to the number of events that may happen we distinguish the single episode analysis and multiple episode analysis, whereas the single episode analysis is the most basic model of the event history tracking analysis
- the stochastic process, which is the subject matter of the analysis, is considered in three basic areas:
  - Time that the occurrence of the distinguished states (events) is expected,
  - The intensity of transitions between distinguished states,
  - Number and sequence of events.
- The basic analytical structure in the event history analysis is the state space and the time axis. The state space is discrete and the time measurement can be continued or discrete. The time axis itself can be defined in two ways: as a calendar time or as a relative time. The state of entry is common for all individuals of the population studied which is defined by common experience by all individuals at the moment of T_0 of an event (it is then a cohort), the occurrence of the bankruptcy event is eliminating an enterprise from a cohort of active enterprises. It is then a final event (exit state) for a single episode model.

Hazard rate model

- Interval dependent variable – Y – time to event (days or months)
- Model of intensities, the value may exceed the range [0-1]



Event history analysis

- Methods of estimation distinguish between parametric and non-parametric methods. This is based on assumptions about the functional form of time distribution. If there are no such assumptions, non-parametrical methods are applied with the classical example of life table models. Non-parametric analysis gives information about changes of individual behaviours schemes in time.
- In parametric approach the time between events is assumed to be a random variable with specific distribution. The most frequently used distributions are: exponential, Weibull, Gompertz. In parametric analysis regression methods are used including the influence of time on hazard rate and the inclusion of explanatory variables and heterogeneity of the population.
- The combination of two approaches is named semi-parametric approach (Cox regression model). The parametric component is based on specified influence of explanatory variables on the hazard rate, but non-parametric component does not specify functional distribution of the time.
- Censoring is very characteristic for event history data. If information is not available then it is censored. The most typical is right censoring when the time till event is not known but it is longer than observation period.

Event history analysis

Take the interval variable T as the time till event occurrence since the time t_0. Distribution of variable *T* may be described in a few different ways apart from density and cumulative function also by survival and hazard functions.

- **survival function**

$$S(t) = P(T > t)$$

where *S*(*t*) means unconditional probability that event occurs after time *t*, so the enterprise will survive at least till time *t.* This function describes the survival pattern in the population.

- **hazard function**

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t \mid T \ge t)}{\Delta t}$$

where *h*(*t*) is conditional density of time to event occurrence (on condition that the event did not occur till time *t*), so *h*(*t*)Δ*t* means (approximately) probability that the event occurs in a very short period of time (*t*, *t*+Δ*t*), on condition that the individual survived at least till time *t*.

Event history analysis – Cox regression model

The most frequently used model is semi-parametric proportional hazards Cox regression model. For Cox regression model the hazard function is given by:

$$h(t \mid x_1,\ldots,x_k) = h_0(t)\exp(\alpha_1 x_1 + \ldots + \alpha_k x_k)$$

where:

$h_0(t)$ means base hazard, parametrically non-specified function of time and $X_1, X_2, \ldots X_k$- means explanatory variables (including time dependent variables).

The main advantage of the Cox model is assessment of variables influence on the process without necessity of base hazard $h_0(t)$ specification. The main disadvantage of Cox model is hazard proportionality assumption. This assumption imposes that for each pair of individuals in any time the hazard rate is fixed. This problem may be solved by including additional time dependent variables. For checking the proportionality assumption, the easy way is to include the interaction with time. The significance of these parameters confirms that the proportionality assumption is violated. In this case the model is named non-proportional hazards Cox regression model. Results of Cox model estimation are parameters describing the influence of explanatory variables on the probability of event occurrence and on the base hazard.

Advantages and disadvantages of event history analysis – Cox regression model

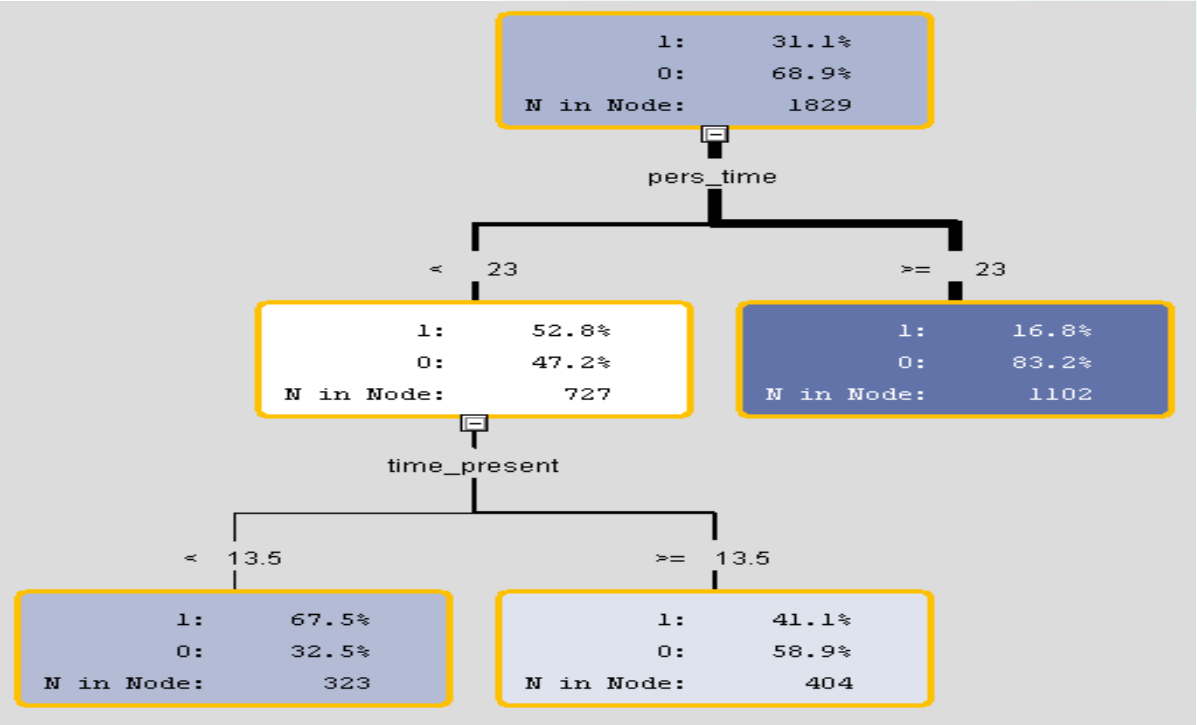
ADVANTAGES	DISADVANTAGES
• apart from the question about „if" we ask the question about „when", • includes censored information about the customer, • no need of fixed time observation period for default observation, • „dynamic" prediction of probability of the event • possibility of including the macroeconomic changes in the model (time varying variables).	• proportionality assumptions, • all assumptions used in regression models: normality assumption, noncollinearity assumption, etc. • non-resistant to missing data, • necessity of the information about the exact time of the event (default)

DATA MINING – decision trees

- A decision tree is a non-parametric predictive method. Observations are classified by assigning them into groups; therefore it determines that the probability of event occurrence is being calculated at the group level. Classification tree can be treated as a segmentation model with supervision (dependent variable).
- A classification is a recursive partition algorithm splitting the group into two subgroups (in each successive step), to ensure uniformity of observations in these subgroups. This model does not require the earlier selection of variables.
- In the preliminary analysis a large set of observations is required when building a tree as well as a relevant number of cases of variable Y (i.e. the number of events).
- Possible unusual observation may distort the results. The main danger when using the decision tree models is the tendency to over-fitting which makes that the resulting model is unstable. This instability means that the classification rules and the estimated probability of an event do not work on the independent data.

Desision tree

- Requires:
  - ✓ A big sample
  - ✓ A lot of defaults (Y)
- Outliers may distort the results
- Overfitting → unstable model



Decision tree

- A decision tree contains the so called root, i.e. the main element, including the entire data set, nodes and sub-segments formed by splitting the data according to the used rules. A tree branch of creates the node with further sub segments. The final division element is called a leaf which is the final segment, which is no longer splitted. Each observation of the output file is being assigned only to one end leaf. A classic decision tree model, for a binary dependent variable, contains the following items (all items are estimated on the training set):
  - nodes definitions, or the principles of assigning each observation to the output to the final leaf,
  - probability (posterior) for each end leaf (ratio of modelled occurrences of a binary variable in each final leaf),
  - assigning level of the dependent variable in the model for each final leaf.
- Decision rule can be based on maximizing the profits, minimizing the costs or minimizing the misclassification error.
- Decision tree, unlike binary logistic regression, does not contain any equations or coefficients, and is based only on the data set allocation rules. The rules generated by the model can be used in prediction without the dependent variable (the result is a binary decision).

Decision tree

- The basic ways of measuring the quality of distribution for dependent binary or discrete variables with few categories are as follows: the degree of separation achieved by the division (measured by p-value Pearson's chi-square test) or the degree of pollution reduction achieved by the separation (measured by the entropy reduction or Gini coefficient).
- Stopping splitting criterion can be as follows: the level of significance p-value divisions, leaf size (minimum size of the final leaf), size distribution (node size), or the maximum size of the path splits (maximum path length).
- After building the decision tree model with the selected method, the next step is cutting the tree to the correct size. It is done in stages. Firstly, one division is cut off, then all possible combinations of the trees are checked and the best of them are chosen. Then another division is cut and the best tree is checked (already shortened twice), etc. With the increase in the number of leaves, the tree value will increase at the beginning but after reaching a certain point, the growth will not be visible, or even a drop can occur. It is the optimal size of a tree.

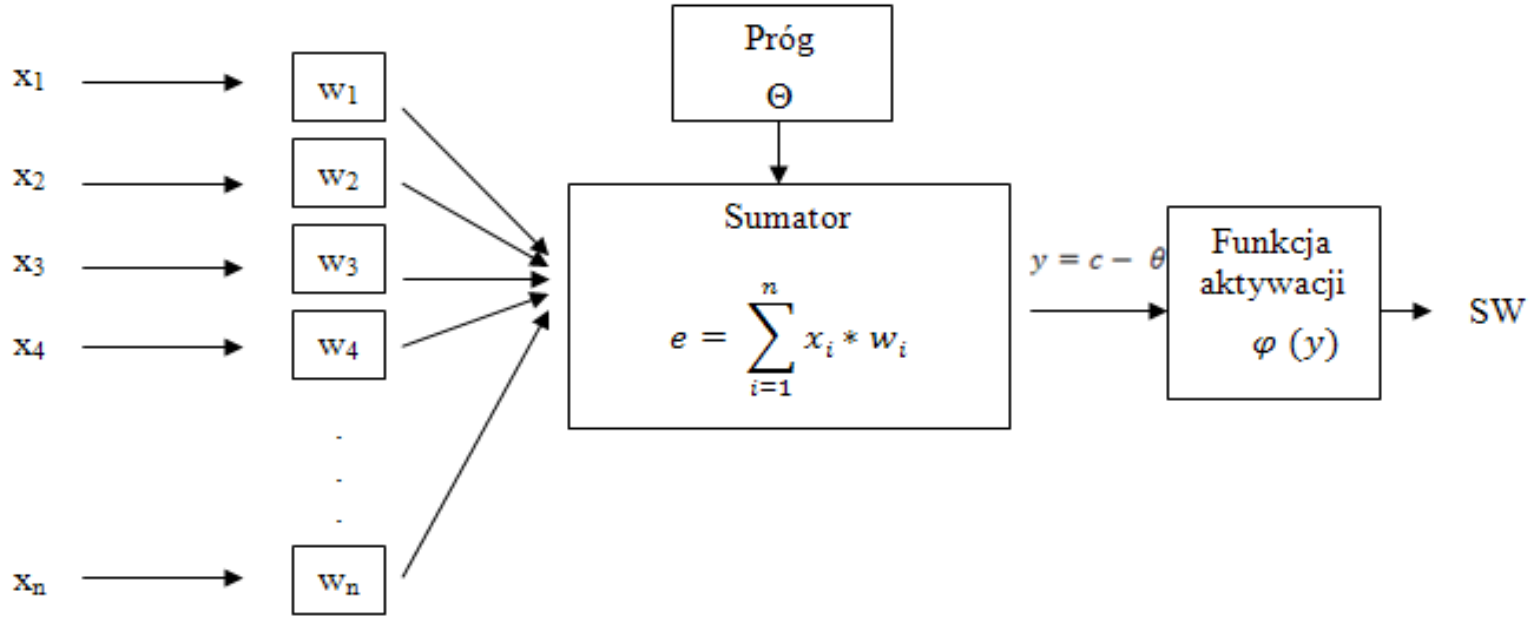
Advantages and disadvantages of decision trees

ADVANTAGES	DISADVANTAGES
• fast adoption of the model to dynamic changes, • easy interpretation and visualization of the data, • resistant to missing data, • no need of normality assumption and no assumption about equal variances between groups, • possible matching with other methods and models, • explanatory variables can be nominal, • possibility of nonlinear dependencies modeling, • automatically selects the significant explanatory variables.	• very often unstable, partitioning may influence the results, • at the one level only one variable is included for splitting, • big training sample is required to stabilize the tree, • the final probability is estimated on the final leaf level (the pooling method), • overfitting risk – very good classification on the training sample, poor on the testing sample.

DATA MINING – neural network

- An artificial neural network is built by neurons (information processing elements) and the connections between them (weights modified during the learning process).
- An artificial neural network is, in fact, a simplified model of the human brain.
- A single artificial neuron has multiple inputs x_n, n=1,2,…,n, and one output. Selected explanatory variables are neuron inputs. Selection of variables is based on the chosen method, such as the factor analysis or principal components method.
- For each variable a specific weight is being assigned– w_n. Once they are determined, the total neuron stimulation e is calculated that is the sum of the products of the explanatory variables and their weights (so called activation function).
- The output value of the neuron depends on the total neuron stimulation, which in turn is obtained by using a suitable activation function $\varphi(y)$. The format of this function determines the type of neuron. For binary endogenous variable the activation function of the output layer is a logistic function, which narrows the range of estimates to [0;1], which makes it possible to interpret the result in terms of the event occurrence probability.

Scheme



B. Prusak, *Nowoczesne metody prognozowania zagrożenia finansowego przedsiębiorstw*, Difin, Warsaw 2005, p. 57.

Neural network

- The capacity of any single artificial neuron is small due to the small computational capabilities and the ability to store a small amount of information. For this reason, the artificial neural networks, consisting of a large number of interconnected neurons, are widely used.
- We can distinguish the following neural networks:
  - two layer - consisting of input and output layers
  - multilayer - consisting of input and output layers, and hidden layers between them.

- The most frequently used neural network is called MLP – *Multi Layer Perception* with one hidden layer.
- One hidden layer is sufficient for modeling all nonlinear dependencies but interval type only.
- Steps of neural network set up require: dependent variable specification, independent variables specification, partitioning, selection of architecture, training, testing.

Advantages and disadvantages of neural networks

ADVANTAGES	DISADVANTAGES
• possibility of fast adoption to dynamic changes, • information may be chaotic and truncated, • no need of normality assumption and no assumption about equal variances between groups, • parallel information included and used in a model, • universal method of estimation, • possible matching with other methods and models, • explanatory variables can be nominal, • nonlinear dependencies and correlations.	• long training time in case of complicated networks, • possibility of non-convergence of the model, • difficult and complicated way of weights estimation, • explanatory variables must be selected, • overfitting, over trained neural networks– very good classification on training set but very poor classification on independent sample (test sample), • Subjective selection of network architecture and optimization algorithms, • black-box, no interpretation of parameters.

Comparision

- **The logistic regression** is placed between the discriminant analysis and the neural network, considering the implementation difficulty. In the logistic regression the assumptions about explanatory variables are not so strict as in the case of discriminant analysis. However it **requires big samples** for precise estimation and a high quality of classification. It is not resistant to missing values.
- **The linear discriminant analysis** model is adequate for smaller samples, smaller databases were characteristic for early 90s. Higher volumes available now in many Banks may develop their own methods using databases and more advanced statistical techniques such as neural networks, logistic regression, decision trees.
- There are not information or technical limitations nowadays. The dynamic development of advanced models and techniques should be observed in rating models development in Banks. The meaning of **event history analysis** and **data mining** analysis will be growing.

Empirical example

- The sample consisted of 6078 financial statements FS for 2342 Medium and Large Enterprises (above 8 mio turnover) in Poland, for which 760 defaults (bankruptcies) were identified over period of 2 years starting from date of FS.
- Financial Statements covered years 2002-2011, missings (<1% cases) were imputed with mean value. Extreme values for variables (ratios) were replaced with the value of 5 and 95 percentil.
- For a balanced sample only defaults (760) and randomly selected 760 non-defaults were selected. The final sample consisted of 1520 individuals, with proportion of 1:1 good and bad enterprises.
- 20 financial ratios proposed for financial analysis of enterprise by banking analytics were used (Zaleska 2012).
- The list of ratios was selected by two steps procedure:
  - Step1. All correlated ratios were eliminated using Hellwig parametric method (r > 0,7),
  - Step 2. Using univariate discriminant analysis the discriminative power of ratios were estimated and all ratios with AR < 0.2 were eliminated.
- The final list of ratios consisted of 7 ratios.

Ratio			AR
PL_PS	quick ratio	Current assets -stocks / short term liabilities	0,308
EF_ROA	profitability of assets ratio	net profit / total assets	0,357
SB_AKW	Coverage of assets by equity	equity/ total assets	0,478
OD_PODE	debt coverage by ebit	(net profit+ tax) / (interests+ capital payments)	0,224
AK_CRKO	working capital cycle	net working capital/ net income from sales x 360	0,256
AK_UNAO	share of receivables in assets	(long and short term receivables) / total assets	0,241
AK_CRZK	liabilities cycle	average trade liabilities / net income from sales x 365	0,293

Discriminant function

Linear discriminant function for:		
	0	1
Intercept	-5.88616	-5.28586
PL_PS	0.45456	0.17347
EF_ROA	0.25585	-6.06969
SB_AKW	14.38695	11.36781
OD_PODE	-0.00190	-0.00135
AK_CRKO	-0.00632	-0.00560
AK_UNAO	10.77621	12.53975
AK_CRZK	0.02961	0.03245

Logistic regression

Parameter	DF	Estimation	Err	Chi-sqr	P-value
Intercept	1	0.5879	0.2038	8.3192	0.0039
PL_PS	1	-0.5763	0.1241	21.5574	<.0001
EF_ROA	1	-6.3861	0.7927	64.9069	<.0001
SB_AKW	1	-2.6665	0.3474	58.9105	<.0001
OD_PODE	1	-0.00009	0.00134	0.0040	0.9493
AK_CRKO	1	0.00145	0.000849	2.9182	0.0876
AK_UNAO	1	2.1120	0.3563	35.1297	<.0001
AK_CRZK	1	0.00246	0.000945	6.8032	0.0091

Odds ratio			
Effect	OR	Wald CL95%	
PL_PS	0.562	0.441	0.717
EF_ROA	0.002	<0.001	0.008
SB_AKW	0.069	0.035	0.137
OD_PODE	1.000	0.997	1.003
AK_CRKO	1.001	1.000	1.003
AK_UNAO	8.265	4.111	16.617
AK_CRZK	1.002	1.001	1.004

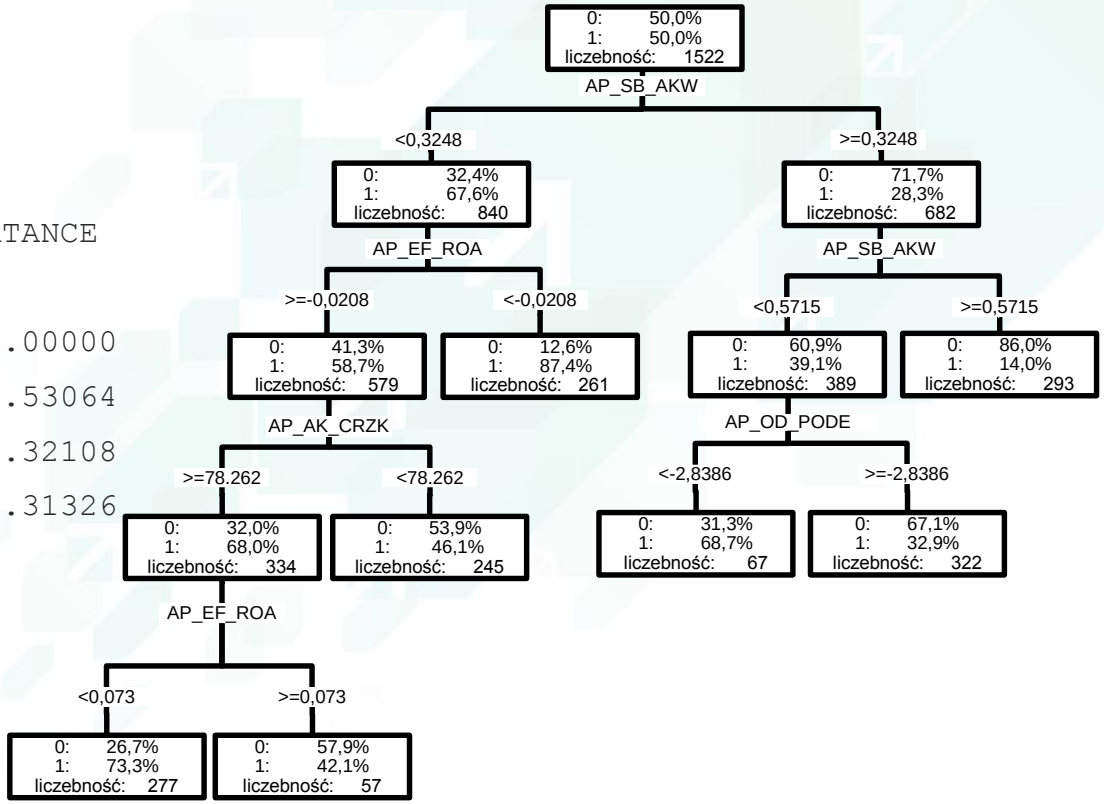
Cox regression survival model

Parameter	DF	Estimation	Err	Chi-sqr	P-value	HR
PL_PS	1	-0.52295	0.09680	29.1859	<.0001	0.593
EF_ROA	1	-4.59334	0.48306	90.4179	<.0001	0.010
SB_AKW	1	-1.69140	0.24218	48.7759	<.0001	0.184
OD_PODE	1	-0.0000975	0.00093	0.0109	0.9169	1.000
AK_CRKO	1	0.00150	0.00051	8.6006	0.0034	1.001
AK_UNAO	1	1.06736	0.21328	25.0442	<.0001	2.908
AK_CRZK	1	0.00109	0.00054	4.1235	0.0423	1.001

Decision tree

- F test with 0.2 significance level used for splitting
- Minimal leaf size 50 units
- Significance of variables:

Obs	NAME	NRULES	IMPORTANCE
1	AP_SB_AKW	2	1.00000
2	AP_EF_ROA	2	0.53064
3	AP_OD_PODE	1	0.32108
4	AP_AK_CRZK	1	0.31326



Neural network

Procedure NEURAL

Optimization results

N	Parameter	Estim	Gradient function target
1	AP_AK_CRKO_H11	-0.100862	0.007724
2	AP_AK_CRZK_H11	-0.074365	0.004461
3	AP_AK_UNAO_H11	0.257176	-0.002191
4	AP_EF_ROA_H11	-0.010791	-0.000592
5	AP_OD_PODE_H11	0.017423	-0.000635
6	AP_PL_PS_H11	-0.063980	-0.000504
7	AP_SB_AKW_H11	-0.260593	-0.000098949
8	AP_AK_CRKO_H12	1.962116	0.004406
9	AP_AK_CRZK_H12	-1.126914	0.003138
10	AP_AK_UNAO_H12	0.208302	-0.001972
11	AP_EF_ROA_H12	-1.698166	-0.000030469
12	AP_OD_PODE_H12	-0.416528	0.000129
13	AP_PL_PS_H12	-0.963936	-0.000226
14	AP_SB_AKW_H12	0.734915	0.000420
15	AP_AK_CRKO_H13	-0.954705	0.000081119
16	AP_AK_CRZK_H13	-1.942790	0.000186
17	AP_AK_UNAO_H13	-0.354209	0.000228
18	AP_EF_ROA_H13	1.441217	-0.000051045
19	AP_OD_PODE_H13	4.470955	-0.000096755
20	AP_PL_PS_H13	2.924606	-0.000225
21	AP_SB_AKW_H13	0.183680	-0.000141
22	BIAS_H11	0.734884	0.001664
23	BIAS_H12	-3.760444	0.001824
24	BIAS_H13	2.045224	-0.000544
25	H11_CZY_D1	3.747302	-0.000085812
26	H12_CZY_D1	1.189661	-0.000237
27	H13_CZY_D1	-1.034454	-0.000500
28	BIAS_CZY_D1	-0.968288	0.000585

Value of final function = 0.51238767

Empirical example - comparison

Model	Accuracy Ratio	I type error	II type error
Logistic Regression	0,602	25,40%	29,40%
Discriminant function	-	**21,68%**	32,59%
COX Regression	0,601	**22,47%**	33,11%
Decision Tree	0,584	37,32%	16,82%
Neural Network	0,646	25,09%	26,94%

Rating model

- A **rating model** is a basic tool in the credit application process – underwriting and in credit risk management system for determining enterprises creditworthiness
- In recent years we observe the speed in development of information system and data warehouses. It makes possible for Banks to develop their own rating models. Most of those models are statistical or hybrid models.
- Changes in Supervisory regulations (CRD IV) caused significant changes in the range of use of methods and techniques of models development.
- At the same time those amendments triggered changes and corrections in qualitative rating and transactions' risk assessment. This caused a lot of issues to be addressed in rating models and systems development.

Rating model development

Financial rating – quantitative part
- financial ratios
- multivariate methods (regression, discriminant, data mining)

Qualitative questionnaire – qualitative part
- as separate model
- as part of common model
- calibration?

- Warning signals
- Additional information
- Additional rules

how many signals?
any knock-out rules?
process/rating

Overrules/Overrides

How to control?
Is it needed?

Group Rating or External Rating
- if exists, +/- rating, consolidated FS

Final rating of the enterprise

Model/system development - PROJECT

FINANCIAL RATING
Experts from risk analysis, definition of ratios,
Statistical methods – data analyst

→

QUALITATIVE RATING
Experts from risk analysis – experts methods, brain storms. Combination with Financial Part – statistician (data analyst)

→

OTHER Elements (big role of expert judgment) :
group rating, transaction assessment, final rating combination, overruling.

RATING MODEL

RATING RULES DESCRIPTION
Methodology procedure – Procedures and Processes

→

MANAGEMNET BOARD – approval for internal usage of models and methods, methodologies

→

IT – implementation in internal information systems, if there is no system – it must be externally provided

Basic elements of a rating model

Goal
Segmentation
Data:
- collection
- quality
- time frame
- default definition
- default lag

Definition of ratios
Univariate analysis
- Transformations (standardization, RDF)
- Discriminative power
- Correlation (selection, factor analysis)

Multivariate analysis
- Different methods (data availability)

Qualitative part:
- List of questions
- Experts (questionnaires)
- Combination with financial part
- Expert judgment

Overruling/higher risk criteria
- Expert judgment (min/max change)
- Control of overrules

Group
- Definition of group
- Expert judgment (min/max change)
- haircut

Assessment of transaction risk

Types of transactions:

Overdrafts:

Object/Activity (focused on goal such as financing contracts, the source of payments)
Entities (current activity financing)

Assessment:
- Duration of transaction,
- Volume of transaction
- Scheme of payments
- Other risks: currency, interest rate.

Investments:

Development (simple way)
Development (standard way)
Replacement investment

Assessment:
- Own contribution
- Duration
- Grace period
- Scheme of payments
- Debt to Income
- Other risks: currency, interest rate.

Final rating assessment

	Transaction			
Rating				

The Third International Workshop

aptak@sgh.waw.pl

Advanced Analytics and Data Science

Partner

Sponsor

Media Partners