“Evaluation of empirical attributes for credit risk forecasting from numerical data”

AUTHORS	Augustinos I. Dimitras Stelios Papadakis Alexandros Garefalakis

ARTICLE INFO	Augustinos I. Dimitras, Stelios Papadakis and Alexandros Garefalakis (2017). Evaluation of empirical attributes for credit risk forecasting from numerical data. *Investment Management and Financial Innovations*, *14*(1), 9-18. doi:10.21511/imfi.14(1).2017.01
DOI	http://dx.doi.org/10.21511/imfi.14(1).2017.01
RELEASED ON	Friday, 31 March 2017

JOURNAL	"Investment Management and Financial Innovations"
ISSN PRINT	1810-4967
ISSN ONLINE	1812-9358
PUBLISHER	LLC “Consulting Publishing Company “Business Perspectives”
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
20	1	3

Augustinos Dimitras (Greece), Stelios Papadakis (Greece), Alexandros Garefalakis (Greece)

Evaluation of empirical attributes for credit risk forecasting from numerical data

Abstract

In this research, the authors proposed a new method to evaluate borrowers' credit risk and quality of financial statements information provided. They use qualitative and quantitative criteria to measure the quality and the reliability of its credit customers. Under this statement, the authors evaluate 35 features that are empirically utilized for forecasting the borrowers' credit behavior of a Greek Bank. These features are initially selected according to universally accepted criteria. A set of historical data was collected and an extensive data analysis is performed by using non parametric models. Our analysis revealed that building simplified model by using only three out of the thirty five initially selected features one can achieve the same or slightly better forecasting accuracy when compared to the one achieved by the model uses all the initial features. Also, experimentally verified claim that universally accepted criteria can't be globally used to achieve optimal results is discussed.

Keywords: credit risk, computational intelligence, management commentary, quantitative and qualitative criteria, Management Commentary Index.
JEL Classification: E5, C63, M41.

Introduction

Banking activity is displayed to miscellaneous hazards. Understanding and evaluating these hazards is urgent for bank administration, and additionally, for the security of the entire economy (Sieczka and Hołyst 2009). Banks have a tendency to loan firms with high credit quality and not to loan low credit quality firms. So the most imperative figure deciding loaning practices is credit risk (Daniels and Ramirez 2008), (Huang et al. 2007).

As indicated by Duff and Einig (2009), research taking into account credit risk has been a standout amongst the most dynamic zones of late economic research, with noteworthy endeavors conveyed to break down the significance, role, and impact of credit ratings. Credit risk examination has pulled in much consideration from budgetary foundations because of the current money related emergencies and administrative worries of Basel II (Basel Committee on Banking Supervision, 2006). Moreover, a business competition for acquiring more market share and benefit turn out to be increasingly forceful as of late, a few establishments go out on a limb to accomplish upper hand in the market. Subsequently, numerous economic establishments experienced an important misfortune of consistent increment of defaults and terrible credits from their partners. In any case, increasingly grown-up population use credit products, for example, mortgages, car and house advance, credit card, and so on, from banks or other money economic foundations. In this manner, a viable credit hazard examination model has been a vital variable for portraying altogether the genuine credit dangers of the chosen bank's advances portfolio.

For the most part, the techniques for client credit risk examination can be essentially seen as two phases. In the first place, when candidates apply for credit, the banks must settle on a choice regardless of whether to allow the credit and the amount to give. The conventional technique for settling on such choices depends on the experience of past loaning choice. Be that as it may, with the expansion of the quantity of candidates and the extraordinary rivalry in the credit business, this customary strategy can't meet the requests of both monetary and productivity viewpoints for economic organizations.

These days, credit scoring is a broadly utilized method that helps the banks to settle on such credit granting choices. Its fundamental thought is to assess the likelihood that how likely the candidate will default, as per the characters recorded in the application frame with a quantitative model in light of data of the past candidates and they acknowledge and dismiss a choice that is made by contrasting the evaluated default likelihood and a legitimate edge.

In the second stage, the moneylenders need to settle on the choices, how to manage the current clients. At which point and how to increment and decrease the clients'credit?

On the off chance that the client begins to fall behind in his reimbursements (i.e. past due obligations), when and what moves ought to be made?

Augustinos Dimitras, Associate Professor, School of Social Science, Hellenic Open University, Patra, Greece.
Stelios Papadakis, Professor, Department of Business Administration, T.E.I. of Crete, Lakonia, Aghios Nikolaos, Crete, Greece.
Alexandros Garefalakis, Adjunct Lecturer, Department of Accounting and Finance, T.E.I. of Crete, Heraklion, Greece.

How the client ought to be treated with respect to hold obligation's practicality? Procedures that assist these choices are called behavior scoring.

The standard of this approach is precisely the same as credit scoring, yet utilizing more data which portrays the client's execution amid some past perception periods. These days, the most vital data to portray the client's performance are gotten from the account of corporate yearly report.

The paper is composed as follows. In Section 2, the structure of key strides in a Bank's credit basic leadership process is portrayed and additionally the best loaning practices to highlight dangers embraced in an advance's assessment stage. In addition, regions of business investigation are set out, to discover clients' economic position and their "specific" operational qualities. In Section 3, the date sample is depicted, with the factors utilized and their particular characteristics. In Section 4, the determination of noteworthy sources of info is set, and in addition the usage of different chosen characteristics approaches made, i.e. Pearson's r, Spearman's ρ, Kendall's τ, and PCA (Song et al., 2010) and multivariate traits assessment. Additionally, discourse and elucidation of the outcomes are displayed. At long last, in Section 5, conclusions are introduced and also additional research.

1. Lending decisions and financial status analysis of firms

1.1. Key steps of lending decision procedure. Borrowers, on a basic model of rating, are divided into two general categories: I) the consistent and II) the borrowers with overdue and problems in repaying their loan obligations. The adoption of this approach, which is consistent with the principles of Basel II, the standardized approach and the internal ratings approach for the measurement of credit risk, places emphasis on the calculation of the expected probability of default (PD) for each of the categories of loans, taking into account customers' historical default data (Pasiouras et al., 2006). The main point, however, for each bank is to find out the right lending decision in the evaluation of credit provided to legal persons. The loan evaluation process is a common process, regardless of the bank organization. However, nowadays, due to the economic crisis, further clarifications of individual characteristics are set in identifying those elements that require further evaluation.

Table 1 describes some basic steps of the lending decision procedure, taking into account best practices in the banking industry worldwide. Each bank initially receives (Step 1) customer's loan application describing the purpose of lending and its characteristics, namely the repayment rate, collaterals provided, etc.

Table 1. Key steps in a bank's loan decision

Steps-procedures	Description	Potential risks
Step 1 Select the appropriate loan product	Choosing an appropriate loan product that meets customer requirements	1. Incorrect loan product that does not cover the actual customer needs. 2. Wrong product pricing and increased probability of default
Step 2 Customer's rating	The bank carries out an initial assessment of the of the borrower's creditworthiness.	Failure to take the real customer's financial status. It results in overestimation or underestimation of the customer's economic capacity
Step 3A Positive lending decision (+)	The bank decides to provide credit to customers, based on the evaluation of its overall financial status	1. Possible failure to take adequate guarantees 2. The bank incorrectly determines the individual loan's characteristics, such as the interest rate, the repayment period, etc.
Step 3B Negative lending decision (-) Reassessment	The bank re-evaluates the second stage due to low credit borrower's quality, inappropriate or little collaterals, etc.	Underestimation of the borrower's real financial status
Step 4A Positive lending decision (+) collaterals' strengthening	Conduct borrower's reassessment in the second stage, with a requirement for collaterals' strengthening	Select loan volume, taking into account the personal and collateral guarantees and not the primary ability to repay the loan
Step 4B Negative judgment (-)	Reject loan's request	Rejection of a good loan's request, which would have been paid without problems. Loss of bank's profits and adverse consequences for the borrower himself

Table 1 highlights the risks undertaken by the bank in the evaluation stage, since the repayment of the credit is purely a quantitative-financial, as well as, a qualitative customer's analysis issue (eg.: solvency, borrower's market position, other loans' quality characteristics, etc.) The final credit decision must be varied in acceptable limits carefully defined per bank for each category of loan product. The described framework decision on a bank loan, which takes into account the best global practice, is for the bank to adopt a flexible strategy that provides adequate and comprehensive information for a customer's credit, minimizing credit risk.

1.2. Customers' financial status analysis. The preparation of information for the decision of credit undertaking is probably the most expensive part of the process of the credit.

Table 2 shows the results of assessing the customer's creditworthiness that determines the level of the credit risk undertaken by the bank. The main issue for the credit decision is to determine the client's financial position.

The use of the standard assessment model, depicted in Table 2, has common features with the systems operating in banks nowadays. For example, to determine the position of a firm, there is a need for using various characteristics (factors), which show its financial position and its "particular" operational characteristics. These factors are grouped in a practical way in nine areas of the baseline analysis (Table 2), i.e. four quantitative and five qualitative characteristics.

Table 2. Areas of business analysis

Number	Quantitative factors	Qualitative factors
1	Profitability	Firm's position in the market (in case of business loans)
2	Financial liquidity	Management Commentary Index (Ma.Co.I)
3	Business activity	Reliability of businessmen
4	Level of sort and long term debt	Nature of business
5	-	Historical customer's credit data

The quantitative factors use only financial data for calculating relevant financial indicators of profitability, liquidity, activity and customer's bank lending. On the other hand, qualitative factor analysis is called assessment of objective factors. These qualities do not show causal relationships; these factors are measured only in a subjective manner, based on solid criteria. The most important qualitative factor is the Management Commentary Index (Ma.Co.I) which measures the disclosure quality of narrative portion of annual reports of firms (Garefalakis et al., 2016).

2. Data and mathematical formulation

A Greek Bank provides loans as products to borrowers. The bank keeps historical records of the behavior of past borrowers (Kosmidou et al., 2007). Each record corresponds to a specific borrower and includes a number of measured attributes of him and the payoff status of the loan he received. These attributes, namely candidate attributes or candidate inputs, are in advanced decided by an expert in the field of application. This decision is based either on expert's experience and/or his intuition and involves the factors affecting the output of the system according to his intuition/experience.

2.1. Sample and core characteristics. We use unique quantitative and qualitative data from 150 firms of a Greek banking institution in retail and corporate banking. The data used include, in particular:

- *The Credit Risk Rate (CRR)* (x_1) that takes values 0 for low, 1 for acceptable, 2 for acceptable with caution, and 3 for high credit risk.
- *The Exposure at Default (EAD)* (x_2) is defined as the gross exposure upon default of an obligor.
- *Probability of default (PD)* (x_3) describing the likelihood of a default over a particular time horizon. It is the likelihood that a borrower will be unable to meet its debt obligations.
- *Expected Loss (EL)* (x_4) is the average credit loss that a bank would expect from an exposure or a portfolio over a given period of time. *Expected Loss (EL)* is estimated as the product of *Exposure at Default (EAD)*, *Probability of default (PD)* and *Loss Given Default (LGD)*. Herein we take a regulatory estimate for LGD equals to 45% under the Basel Committee (2006) requirements.
- *Variables* x_5-x_{24} are financial accounts and quantitative indicators taken out from the borrowers' financial statements.

More specifically:

x_5 *Fixed Assets (FA).*

x_6 *Equity Capital (EQU).*

x_7 *Bank Loans (BLs).*

x_8 *Total Sales (TS).*

x_9 *Cash Deposits (CD).*

x_{10} *Gross Profit Margin (GPM).*

x_{11} *Net Profit Margin (NPM).*

x_{12} *Return on Total Assets (ROA).*

x_{13} *Return on Equity (ROE).*

x_{14} *Asset Turnover Ratio (ATR).*

x_{15} *Sales to Equity Ratio (SER).*

x_{16} *Fixed-Asset Turnover Ratio (FATR).*

x_{17} *Debt-to-Equity Ratio (DER).*

x_{18} *Liquidity Ratio (LR).*

x_{19} *Acid Liquidity Ratio (ALR).*

x_{20} *Inventory Turnover (IT).*

x_{21} *Working Capital Turnover Ratio (WCTR).*

x_{22} *Debt-to-Equity Ratio (DER).*

x_{23} *Debt Ratio (DR).*

x_{24} *Debt to Capital Ratio (DCR).*

- *Ma.Co.I Narrative Quality* (x_{25}), a binary variable that takes the value 1 if the quality of narrative part of corporate annual reports is more than 50% and "0" if the quality of narrative part of corporate annual reports is less than 50% (Garefalakis et al., 2016).
- *Customer's Characterization* (x_{26}), a variable that takes discrete values 1 for very credible, 2 for credible, 3 for satisfactory, 4 for adequate, and 5 for inadequate customers[1]. Customer's Characterization *(x26)*, a variable that takes discrete values of:

1. for absolute credible borrowers, with no over due ever listed in their records,
2. for credible borrowers, with no over due ever listed in their records, but with no prior cooperation with the bank,
3. for satisfactory, with minor overdue in payments listed in their records, less than 30 days,
4. for adequate, with overdue debt over 30 until 89 days in payments listed in their records,
5. for inadequate borrowers, with overdue listed for at least 91 days and more.

- *The weighted average of annual interest rate* (x_{27}) is the average bank's interest rate spread for the last 3 year period of lending bank inside information.
- *Collaterals* (x_{28}) taken for loans' guarantees in euros.
- *The Loan to Value Ratio (LTV)* (x_{29}) is used by banks to express the ratio of a loan to the value of an asset purchased.
- *Obligor type* (x_{30}) that takes the value of 0 for retail, and 1 for corporate borrower.
- *Collateral type* (x_{31}) that takes the values of 1 for urban property, 2 for commercial property, 3 for other types of property, and 4 for none.
- *Loans' Maturity* (x_{32}) that takes the values of 0 for short-term, 1 for mixed term and 2 for long term debt.
- *Firms' Exporting Activity* (x_{33}) that takes the value 0 for domestic activity and 1 for domestic and exported activity.
- *Firms' Sectors* (x_{34}) that take the values 1 for agricultural, 2 for constructions, 3 for hotels and tourism, 4 for manufacture, 5 for other industries, 6 for retail commerce, 7 for services, and 8 for Wholesale.
- *Firm's Legal Form* (x_{35}) that takes the values 1 for sole proprietorship, 2 for cooperative, 3 for municipal public companies, 4 for private limited company, 5 for limited partnership, 6 for limited liability, and 7 for unlimited companies.

The sample consists of 150 borrowers, solely firms that keep "C" category accounting books (in the Greek National Accounting System) and categorized based on Basel II criteria, as exposures to retail or corporate banking; in that case, such loans are covered by mortgages on commercial or residential property and cash collaterals. The CRR is the result of calculating the degree of credit risk using a commercial software and takes values between 0.07 (AA +) and 1 (CCC) of Standard & Poor's rating scale. These values correspond to the respective levels of credit risk assessment classification ranging from "Minimum" to "High" credit risk, respectively. Regarding the used financial ratios and variables taken into account the average values as derived from firms' published financial statements for the three consecutive years 2009-2011.

Regarding the qualitative customer data, exported from the total "financial" status of the borrower. We enter the value 1 for the category "Satisfactory quality data" when all firms' qualitative data are positively evaluated (e.g., quality of cooperation, professionalism, successors continuity, market position, etc.) and account for more than 60 % of the total information taken and we put the value 0 where customer's quality data range is less than 60 % of the total information.

Also, regarding the due borrowers, we take into account debts during the last year (2011) of the study in relation to the reference period (01/01/2009 -12/31/2011). We enter a value of 1 for non-existence of a customer's overdue debt, as well as other information taken from the Tiresias system, the main Greek Default Financial Obligations & Mortgages and Prenotations to Mortgages System that contains data concerning bounced checks, unpaid bills of exchange, mortgages and presentations to mortgages. In contrast, we enter 0, where borrowers show due debt greater than 90 days or unfavorable data in Tiresias system.

Finally, with respect to the variable associated with the 'maturity' in firms' loans, we take into account the repayment period and a customers' separation in those with short-term, mixed-term and long-term lending. Thus, loans with annual recycling capital and interest (e.g., credit limits using overdrafts) are accounted for as short-term lending, and we enter the value 0. Where borrowers receive long-term funding (usually more than two years), we enter the value 2 and in intermediate cases, where borrowers have and short- and long-term debt, we enter the value 1. Long-term loans are considered as the highest ones in terms of credit risk by the bank due to the time horizon of repayment that may be changed significantly during the years to come.

[1] Bank inside information.

The usefulness of economic fundamentals and financial indicators of a company as explanatory variables for the assessment of credit risk has been shown in various studies, including Benos and Papanastasopoulos (2007), Doumpos and Zopounidis (2001), Fernandes (2005) among others. The choice of these variables was based on this literature and the validity of those financial indicators. The methodological framework for the variables draws evidence from both the hybrid creditworthiness model of Benos and Papanastasopoulos (2007) and from key characteristics of Risk Calc and KMV EDF Risk Calc (v.3,1) software.

2.2. Mathematical formulation. The bank aims to predict the payoff status of a future borrower, provided that his attributes are given. This problem can be formulated as a classification problem. Each past borrower is an *instance* consisting of an input vector of his attributes and a label $\{\pm 1\}$, which denotes that the borrower was trustworthy $(+1)$ or not (-1). A function that assigns a label to an input vector can be constructed from the given database by computational intelligence techniques (Pasiouras and Tanna, 2010). From that point of view, the prediction of the behavior of a new customer can be formulated as a typical classification problem according to the following mathematical formulation.

Let $X \subseteq R^m$ be the set of ordered vectors $x_i = [x_{i,1}, x_{i,2}, ..., x_{i,m}]\, i = 1,2,...,n$. Each vector corresponds to a particular borrower, encoding his measured attributes as real numbers. If n borrowers are available and m attributes per borrower are recorded, then, the dimensions of each vector equal m and the cardinality of X equals n. Let $L = \{\pm 1\}$ the set of labels that encodes the behavior of the customers. We assume that a label $+1$ denotes a trustworthy borrower, which, in turn, means that the borrower fully repaid the entire amount, timely. Controversially, a label (-1) denotes that the respective borrower was inconsistent in repaying the loan. The modeling process aims to build a function $f : X \to L$, which assigns a label $l \in \{\pm 1\}$ to a given input datum $x \in X$. In matrix notation, a two dimensional $n \times m$ matrix X stores the attribute values of historic customers, while an $n \times 1$ matrix L stores the respective labels. The whole historic data set can be stored in a matrix $D = [X \mid L]$ consisting of the attributes matrix X augmented by L. Each row of D corresponds to a specific borrower encoding both his attributes and his behavior. In terms of computational intelligence, a model considered as *black-box* can be used to implement the function f as *its transfer function*. The model accepts an input datum x_i and produces an output value $\tilde{y}_i \in \{\pm 1\}$.

The aim of the modeling is $\tilde{y}_i = \ell_i \in \{\pm 1\} \forall i$. In that respect the model reliably identifies the input-output relation of the given dataset. As a result, we can make the assumption that the model can be used to predict the label (output value) of a new *unseen* datum (input vector) correctly.

3. Attribute evaluation

3.1. Data pre-processing. A meaningful preprocessing step is the normalization of the given dataset. The normalization ensures that the contribution of each input to the computation of metrics is irrespective of its actual range; it might be useful both in the data analysis and the modeling performance (Sola and Sevilla, 1997). In general, data normalization is an affine transformation (a linear combination plus a constant term) of each attribute value $x_{i,j}$ from its actual range $[x_j^+, x_j^-]$ into [0, 1].

If the domain of normalization is the range [0, 1], then, the transformation can be computed by Eq. (1).

$$\hat{x}_{i,j} = \frac{x_{i,j} - x_j^-}{x_j^+ - x_j^-} \in [0,1]. \quad (1)$$

In order to simplify the mathematical notation, we will use the symbol $x_{i,j}$ instead of $\hat{x}_{i,j}$ to denote the normalized value of the respective attribute for the rest of the paper.

After normalizing the data, the evaluation of candidate attributes, which will be used as inputs to the model, follows. Since the initial set of attributes (i.e., candidate inputs) is intuitively/empirically selected by a human expert, the possibility of wrong decisions always exists. A human expert may select inputs, which are either redundant or mutually dependent on one another.

3.2. Feature evaluation approaches. The input selection task involves the selection of those of the candidate inputs, which significantly affect the output of the system. The selection of *significant inputs* is based on the collected observational data and it is usually carried out by statistical processing (filter based input selection) or by non-parametric models employed as *wrappers*. Moreover, third category of input selection approaches is the *embedded approaches*, where the identification of significant inputs follows from the model's construction process. It is empirically recognized (Hall and Smith, 1998; Kohavi and John, 1997) that the wrapper based methods provide more exact solutions than the filter based ones. However, wrapper based methods are model dependent and their results de-

pend on the particular model, applied as the *wrapper*. If a specific feature doesn't significantly affects the performance of the selected wrapper, then, this feature is considered as meaningless and is eliminated. However, this assumption implies that the wrapper adequately identifies the dataset, which is not always true. Another drawback of wrapper methods is the lack of interpretation of why a particular feature is rejected or not. This drawback is more intense if the input output relation is non-linear for a particular input. Filter based methods are considered as less accurate, but they have the advantage of providing model independent and easily interpretative results. This aspect is important in financial applications, where the results of any processing should be explanatory.

3.3. Filter based attribute evaluation. *3.3.1. Pearson's correlation coefficients.* A widely used filter based method is the calculation of Pearson's coefficients. A Pearson's coefficient captures the linear correlation between two random variables. Although a Pearson's coefficient is limited to the calculation of the linear correlation between two random variables, it releases the advantage of straightforward and easily interpretative results, even for people who are not experts in the particular domain of application. We compute the Pearson's coefficient r_j for each attribute x_j by Eq. (1):

$$r_j = \frac{\sum_{i=1}^{N}(x_{i,j} - \bar{f}_j)(l_i - \bar{l})}{\sqrt{\sum_{i=1}^{N}(x_{i,j} - \bar{f}_i)^2 \cdot (l_i - \bar{l})^2}}, r_j \in [-1,1] \subseteq R, \quad (1)$$

where $\bar{f}_j$ is the mean value of the j^{th} feature (i.e., the j^{th} column of D); $l_i \in \{\pm 1\}$ is the label of instance i; and $\bar{\ell}$ is the mean value of labels (i.e. the last column of D). A value of r_j around ± 1 denotes a strong linear interdependence between the attribute x_j and the output. Controversially, a value around zero denotes linear independence. The sign of r_j denotes whether the linear relation is ascending or descending, respectively. Since

we are only interested in the magnitude of the dependence, the sign in Eq. (1) can be omitted by taking the absolute value $|r_j|$ of r_j. The set of the candidate attributes includes thirty five attributes (Table 3), intuitively selected by banking experts.

Pearson's approach requires that the data are normally distributed, besides, the assumption of an existing linear relation between the random variables being probed. We use the normality test introduced in Agostino (1971), Bowman and Shenton (1975) to test the null hypothesis: $H_0 = \{$*The sample comes from a normal distribution*$\}$. The $p-\text{value}$, which expresses a two-sided chi squared probability for the hypothesis test, is computed for each attribute (random variable). The p value was less than 0.05 for all attributes except from $x_{32}, x_{27}, x_3, x_{30}$ for which the p - value was 0.96, 0.96, 0.78, 0.07 respectively. The null hypothesis was rejected for most of the attributes and hence, Pearson's approach should not be directly used to decide the significance of attributes.

3.3.2. Spearman's rank correlation coefficients. An alternative non-parametric statistic is the Spearman's ρ rank-ordered correlation statistic (Corder and Foreman, 2009), which measures the monotonic relationship between two random variables. Although, the existence of a monotonic relation between the random variables is an underlined assumption of Spearman's approach, however, this approach is less strict than Pearson's correlation coefficients.

When no duplicate values exist between the random variable x_j and the labels L, then, Spearman's ρ can be computed by the following equation:

$$\rho_j = \frac{6\sum_{i=1}^{n}[R_x(x_{i,j}) - R_L(l_i)]^2}{n(n^2-1)}, \quad (2)$$

where n is the number of instances; j denotes the j^{th} attribute; $R_x(x_{ij})$ denotes the rank of attribute value x_{ij} when sorted in ascending order; and $R_L(\ell_i)$ denotes the rank of ℓ_i when sorted in ascending order after the sorting of x_{ij}. If duplicate values exist, then, the Eq. (1) should be used on the ranked values of $x_{i,j}, \ell_i$, instead of Eq. (2) for computing ρ_j (Corder and Foreman, 2009).

The next issue is to decide a threshold of significance, below which an attribute is rejected as meaningless and above which the attribute is selected as significant. We follow two approaches to decide on the significance of an attribute. The first approach ignores the assumptions of Pearson's and Spearman's approaches and is based on the well-known student's statistical test to check the null hypothesis: $H_0 = \{$*The probed random variables are by chance correlated*$\}$. We check for two levels of significance: for $\alpha = 0.05$ and for $\alpha = 0.01$. Each attribute is evaluated according to its $p-$value, which roughly indicates the probability that an uncorrelated system

produces data sets, which have a correlation at least as extreme as the one computed from the given data sets. The attributes with $p \leq \alpha$ are selected as important ones because the null hypothesis is rejected with high probability for them. Next, the selected attributes are ranked in descending order according to their $|\rho|$ value.

In the second approach, we use the above mentioned statistical methods only as a ranking tool because of their implied assumptions. The second approach exploits the ranking which the statistical tests provide, but employs a non-parametric, non-linear model to evaluate the significant attributes. Tenfold cross-validation of the model is performed as follows: the initial data set is divided into a learning set, including the 90 percent of initial data, and a testing set, including the rest 10%. Next, the learning set is subdivided into the training set, including the 90 percent of the learning set and the validation set including the rest 10% of the learning set (Figure 1).

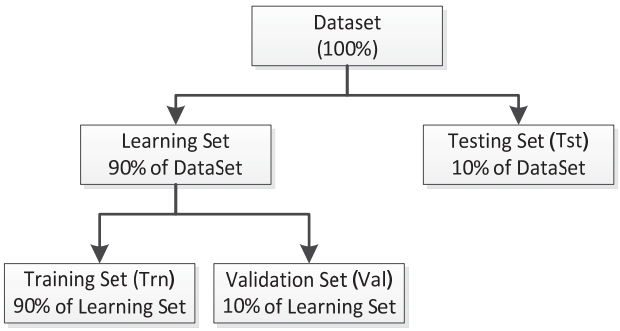

Fig. 1. Tenfold cross-validation

The attributes are sorted in descending order according to their correlation to the output and then they are progressively inserted to a support vector machine model as inputs. A tenfold cross validation is performed on the learning set for each new attribute and the average success classification rate on the validation set is monitored. The subset of attributes, which provides the maximum average success classification rate on the ten validation sets is selected.

Table 3. The selected attributes presented in descending order according to their |r| value

	Selected attributes	Average Tenfold Tst Success Rate
Pearson's r		
0.05	$x_{26}, x_{33}, x_{25}, x_{23}, x_5$	67.33%
0.01	x_{26}, x_{33}	70.00%
Cross-validation	x_{26}, x_{33}, x_{25}	71.33%
Spearman's ρ		
0.05	$x_{26}, x_{33}, x_{25}, x_{23}, x_{22}, x_3$	65.33%
0.01	x_{26}, x_{33}	70.00%
Cross-validation	x_{26}, x_{33}, x_{25}	71.33%
Kendall's τ		
0.05	$x_{26}, x_{33}, x_{25}, x_{23}, x_{22}, x_3$	65.33%
0.01	x_{26}, x_{33}	70.00%
Cross-validation	x_{26}, x_{33}, x_{25}	71.33%
PCA (Song et al., 2010)		
Λ=0.72	$x_{25}, x_{32}, x_{26}, x_{10}, x_{15}, x_{14}, x_{35}, x_{27}, x_8, x_7 \; x_{16}, x_{12}$	68.00%
Cross-validation	$x_{25}, x_{32}, x_{26}, x_{10}, x_{15}, x_{14}, x_{35}, x_{27}$	70.00%

Notes: the selected attributes presented in descending order according to their |r| value when using Pearson's coefficients; Spearman's |ρ| correlation coefficients; and Kendall's |τ|. In every case a student's t-test was performed to check the null hypothesis for significance levels p = 0.05 and p = 0.01. Additionally, a selection based on tenfold cross validation on the learning set was performed. The average tenfold success classification rate on the testing set is given in the last column of every case. It is clear that the cross validation based selection delivered more representative attributes.

Finally, a tenfold cross-validation is performed on the initial data set and the average success classification rate on the testing set is used as the final criterion for the selection. We highlight that the data of the testing set (*Tst*) were neither used in the construction of the svm (Smola and Schölkopf, 1998; Vapnik, 1992; Vapnik, 2000) model nor in the statistical tests, applied.

3.4. Multivariate attribute evaluation. Besides the specific assumptions that filter based approaches require, these approaches have the additional drawback that the attributes are evaluated one by one. It includes the risk of sub-optimal solutions, because one attribute itself may be characterized as non-important when evaluated alone, but it might be important when evaluated with another one jointly. On the contrary, an attribute may be important when evaluated alone, while the same attribute might be not important when evaluated with another one jointly. To this end, multivariate analysis of variance (Grimm and Yarnold, 1995; (Stevens, 2012) was widely used for processing more than one variable simultaneously.

Principal Components Analysis (PCA) (Diamantaras and Kung, 1996; Kung and Diamantaras, 1991) is a variable-reduction technique that shares many similarities to exploratory factor analysis (Thompson, 2004). Its aim is to reduce a larger set of variables-attributes into a smaller set of "artificial variables", called "principal components", which account for most of the variance in the original variables. PCA is mainly used in an exploratory way. If one is interested in reducing the observed variables down to their principal components while maximizing the variance accounted for in the variables by the components, then he should be using PCA.

Factor analysis (Thompson, 2004) is a multivariate method based on Principal Components Analysis (PCA). PCA projects the original input space to a new space of orthogonal variables which approximately conveys the same information as the original one. Factor analysis extracts a subset of specific variables (usually fewer) from the original set. The eigenvalues of the covariance matrix of the original data are computed at first. Next, the eigenvalues are sorted in descending order and are normalized such that their convex sum equals one (100% of total variance). The normalized eigenvalues are named *latent* variables. Consider m eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$, sorted in descending order ($\lambda_i > \lambda_j$ $\forall i > j$). The criterion for selecting the q most important variables is that the cumulative sum of the selected eigenvalues is up to a predefined threshold $\Lambda \in [0,1]$. That is, select the first q eigenvalues such that:

$$\frac{\sum_{i=1}^{q} \lambda_i}{\sum_{j=1}^{m} \lambda_j} \leq \Lambda, \qquad (3)$$

which means that the total variance the selected variables express is up to $\Lambda \cdot 100\%$ of the total variance the original variables explain. Selecting the q most important, eigenvalues which satisfy Eq. (3) does not provides direct information on which of the original values are important. In order to identify the q most important original variables, we examine the absolute values of the coefficients of the respective q eigenvectors as in Song et al. (2010). For $\Lambda = 0.95$, we got $q = 17$. Actually, PCA computes new features by rotating the original axis, thus, transforming the original space to a new orthogonal *feature space*, linearly. If the q most important eigenvectors are stored into an $m \times q$ matrix V, the transformation of the original $n \times m$ input space X is calculated by $\tilde{X} = X \cdot V$, which is an $n \times q$ matrix. Each original datum stored in row X_i in the original $m-$ dimensional space is linearly mapped to row $\tilde{X}_i$ in the new $q-$ dimensional feature space, where $q < m$. We performed tenfold cross-validation by using a svm model on $\tilde{X}$. The model had q inputs, while the label of each datum was preserved.

The average success classification rate on the rotated testing data was 70,66 for $q = 17$, which was the same as the one achieved when included all the m variables. We conclude that although in our problem linear PCA (and subsequently factor analysis) performed significant dimensionality reduction ($\simeq 50\%$), however, it failed to identify the best attributes. It failed both quantitatively, in terms of average success classification rate, and qualitatively in terms of which of the original variables were exactly the most important ones.

3.5. Discussion and interpretation of the results. The average tenfold cross-validation success classification rate and the respective attribute for $\alpha = 0.05$ and $\alpha = 0.01$ are summarized in Table 3. The selection based on the statistical test is computationally more efficient and more intuitive. However, it is based on assumptions that we don't know whether they are satisfied or not. Moreover, the α value affects the final outcome and it is an extra parameter being decided. Cross-validation is more computationally expensive but provides more accurate results since it is independent from the assumptions of the statistical tests.

It is obvious that results are very promising regarding credit analysis perspectives and key variables' selection that remain important for a credit officer's thorough decision whether one can proceed to customer's lending or not. First of all, we see that *Pearson's r* is surprisingly good, near-

ly optimal results when compared to cross-validation, despite the fact that most of the attributes, considered as random variables, do not follow the normal distribution.

In the optimum set of results, three out of thirty five attributes are selected in cross-validation option with average Tenfold Tst Success Rate at 71.33 percent. This is a very constructive result, where one can get. Only three out of thirty five attributes are enough for taking out the most influential information needed for banking authorities in order to take proper lending decisions.

Another core conclusion is that alpha parameter on statistical methods is difficult to be decided; generally, a value at 0.05 provides more attributes outcomes than required in any selected attribute of the research, while *a* at 0.01 delivers less attributes than actually required. This issue is solved by the use of cross validation estimator performance. Furthermore, all statistical methods appeared more effective and robust than the simplest of the true eigenvector-based multivariate analyses (PCA). Moreover, all statistical methods detect significant attributes, as it was verified by cross validation; also all methods failed to detect the optimal attribute set. Generally the PCA method provides more attributes (nine attributes for PCA-(Song et al., 2010)) options for cross validation analysis) and fewer representatives.

The average tenfold success rate for all attributes lies at 68.66 percent. Optimal subset of attributes delivered the most accurate results in terms of generalization are in *Pearson's r* and *Spearman's* ρ options for the x_{26}, x_{33} and x_{25} attributes. It is in any terms visible that qualitative attributes are those that express best sample's credit quality and provide maximum success rate in any case. This is very close to what Greek banking market experts' support that except from borrowers' core financial positions, factors such as quality of cooperation, as well as good credit history records with the bank are essential for credit quality's assessment. Also, borrowers (firms) with exporting activity, apparently, tend to acquire higher credibility rates than those with solely domestic activity.

Conclusion

Traditional practices rely too much on credit quality indicators such as delinquency, nonaccrual, and risk rating trends. Banks have found that these indicators do not always provide sufficient information for a borrower's credit quality. Both collateral and capital can act as a form of credit risk mitigation, especially in credit forms for both retail and corporate borrowers. The exchange of collateral is a key risk mitigation technique that provides core elements in credit lines given in almost any bank.

On the other hand, one of the main criticisms to be made of the up-to-date credit risk management practices is that these techniques include very limited use of specific kind of information taken by an overall borrower's assessment that takes into account the qualitative information about the counterparties. The undervaluation of the qualitative information's importance in the existing credit risk models, based largely on quantitative inputs, such as financial ratios and relative analysis made, will undoubtedly have to be reconsidered in the near future. Qualitative criteria are essential for the credit quality assessment. From this perspective, the most notable contribution of this study is the inclusion of qualitative information to credit risk modeling.

This paper investigates the determinants of a variety of financial and non-financial factors contributing in almost any credit decision. A micro-analysis is made taking under consideration a loan portfolio with reference to Greek firms. In this part of the research, we identify core elements, both quantitatively and qualitatively, that play major part in taking good lending decisions within banking institutions. Using several computational intelligence techniques in a data set from a Greek bank, we find very thorough results for the bank, management towards mitigating credit risk in loans portfolios.

More specifically, it is revealed that building a simplified model by using appropriate information out of several criteria, we find that only 3 out of the 35 initially selected features one can achieve are enough for through lending decisions. These criteria can produce the same or slightly better forecasting accuracy when compared to the forecasting accuracy achieved by a model, which uses all the 35 features (qualitative and quantitative ones). The main contribution of this study to the literature is the consideration of only two firms' qualitative attributes (i.e., 1. Customer's Characterization and 2. Ma.Co.I Narrative Quality), as well as firms' ability in developing exporting activity; these three separated criteria tend to become significantly conclusive for credit decisions and in any case, they can provide adequate information for credit officers to mitigate bank's credit risk.

From the experimental results, we observe that many of the intuitively selected attributes are redundant, while the generalization performance of many classifiers by using the selected attributes is rather poor. This observation leads us also to conclude that the initially selected attributes can be further enriched so that the decision on the behavior of a feature borrower should be also based upon other representative attributes, as well that take into account more custom made and focused customers'

characteristics, corresponding to focused entrepreneurial environment for the selected borrowers. Furthermore, a richer set of training instances might lead to more accurate results.

In any terms, it is generally acceptable that there is no global credit quality system that fits for all cases nearly for every loan's portfolio selected. This is another major conclusion that may be enriched in future research towards the understanding of a better and more conclusive segmentation of banks' loan portfolios, based on certain and robust banking and market oriented features.

References

1. D' Agostino, R. B. (1971). An omnibus test of normality for moderate and large size samples. *Biometrika*, *58*, 341-348.
2. Basel Committee on Banking Supervision. (2006). *Observed range of practice in key elements of Advanced Measurement Approaches (AMA)*.
3. Benos, A., and Papanastasopoulos, G. (2007). Extending the Merton model: A hybrid approach to assessing credit quality. *Mathematical and computer modelling*, *46*, 47-68.
4. Bowman, K. O., and Shenton, L. R. (1975). Omnibus test contours for departures from normality based on b1 and b2, *Biometrika*, *62*, 243-250.
5. Corder, G. W., and Foreman, D. I. (2009). *Nonparametric statistics for non-statisticians: a step-by-step approach*. John Wiley and Sons.
6. Daniels, K., and Ramirez, G. (2008). Information, credit risk, lender specialization and loan pricing: Evidence from the DIP financing market. *Journal of Financial Services Research*, *34*, 35-59.
7. Diamantaras, K. I., and Kung, S. Y. (1996). *Principal component neural networks: theory and applications*, John Wiley and Sons, Inc.
8. Doumpos, M., and Zopounidis, C. (2001). Assessing financial risks using a multicriteria sorting procedure: the case of country risk assessment. *Omega*, *29*, 97-109.
9. Duff, A., and Einig, S. (2009). Credit ratings quality: The perceptions of market participants and other interested parties. *The British Accounting Review*, *41*, pp. 141-53.
10. Fernandes, J. (2005). Corporate credit risk modeling: Quantitative rating system and probability of default estimation. *Available at SSRN 722941*.
11. Garefalakis, A. Dimitras, A. Floros, C., Lemonakis, C. (2016). How Narrative Reporting changed the Business World: Providing a new measurement tool. *Corporate Ownership and Control*, *13*, 317-334.
12. Grimm, L. G., and Yarnold, P. R. (1995) *Reading and understanding multivariate statistics.* American Psychological Association.
13. Hall, M. A., and Smith, L. A. (1998). *Practical feature subset selection for machine learning.*
14. Huang, C. L., Chen, M. C., and Wang, C.J. (2007). Credit scoring with a data mining approach based on support vector machines. *Expert systems with applications*, *33*(4), 847-56.
15. Kohavi, R., and John, G.H. (1997). Wrappers for feature subset selection. *Artificial intelligence*, *97*, 273-324.
16. Kosmidou, K., Pasiouras, F., Tsaklanganos, A. (2007). Domestic and multinational determinants of foreign bank profits: The case of Greek banks operating abroad. *Journal of Multinational Financial Management*, *17*, 1-15.
17. Kung, S. Y., and Diamantaras, K. I. (1991). Neural networks for extracting unsymmetric principal components. *Neural Networks for Signal Processing [1991], Proceedings of the 1991 IEEE Workshop* (pp. 50-59). IEEE.
18. Pasiouras, F., Gaganis, C., Zopounidis, C. (2006). The impact of bank regulations, supervision, market structure and bank characteristics on individual bank ratings: a cross-country analysis. *Review of Quantitative Finance and Accounting*, *27*(4), 403-438.
19. Pasiouras, F., Tanna, S. (2010). *The Prediction of bank acquisition targets with discriminant and logit analyses: methodological issues and empirical evidence Research in International Business and Finance*, *24*, 39-61.
20. Sieczka, P., and Hołyst, J. A. (2009). Collective firm bankruptcies and phase transition in rating dynamics. *The European Physical Journal B-Condensed Matter and Complex Systems*, *71*(4), 461-466.
21. Smola, A. J., and Schölkopf, B. (1998). *Learning with kernels*. Citeseer.
22. Sola, J., and Sevilla, J. (1997). Importance of input data normalization for the application of neural networks to complex industrial problems. *Nuclear Science, IEEE Transactions on*, *44*, 1464-1468.
23. Song, F., Guo, Z., and Mei, D. (2010). *Feature selection using principal component analysis*, *1*, 27-30. Presented at the System Science, Engineering Design and Manufacturing Informatization (ICSEM), 2010 International Conference on, IEEE.
24. Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. Routledge.
25. Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* American Psychological Association.
26. Vapnik, V. (1992). Principles of risk minimization for learning theory. *Advances in neural information processing systems*, 831-838.
27. Vapnik, V. (2000). *The nature of statistical learning theory*. Springer Science and Business Media.