Available online at www.sciencedirect.com
SCIENCE DIRECT®
Decision Support Systems xx (2005) xxx–xxx
Decision Support Systems
www.elsevier.com/locate/dss

A process model to develop an internal rating system: Sovereign credit ratings

Tony Van Gestel [a,d,*], Bart Baesens [b,*], Peter Van Dijcke [c], Joao Garcia [a], Johan A.K. Suykens [d], Jan Vanthienen [e]

[a] *Credit Risk Modelling, Group Risk Management, Dexia Group, Square Meeus 1, B-1000 Brussel, Belgium*
[b] *School of Management, University of Southampton, Southampton SO17 1BJ, UK*
[c] *Research, Dexia Bank Belgium, Av. Galilei 30, B-1000 Brussel, Belgium*
[d] *K.U.Leuven, Department of Electrical Engineering, ESAT-SCD-SISTA, Kasteelpark Arenberg 10, B-3001 Leuven (Heverlee), Belgium*
[e] *K.U.Leuven, Department of Applied Economic Sciences, Naamsestraat 69, B-3000 Leuven, Belgium*

Received 2 May 2005; received in revised form 5 October 2005; accepted 12 October 2005

Abstract

The Basel II capital accord encourages financial institutions to develop rating systems for assessing the risk of default of their credit portfolios in order to better calculate the minimum regulatory capital needed to cover unexpected losses. In the internal ratings based approach, financial institutions are allowed to build their own models based on collected data. In this paper, a generic process model to develop an advanced internal rating system is presented in the context of country risk analysis of developed and developing countries. In the modelling step, a new, gradual approach is suggested to augment the well-known ordinal logistic regression model with a kernel based learning capability, hereby yielding models which are at the same time both accurate and readable. The estimated models are extensively evaluated and validated taking into account several criteria. Furthermore, it is shown how these models can be transformed into user-friendly and easy to understand scorecards.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Internal rating system; Process model; Support vector machines; Sovereign ratings

1. Introduction

The recently put forward Basel II capital accord provides guidelines for the calculation of the minimum required regulatory capital needed to be set aside to recover from defaulted loans or obligations [4]. One of the key recommendations encourages financial institutions to build rating based risk systems that quantify the default and/or recovery risk of their credit assets. In contrast to the standardized approach, where banks can rely on external ratings, the internal ratings based (IRB) approach catalyzes the development of customized ratings based on collected data and advanced statistical modelling. In this paper, we will present a process model to develop rating models and apply it to design a model for country risk.

The aim of country risk analysis is to identify those countries that will be unable to meet their commitments

* Corresponding authors.
E-mail addresses: VanGestel@dexia.com (T. Van Gestel), B.M.M.Baesens@soton.ac.uk (B. Baesens), Peter.VanDijcke@dexia.be (P. Van Dijcke), Joao.Garcia@dexia.com (J. Garcia), Johan.Suykens@esat.kuleuven.ac.be (J.A.K. Suykens), Jan.Vanthienen@econ.kuleuven.ac.be (J. Vanthienen).

0167-9236/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.dss.2005.10.001

on external debt, i.e. debt owed to non-residents. This is typically tackled by assigning ratings to countries reflecting a country's ability and willingness to service and repay its external financial obligations [8,15]. A strong credit risk rating creates a financially favorable climate whereas a low credit rating usually leads to a

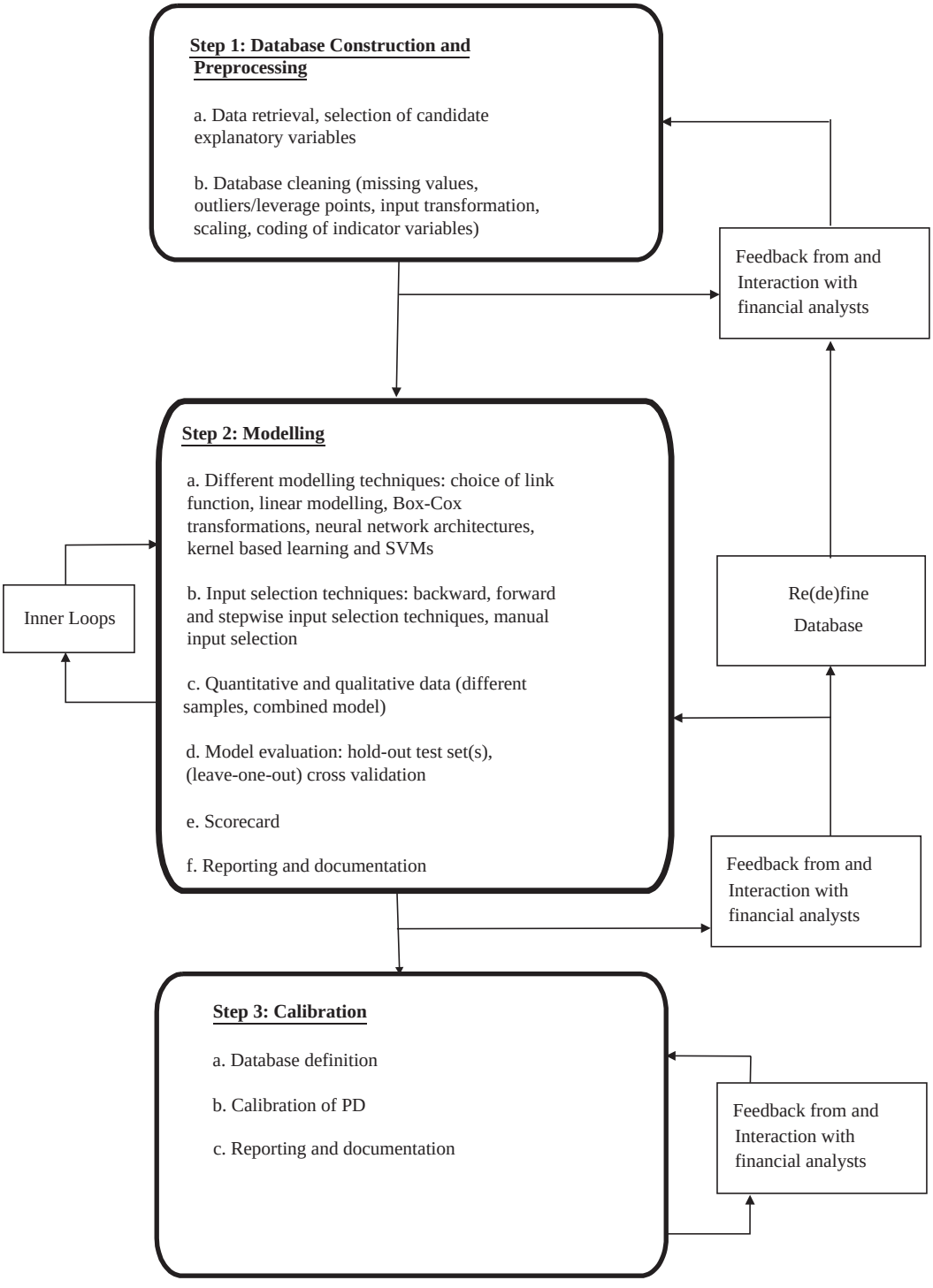


Fig. 1. A process model for developing an internal rating system for mapping to external ratings (or default data) and calibrating it for the internal rating based (advanced) approach. See text for details.

reversal of capital flows and an economic downturn. A good country rating is a key success factor of the availability of international financing since it directly influences the interest rate at which countries can borrow on the international financial market. It may also impact the rating of its banks and companies and is reported to be correlated with national stock returns.

Credit rating agencies have developed models to estimate country risk ratings. The most popular are Moody's Investor Service, Standard & Poor's and Fitch [8]. The external ratings are typically alpha-numerically encoded[1] and are constructed using quantitative economic, social, and political factors and their interactions as well as judgmental aspects and future projections. A drawback impeding the practical use of these external ratings (for the IRB approach) is that most agencies nowadays adopt rating systems that are, for obvious reasons, not disclosed, in the sense that only the output rating is provided and not how it is computed or how the independent/explanatory variables influence the rating. It is the purpose of this paper to build a white-box internal country rating system that is both transparent and easy to understand and that can be applied to both externally and not externally rated countries. For internal reasons, the system will try to mimic the ratings provided by Moody's. The ratings of the different agencies are usually very similar. They are considered as the best measure of a country's credit risk available nowadays as internal default data is missing [4,8].

The system will be built following the process model depicted in Fig. 1. In Step 1, the database with 63 candidate explanatory is constructed and cleaned on which the rating model will be estimated. In Step 2, the rating model is estimated using different regression techniques. An important issue here is the interaction with the financial analysts, e.g., to take into account their experience for selecting the set of explanatory variables that is optimal from both the statistical and the economical perspective. The calibration of the IRB risk system is done in Step 3, fixing the probability of default (PD) in order to calculate the risk weights and regulatory capital.

The rating model is the cornerstone of the IRB approach. Modelling techniques that have been used to assess country risk are, e.g., ordinary least squares regression, logistic regression, decision trees and neural networks [8,10,15,19,25]. In this paper, a stepwise and gradual approach is followed to find a trade-off between simple techniques with excellent readability, but restricted model flexibility and complexity, and advanced techniques with reduced readability but extended flexibility and generalization behavior. First a linear ordinal logistic regression model [17] is estimated, which is the benchmark statistical technique. Next an intrinsically linear model is built by considering univariate nonlinear transformations of the explanatory variables [6,30]. Finally, a kernel based technique called Support Vector Machines (SVMs) is introduced to construct an advanced nonlinear model on top that captures the remaining multivariate nonlinear relations in the data [24,29]. The approach is visualized in Fig. 2, where it is seen that the generalization capacity increases, while the model readability decreases.

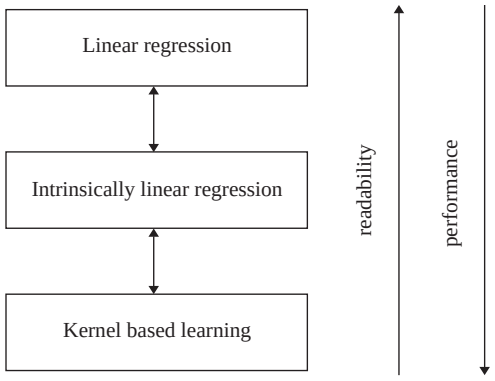


Fig. 2. Gradual combination of linear regression, intrinsically linear regression and kernel based learning (SVMs) with increasing modelling capacity and decreasing readability.

This paper is organized as follows. In Section 2 the modelling techniques[2] are described. The process model is explained in Section 3 and applied to design the country rating model. Conclusions are drawn in Section 4.

2. Combining linear and nonlinear ordinal logistic regression

2.1. Linear ordinal logistic regression

For binary classification problems like bankruptcy prediction, ordinary least squares[3] and logistic regres-

[1] Moody's uses, e.g. Aaa (best credit), Aa1, Aa2, …, C (worst credit before default), while S&P uses AAA, AA+, AA, …, C, respectively.

[2] The reader whose main focus is not on the statistical part may start reading with Section 3.

[3] For binary classification problems, ordinary least squares regression corresponds to Fisher Discriminant Analysis and Canonical Correlation Analysis [1,27].

sion [18] are key techniques to build a discriminant function between two classes: class 1 (defaults) and class 2 (non-defaults). Logistic regression is typically preferred because: its model formulation is specific to a binary classification problem (defaults/non-defaults); it is empirically observed to exhibit better generalization behavior than least squares regression [3,28] and it is known to be more robust to deviations from multivariate Gaussian distributed classes. The ordinal logistic regression (OLR) model [17] is an extension of the binary logistic regression model for ordinal multi-class categorization problems, like e.g., class nr. 1 (very good), class nr. 2 (good), class nr. 3 (medium), class nr. 4 (bad) and class nr. 5 (very bad). Hence, ordinal logistic regression is an interesting technique[4] to model external ratings.

In the cumulative OLR model, the cumulative probability of the rating y is given by:

$$P(y \le i) = 1/(1+\exp(-\theta_i + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n)), \qquad i = 1, \ldots, m, \tag{1}$$

with the vector $\boldsymbol{x}=[x_1, x_2, \ldots, x_n]^T$ of n explanatory variables $x_1, x_2, \ldots, x_n$ and the corresponding coefficient vector $\beta=[\beta_1, \beta_2, \ldots, \beta_n]^T$. Because $P(y \le m)=1$, the parameter θ_m is equal to ∞. The latent variable z is the linear combination of the explanatory variables x_i, $(i=1, \ldots, n)$:

$$z = -\beta_1 x_1 - \beta_2 x_2 - \ldots \beta_n x_n = -\boldsymbol{\beta}^T \boldsymbol{x}, \tag{2}$$

and summarizes the financial information of the risk entity. From the cumulative probabilities $P(y \le i)$, with $i=1, \ldots, m$, one obtains the probabilities $P(y=i)$ as $P(y=1)=P(y \le 1)$, $P(y=i)=P(y \le i)-P(y \le i-1)$ for $1<i<m$ and $P(y=m)=1-P(y \le m-1)$.

Given a training data set $D=\{\boldsymbol{x}_i, y_i\}_{i=1}^N$ of N data points, the parameters $\theta_1, \theta_2, \ldots, \theta_m$ and $\beta_1, \beta_2, \ldots, \beta_n$ are estimated minimizing the negative log likelihood (NLL):

$$\left(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m; \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_n\right) = argmin\, NLL(\boldsymbol{\theta}, \boldsymbol{\beta}) = -\sum_{i=1}^{N} \log(P(y=y_i)), \tag{3}$$

with $\theta_m=\infty$ and $y_i \in \{1, \ldots, m\}$. As a result of the maximum likelihood optimization, not only the optimal parameters are obtained, but also the standard errors (square roots of the diagonal elements of the inverse Hessian) and the corresponding p-values (z-test). The model deviance (dev) is equal to twice the negative log likelihood in the optimum and can be used for model comparison, e.g., using an appropriate information criterion [5,24]. The statistical relevance of input i can be assessed from its p-value of the hypothesis test H0 ($\beta_i=0$). It is also reported here by the difference in model deviance[5] between the full model $\mathcal{M}_1$ (with inputs 1, …, $i-1$, i, $i+1$, …, m) and the reduced model $\mathcal{M}_0$ without the corresponding input (inputs 1, …, $i-1$, $i+1$, …, m). The Bayes factor $\mathcal{B}_{10}$ is approximated via

$$2\log(\mathcal{B}_{10}) \approx \mathrm{dev}(\mathcal{M}_0) - \mathrm{dev}(\mathcal{M}_1) = \Delta\mathrm{dev} \tag{4}$$

and indicates the model improvement and has to be sufficiently large as indicated in Ref. [16]: $0 \le 2\log(\mathcal{B}_{10}) < 2$ not worth more than a bare mention, $2 \le 2\log(\mathcal{B}_{10}) < 5$ positive evidence against H0 hypothesis of no improvement, $5 \le 2\log(\mathcal{B}_{10}) < 10$ strong evidence and $10 \le 2\log(\mathcal{B}_{10})$ decisive evidence.

2.2. Intrinsically linear ordinal logistic regression

In the linear model (2), a ratio x_i influences the latent variable z in a linear way. However, it can be argued that a change of a ratio with 1% should not always have the same influence on the score and risk [2], e.g., an increase of 10% of debt to exports from 50% to 60% is reported not to have the same impact on the economic growth as an increase from 200% to 210% [20]; and, hence, may influence the country risk differently.

Therefore, one often suggests to estimate univariate nonlinear transformations ($x_i \mapsto f_i(x_i)$) for some of the independent variables [6]. Applying the transformation to ratios $m+1, \ldots, n$, the z-score (Eq. (2)) becomes

$$z = -\beta_1 x_1 - \ldots - \beta_m x_m - \beta_{m+1} f_{m+1}(x_{m+1}) - \ldots - \beta_n f_n(x_n). \tag{5}$$

This model is called intrinsically linear in the sense that after applying the nonlinear transformation to the explanatory variables, a linear model is being fit [6]. A nonlinear transformation of the explanatory variables is applied only when it is reasonable from both a financial

[4] It is also possible to apply least squares regression to the classes, but the result may or may not depend on the numerical coding of the classes; e.g., for the 5 classes one may choose codings (1, 2, 3, 4, 5) or (1, 2, 4, 8, 16) which yield possibly different results. The ordinal logistic regression formulation is independent of the classes and therefore often preferred from a theoretical perspective. Additionally, it includes a probabilistic interpretation that indicates how sure the model is on a rating decision.

[5] It is preferred to report the deviance as it is straightforward to compute the appropriate complexity criteria from the deviance.

as well as a statistical perspective as will be illustrated in Section 3.

The Box–Cox power transformations are a well-known type of transformation to improve symmetry, normality or model fit [6,30]. However, these transformations are only defined for positive values $x>0$. Recently, an alternative family of transformations [30] has been proposed that is of the same form as the Box–Cox transformations and is also valid for negative values:

$$f(x;\lambda)=\begin{cases} x\geq 0: & \left((1+x)^{\lambda}-1\right)/\lambda & (\lambda\neq 0) \;\&\; \log(x+1) & (\lambda=0) \\ x<0: & -\left((1-x)^{2-\lambda}-1\right)/(2-\lambda) & (\lambda\neq 2) \;\&\; -\log(-x+1) & (\lambda=2). \end{cases} \tag{6}$$

It can be easily verified that for $\lambda=1$ the identity transformation $x\mapsto x$ is obtained. If $\lambda=0$ ($\lambda=2$), the log transform is applied to the positive (negative) values, whereas negative (positive) values are transformed accordingly via a smooth transition between positive and negative values. The tuning parameters λ_i, $i=m+1, \ldots, n$ of the nonlinear transformations can be selected based on expert knowledge or can be estimated from the training data, as is described in Appendix A.

2.3. Support Vector Machines

Given its universal approximation property [5,24], the Multilayer Perceptron (MLP) neural network is a popular neural network for both regression and classification and has often been used in financial contexts such as bankruptcy prediction and credit scoring (see, e.g., Refs. [3,13,21,26]). Although nowadays there exist good training algorithms (e.g. Bayesian inference) [5,24] to design the MLP, there are still a number of drawbacks, like the choice of the architecture of the MLP and the existence of multiple local minima, which imply that the estimated parameters may not be uniquely determined. Recently, a new learning technique emerged, called Support Vector Machines (SVMs) and related kernel based learning methods in general, in which the solution is unique and follows from a convex optimization problem [24,28,29].

SVMs were first derived for the binary classification problem with class labels -1 and $+1$. The classifier has the form

$$y(\boldsymbol{x}) = \operatorname{sign}\left[\boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x}) + b\right], \tag{7}$$

where the coefficient vector $\boldsymbol{w}\in\mathcal{R}^{n_\varphi}$ and bias term b have to be estimated from the data. The corresponding score function is equal to $z=\boldsymbol{\varphi}(\boldsymbol{x})+b$. The nonlinear function

$$\boldsymbol{\varphi}(\cdot) : \mathcal{R}^n \rightarrow \mathcal{R}^{n_\varphi} : \boldsymbol{x}\mapsto\boldsymbol{\varphi}(\boldsymbol{x}) \tag{8}$$

maps the input space to a high (possibly infinite) dimensional feature space (see Fig. 3). In this feature space, a linear separating hyperplane $\boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x})+b=0$ is then constructed applying linear methodology. In SVMs, the classifier is obtained from a convex quadratic programming (QP) problem in the parameters $\boldsymbol{w}$ and b subject to $2N$ constraints as explained in Appendix B. A key element of nonlinear SVMs and kernel based learning in general is that the nonlinear mapping $\boldsymbol{\varphi}(\cdot)$ and the weight vector $\boldsymbol{w}$ are never calculated explicitly. Instead, Mercer's theorem

$$K(\boldsymbol{x}_i,\boldsymbol{x}_j) = \boldsymbol{\varphi}(\boldsymbol{x}_i)^T\boldsymbol{\varphi}(\boldsymbol{x}_j) \tag{9}$$

is applied to relate the mapping $\boldsymbol{\varphi}(\cdot)$ with the symmetric and positive definite kernel function K. For $K(\boldsymbol{x}_i,\boldsymbol{x})$ one typically has the following choices: $K(\boldsymbol{x}_i,\boldsymbol{x})=\boldsymbol{x}_i^T\boldsymbol{x}$ (linear kernel); $K(\boldsymbol{x}_i,\boldsymbol{x})=(\boldsymbol{x}_i^T\boldsymbol{x}+\eta)^d$ (polynomial SVM of degree d with η a positive real constant); $K(\boldsymbol{x}_i,\boldsymbol{x})=\exp(-\|\boldsymbol{x}-\boldsymbol{x}_i\|_2^2/\sigma^2)$ (RBF-kernel with bandwidth parameter σ). Constructing the Lagrangian of the QP problem, one can eliminate $\boldsymbol{w}$ from the conditions of optimality in the saddle point of the Lagrangian and formulate a dual optimization problem in the Lagrange multipliers $\alpha=[\alpha_1,\ldots,\alpha_N]^T\in\mathcal{R}^N$. The resulting classifier is depicted in Fig. 4 and is given by

$$y(\boldsymbol{x}) = \operatorname{sign}\left[\sum_{i=1}^{N}\alpha_i K(\boldsymbol{x},\boldsymbol{x}_i) + b\right], \tag{10}$$

with $z=\sum_{i=1}^{N}\alpha_i K(\boldsymbol{x},\boldsymbol{x}_i)+b$ (see Appendix B for details).

More generally, SVMs and related kernel based learning techniques for QP classification, Fisher Discriminant Analysis, logistic regression and least squares

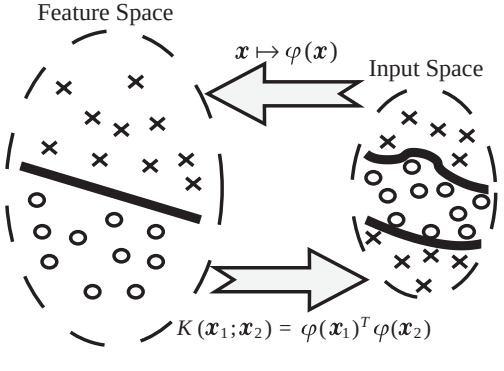


Fig. 3. Illustration of SVM based classification.

ARTICLE IN PRESS

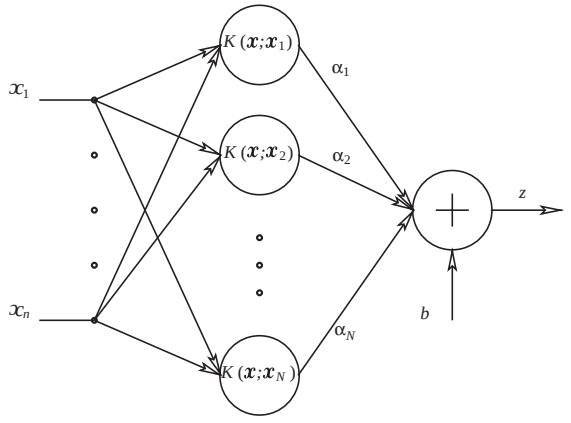


Fig. 4. Network architecture of kernel based estimators.

regression follow more or less the following steps. One starts with mapping the input space to a high dimensional feature space where the mapping itself is implicitly defined by the Mercer condition (9). A (linear) regression or classification technique is then formulated in the (primal) feature space, where one typically uses a regularization term to avoid over-fitting. The Lagrangian is constructed and a (dual) optimization problem is formulated in the Lagrange multipliers. The solution is then expressed in terms of the resulting Lagrange multipliers and the kernel function K.

2.4. Adding SVM terms to the linear model

Given the intrinsically linear model of Eq. (5), more complex nonlinearities can be captured by adding nonlinear SVM terms $w_i\varphi_i(\boldsymbol{x})$ as follows

$$z = -\beta_1 x_1 - \ldots - \beta_m x_m - \beta_{m+1} f(x_{m+1}) - \ldots - \beta_n f(x_n) - w_1\varphi_1(\boldsymbol{x}) - \ldots - w_p\varphi_p(\boldsymbol{x}), \qquad (11)$$

where $w_1\varphi_1(\boldsymbol{x})+\ldots+w_p\varphi_p(\boldsymbol{x})$ will be expressed in terms of the kernel function K. The estimation of the coefficients is done within the OLR framework, using primal–dual relations from Nyström sampling (originally derived for SVMs) as detailed in Appendix B.

Different approaches exist to estimate the parameters β_i, w_i from given training data. A first alternative is to perform a joint estimation of both parameters of the (intrinsically) linear part and the SVM terms. This approach has the advantage that the estimation is done in the space spanned by all the regressors, hereby yielding an optimal solution. A disadvantage of this approach however is that the SVM terms may be preferred by the input selection algorithms over the linear terms, which reduces the readability of the estimated model. Therefore, an alternative approach has been suggested in econometrics. The nonlinear terms of the SVM are estimated by means of partial regression on top of the estimated (intrinsically) linear model. In a least squares regression set-up, this would correspond to modelling the residuals with a nonlinear model. As readability is an important aspect of the model for its successful use and interpretation by the financial analyst, the second approach will be adopted in the empirical section.

3. A process model for developing an internal rating system

The construction of the internal rating system is done according to the process model depicted in Fig. 1. In the first step, a database with a list of candidate explanatory variables and external ratings[6] is constructed. We used the long term foreign currency rating of Moody's, as opposed to the S&P and Fitch ratings, because this rating agency rates the highest number of countries for the set up of the internal rating system. The database is constructed in close collaboration with financial analysts to make sure that all necessary ratios are included in the candidate set of explanatory variables.

The predictive rating model is designed in a second step. First, an appropriate modelling technique is selected. Input selection is considered in order to get more concise, comprehensible and powerful models. Both quantitative and qualitative data will be considered to train the models. The estimated models are then extensively validated using different performance measures and using different (cross-) validation tests [11]. This step is concluded by the scorecard definition, the user guide and the guidelines with, e.g., the perimeter definition and the overruling procedure. One may also define backtesting procedures.

The calibration of the IRB risk system is done in Step 3. The internal ratings are a key input in the internal ratings based risk system. Based on the Vasicek one-factor model, the loss distribution follows the normal inverse distribution and the required regulatory capital is set to the appropriate confidence level [4]. In the internal ratings based approach, the corresponding default probability (PD) per rating needs to be determined. Given the limited default history for

[6] In the case of default data, the target variable is the binary variable default/non-default.

rated sovereigns and given that the observed corporate and sovereign default rates are found not to be significantly different, one may, e.g., opt to apply corporate default rates, eventually adjusted by the historical default rates for sovereigns from the rating agencies. Given the limited default information, one may also infer the risk neutral default probabilities from debt prices, credit derivative prices and equity (index) prices. Since risk neutral probabilities do not always correspond to historical default probabilities, a conversion factor may be applied. For the advanced internal ratings based approach, one also needs to estimate[7] the loss given default (LGD). The calibration of the PD and LGD is not the scope of this paper.

3.1. Step 1: database construction and preprocessing

3.1.1. Database construction

The selection of the candidate explanatory variables is based on the expertise of the financial analysts and on an extensive literature study, a.o., Refs. [8,14]. The retrieved variables are both pure economical and financial variables as well as qualitative ratios. Most of the variables listed in Table 1 were retrieved from two Worldbank databases: World Development Indicators (WDI) and Global Development Finance (GDF), which also contain data from the International Monetary Fund. Additionally, other ratios were retrieved from Moody's (M), Transparency International (TI), and the United Nations Human Development Programme (UN).

Based on the structure of the WDI database, these candidate explanatory variables are subdivided in the following categories: demographic (1–17); economic (18–39); debt (40–56) and markets (57–62). Note that Tables 1 and 2 also report the expected sign (E.S.) of each variable from a univariate macro-economic perspective on the creditworthiness of a country. A positive sign means that the creditworthiness increases when the ratio increases.

The total number of countries and regions with candidate explanatory variables in the database is about 200, of which about 95 were regularly rated. Hence, the data retrieval for the model development was restrained to these countries. The variables are retrieved over a considerable time period, so as to have a rich data set with a sufficient number of observations. It is also important to note that the model should preferably be trained on recent data, as due to non-stationarity, the risk elements and behavior may change. It is decided to use a 6-year time period 1997–2002 in this paper, where the external ratings at the end of each year are considered to ensure that the model runs with input variables that become available before the rating is calculated.

3.1.2. Definition of new explanatory variables

The aim of the internal rating model is to estimate the rating of a country in year $T+1$, given the information from previous years T, $T-1$, $T-2$, $T-3$ and $T-4$. In order to do this, the following derived variables are computed: the 5-year average (AV) $x_{5y}=(x_{T-1}+x_{T-2}+x_{T-3}+x_{T-4})/5$, the last available value (T0) x_T, the relative trend[8] (rTR) $x_{rtr}=(x_T-x_{T-4})/(4x_{T-4})$ and the absolute trend (aTR) $x_{atr}=(x_T-x_{T-4})/4$. When the choice between the last available and the average value of a ratio is equal from a statistical perspective, the use of the last available values is preferred for structural variables, whereas average values are preferred to average out trends for the more volatile variables linked to the business cycle.

3.1.3. Missing values

Missing values are commonly treated using imputation procedures which replace them by the mean or median (for continuous attributes) or the mode (for discrete attributes) of the distribution. Missing values were considered in two ways: countries and variables. The following countries were removed due to the too limited information available: the Bahamas, Macao (China), Lebanon, Qatar, San Marino, Turkmenistan, and the United Arab Emirates. As stated above, the considered time period for the rating is the beginning of 1997 until the beginning of 2003. As not all countries have ratings for the full 6-year period, 511 country–year observations are available. Starting from the 511 country–year observations and the 63 candidate inputs listed in Table 1, with a 5-year history, the number of times a candidate input is missing for the full period was analyzed and reported in Fig. 5a. The variables 43 until 57 are debt variables and have a high number of missing values, which is mainly due to the fact that most debt variables considered are not systematically available for developed countries or advanced industrial countries. As debt variables are believed to be important for predicting creditworthi-

[7] In addition to the PD rating scale, one may also define an LGD rating scale applying LGD scoring.

[8] For reasons of readability and monotonicity, the relative trend will only be used when the denominator has a constant sign.

ARTICLE IN PRESS

Table 1
Variable list

Nr.	Variable	E.S.	Coeff.	S.D.	*P*-value	Δdev	Motiv.
1	Health expenditure per capita (current US$)	+	0.984	0.151	0	6.03	I
2	Health expenditure, total (% of GDP)	+					
3	Improved water source (% of population with access)	+	0.027	0.014	0.048	3.83	II
4	Birth rate, crude (per 1000 people)	+	−0.058	0.016	0.001	4.99	I, III
5	Death rate, crude (per 1000 people)	−	0.189	0.041	0	12.09	III
6	Fertility rate, total (births per woman)	+	−0.409	0.118	0.001	12.14	III
7	Life expectancy at birth, total (years)	+	0.003	0.031	0.918	0.01	IV
8	Mortality rate, under −5 (per 1000 live births)	−	−0.012	0.007	0.072	3.25	IV
9	Gini index (−, T0)	−	−0.007	0.012	0.566	0.32	II
10	Poverty headcount, national (% of population)	−	0.042	0.013	0.001	1.01	II
11	Malnutrition prevalence weight for age (% children under 5)	−	−0.058	0.018	0.001	10.35	II
12	Human development index	+	10.325	1.580	0	7.08	IV
13	Corruption perception index	+					
14	School enrolment primary (% gross)	+	0.008	0.010	0.393	0.73	IV
15	School enrolment secondary (% gross)	+	0.008	0.005	0.079	3.09	IV
16	School enrolment tertiary (% gross)	+	0.016	0.006	0.011	6.39	IV, V
17	Illiteracy rate, adult total (% of people ages 15 and above)	−	−0.014	0.011	0.176	1.8	II
18	Unemployment (% of total labour force)	−	−0.026	0.022	0.256	1.28	IV
19	GDP per unit of energy use (PPP $ per kg of oil equivalent)	+	0.043	0.050	0.385	0.75	IV
20	GDP growth (annual %)	+	0.009	0.034	0.773	0.08	IV
21	GDP per capita (constant 1995 US$)	+	1.086	0.159	0	9.99	VI
22	GDP per capita growth (annual %)	+	0.055	0.031	0.074	3.17	IV
23	GDP per capita (PPP)	+	1.638	0.243	0	8.80	I, V
24	GDP per capita (US$)	+					
25	Gross capital formation (% of GDP)	+					
26	Gross domestic savings (% of GDP)	+	0.016	0.014	0.228	1.45	IV
27	Inflation consumer prices (annual %)	−	0.006	0.035	0.855	0.03	IV
28	Exports of goods and services (% of GDP)	+	0.010	0.005	0.064	3.42	IV
29	Imports of goods and services (% of GDP)	−	0.006	0.005	0.177	1.82	IV
30	Food imports (% of merchandise imports)	−	0.001	0.021	0.998	0.00	IV
31	Interest payments (% of current revenue)	−	−0.021	0.014	0.131	2.28	IV
32	Overall budget balance including grants (% of GDP)	+	0.019	0.038	0.614	0.25	IV
33	Money and quasi money (M2) as % of GDP	−	0.001	0.004	0.787	0.07	IV
34	Money and quasi money (M2) to gross international reserves ratio	−	−0.099	0.042	0.020	5.37	I, III, IV
35	Money and quasi money growth (annual %)	−	−0.004	0.009	0.667	0.18	IV
36	Foreign direct investment net inflows (% of GDP)	+	0.047	0.029	0.107	2.61	IV
37	Food production index	+	−0.001	0.003	0.875	0.02	IV
38	Net barter terms of trade (1995=100)		−0.004	0.010	0.722	0.12	II
39	Current account balance (% of GDP)	+	−0.001	0.019	0.955	0.00	IV
40	Gross international reserves in months of imports	+					
41	Budget balance/GDP (%)	+	0.042	0.031	0.175	1.83	IV
42	Public debt/GDP (%)	−	−0.002	0.003	0.647	0.20	IV
43	Cumulated Debt Forgiveness/GDP	−	−7.120	5.954	0.231	1.42	IV
44	Interest arrears on total long term debt/GDP	−	−1.388	1.199	0.247	1.33	IV
45	Principal arrears on total long term debt/GDP	−	−0.463	0.272	0.088	2.89	IV
46	Interest rescheduled (capitalized) (US$)	−				1.00	II
47	Reserves vs. total debt	+	−0.714	0.837	0.393	0.72	IV
48	Total debt vs. Reserves	−	0.020	0.075	0.789	0.07	IV
49	ST-debt vs. T-debt	−	1.690	0.957	0.077	3.12	IV
50	Total debt service paid/CA (%)	−	0.046	0.028	0.103	2.61	IV
51	(Total debt service paid + ST-debt)/CA (%)	−	0.004	0.008	0.628	0.23	IV
52	Total debt service paid/exports of goods and services (%)	−	−1.152	1.645	0.483	0.49	IV
53	(Total debt service paid + ST-debt)/exports of goods and services (%)	−	0.680	0.771	0.377	0.77	IV
54	(Total debt service paid + ST-debt)/reserves (%)	−	0.018	0.176	0.916	0.01	IV
55	Total debt stocks/GDP	−	2.587	0.632	0.000	16.92	V
56	Total debt stocks/CA	−					
57	Public and publicly guaranteed (PPG) debt (% of GDP)	−	−0.371	0.828	0.654	0.20	IV
58	Gross foreign direct investment (% of GDP)	+	0.014	0.018	0.448	0.57	IV

Table 1 (*continued*)

Nr.	Variable	E.S.	Coeff.	S.D.	*P*-value	Δdev	Motiv.
59	Market capitalisation of listed companies (% of GDP)	+	−0.001	0.002	0.893	0.01	IV
60	Interest rate spread (lending rate minus deposit rate)	−	0.031	0.032	0.338	0.91	IV
61	Real effective exchange rate index (1995 = 100)	+	−0.022	0.011	0.045	4.01	II
62	Real interest rate (%)	−					
63	Risk premium on lending (% gross)	−	−0.036	0.044	0.414	0.66	II

Column 3 indicates the expected sign, columns 4, 5, 6 and 7 report the coefficient, standard deviation, *p*-value and difference in deviance. A motivation for why the variable is not selected in model 3 is given in the last column. I: the introduction of this ratio yields a wrong sign and/or reduces the readability as perceived by the financial analysts; II: the ratio has too many missing values; III: the difference in deviance is too small; IV: the estimated coefficient is statistically not significantly different from zero, the corresponding *p*-value is too high; V: the leave-one-out and/or leave-country-out cross-validation performance decreases when using the ratio into the model; VI: another type of ratio, e.g., the last available value is preferred by the financial analysts without reducing the out-of-sample performance.

ness of a developing country, these ratios are kept and a dummy/indicator variable will be introduced in the modelling step for the developed countries so as to adjust for the missing values and possible other qualitative effects, like, e.g., possible increased financial stability and economic development. Because the candidate inputs 3, 9, 10, 11, 38, 61 and 63 have a high number of missing values, the financial analysts approve to not consider them for input selection. Fig. 5b depicts the percentage of missing values per country during the 5-year period, without taking into account the removed variables. No countries were additionally removed from the database. To all remaining missing values, median imputation was applied.

3.1.4. Input transformations

All size variables are transformed as their distribution is typically far from Gaussian. Since all observations of the size variables health expenditure per capita, GDP per capita (constant 1995 US$), GDP per capita (PPP) and GDP per capita (US$) are positive, the logarithmic transformation ($x \mapsto \log(1+x)$) is applied [8,14].

3.1.5. Outlier handling

Since most candidate explanatory variables are ratios, it is expected that the distributions of these variables may have fat tails with large positive and negative values. These data points usually correspond to leverage points (X-outliers). In order to avoid that these outliers have a negative influence on the model performance, the most extreme points are selected and reduced to the 3σ-borders in a similar way as in the winsorised mean procedure. For the limits $m \pm 3 \times s$, one computes m and s in a robust way using the median and $s = \frac{\mathrm{IQR}(x)}{(2\times 0.6745)}$, with IQR the interquartile range. These limits were also verified by the financial analysts.

3.2. Step 2: modelling

3.2.1. Model requirements and specifications

The model is designed to meet the following requirements:

1. The model has to be stable, meaning that the estimated coefficients are well determined with high confidence and sufficiently low uncertainty. Moreover, each variable should have a significant contribution in the model.
2. The readability of the model is another important performance measure. It should be relatively easy to interpret the model for the financial analysts.
3. The model needs to accurately discriminate the solvent countries from the non-solvent countries. Assuming that the external rating is discriminative, the internal rating should approximate the external rating as good as possible.

The first performance criterion, i.e., stability, is measured in three ways. First, for all coefficients, the *p*-value has to be sufficiently low. Given the number of observations, a *p*-value below 5% is required and it is preferred to have all *p*-values below 1%. Secondly, each variable has to yield a significant improvement in the deviance of the model as reported in Subsection 2.1. Thirdly, it is verified whether the values of the estimated coefficients of the selected variables do not change too much with removal of a country from the data set. This additional check is carried out to avoid that the resulting model would become too dependent on the data sample.

The readability of the ordinal logistic regression model is relatively high. Although it is sometimes noted that "wrong sign problems[9]" are not important

[9] An estimated coefficient has a wrong sign when it is opposite to the sign expected from a financial perspective.

Table 2
Selected explanatory variables in model 1 (linear, all countries), model 2 (linear, developed/developing) and model 3 (intrinsically linear)

Variable	Type		E.S.	Model 1			Model 2			Model 3		
				S.C.	p-value (%)	Δdev	S.C.	p-value (%)	Δdev	S.C.	p-value (%)	Δdev
Health expenditure, total (% of GDP)	T0	Type 0	+	+	0.000	−82.4	+	0.000	−81.7	+	0.000	−93.1
Corruption perception index	T0	Type 0	+	+	0.003	−17.8	+	0.524	−7.8	+	0.000	−21.0
Mortality rate, under −5 (per 1000 live births)	T0	Type 1	−				−	0.000	−43.1	−	0.002	−19.1
School enrolment secondary (% gross)	AV	Type 0	+	+	0.006	−16.2	+	0.741	−7.2			
GDP per capita (US$)	T0	Type 0	+	+	0.001	−21.0	+	0.000	−35.7	+	0.000	−38.1
GDP growth (annual %)	AV	Type 0	+	+	0.000	−36.4	+	0.006	−11.9	+	0.006	−16.5
Gross capital formation (% of GDP)	T0	Type 0	+				+	0.001	−21.4	+	0.006	−16.7
Gross domestic savings (% of GDP)	AV	Type 0	+	+	0.000	−27.1						
Gross domestic savings (% of GDP)	T0	Type 1	−				−	0.258	−9.1			
Inflation consumer prices (annual %)	AV	Type 0	−	−	0.142	−10.2	−	0.002	−18.1	−	0.002	−18.8
Inflation consumer prices (annual %)	TR	Type 2	−							−	0.376	−8.5
Interest payments (% of current revenue)	T0	Type 0	−	−	0.894	−6.9						
Interest payments (% of current revenue)	T0	Type 1	−				−	0.008	−16.7			
Interest payments (% of current revenue)	AV	Type 1	−							−	0.006	−17.2
Current account balance (% of GDP)	T0	Type 1	+							+	0.191	−9.2
Gross international reserves in months of imports	T0	Type 0	+	+	0.001	−19.4						
Gross international reserves in months of imports	T0	Type 2	+				+	0.000	−58.7	+	0.000	−54.8
Public debt/GDP (%)	TR	Type 0	−	−	0.001	−19.6	−	0.000	−26.2	−	0.000	−26.0
Interest arrears on total long term debt/GDP	AV	Type 2	−	−	0.000	−57.8	−	0.000	−31.9	−	0.000	−32.9
Cumulated debt forgiveness/GDP	T0	Type 2	+				+	0.002	−18.4	+	0.000	−40.1
Total debt service paid/CA (%)	AV	Type 2	−	−	0.000	−35.3	−	0.141	−10.3	−	0.236	−9.3
Total debt stocks/CA	TR	Type 2	−				−	0.013	−15.1	−	0.001	−20.5
Total debt stocks/CA	AV	Type 2	−				−	0.019	−14.2	−	0.006	−16.5
Real interest rate (%)	T0	Type 0	−	−	0.000	−46.2						
Real interest rate (%)	T0	Type 1	−				−	0.005	−17.3	−	0.001	−19.6
Real Interest rate (%)	T0	Type 2					−	0.000	−21.6	−	0.004	−17.5
Ind. developed countries			+	+	0.000	−76.1	+	0.000	−120.6	+	0.000	−143.6

The types T0 (most recent observation), AV (5-year average), aTR (absolute trend) and rTR (relative trend) are reported in column 2. Column 3 indicates whether the ratio is used to discriminate between all countries (type 0), developed (type 1) or developing (type 2) countries. The expected sign is reported in column 4. The sign of the estimated coefficients (SC), the p-values and the differences in deviance are then reported for each model.

in a multivariate regression context, due to the correlation between the variables, it is preferred here that the signs are in line with the expectations of the team of financial analysts, so as to enhance the readability of the model. Such approaches are, e.g., also observed in Refs. [8,14].

The classification performances will be computed based on the confusion matrix numbers. These matrices are summarized by the cumulative notch difference, overall classification accuracy and classification accuracy per rating category (Aaa, Aa, A, Baa, ...). These performance measures can be computed using several sampling strategies. Remember that the resulting data set consists of about 6 years of information on 88 countries, yielding a total number of 511 country–year combinations. As this number is relatively low, it is decided not to

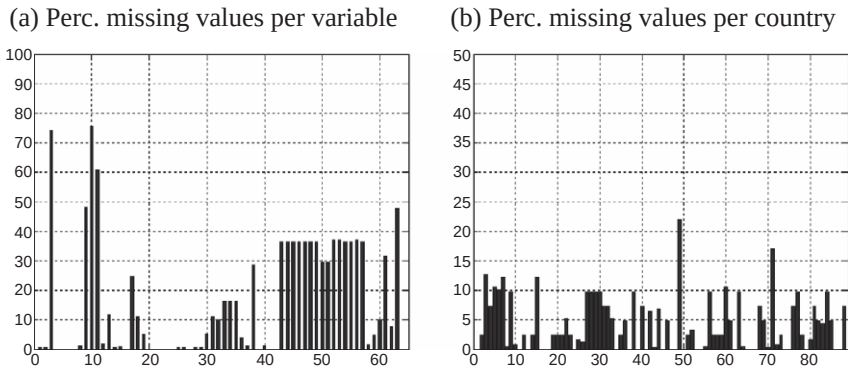


Fig. 5. Percentage of missing values per variable and per country.

split-up the data into a training set, used for estimating the model, and a separate validation set, used for calculating its performance [5]. Instead, the performance of the model will be evaluated using leave-one-out cross-validation [5,11]. Notice, however, that this approach has the disadvantage that already some part of the country information is in the training data set. Therefore, the cross-validation performance whereby all country–year combinations relating to the same country are put into the validation set is also assessed. It basically represents the performance of the rating system on countries on which the model was not trained.

3.2.2. Model estimation

3.2.2.1. General model for all countries (Model 1). An ordinal logistic regression model is first estimated to rate both the developed and the developing countries. Since parsimonious models are generally preferred, backward, forward and stepwise input selection techniques are applied first to explore the data set. The experience of the financial analysts is then extensively used in the model design to steer the input selection process so as to obtain a stable and performing model both in terms of financial and statistical requirements. The results are reported in the columns labelled Model 1 of Table 2. Note that due to confidentiality and non-disclosure agreements, the estimated coefficients are not reported, but all considered inputs have the expected sign and are highly statistically significant (p-value $\leq 1\%$). The leave-one-out and leave-country-out performances[10] are reported in Table 3.

The model finds a balance between demography (3 variables), economy (5 variables), debt (4 variables) and markets (1 variable). From a macro-economic viewpoint, all "classic" variables are represented in the model (GDP, inflation, real interest rate, public debt, ...) [8]. The corruption perception indicator is also found to be significant. A closely related qualitative variable, government effectiveness, was found to be significant in recent studies. The school enrolment is a qualitative variable that reflects the future growth perspective of the country. Education has also a positive impact on health. Health expenditure and the indicator variable (developed/developing countries) seem to have a large impact on the difference in model deviance. Chakraborty showed that health expenditure has a stronger impact on human development and well-being than the growth of per capita income [9]. Hence, investing in health expenditure has important implications both from a social and economical perspective [23]. The indicator variable takes into account the median imputation for the debt variables and the reduced external transfer risk as perceived by the agencies [8]. For the debt variables, both the debt burden and the debt level are important [8,14]. The interest arrears (% of GDP) variable is an indicator of the near past debt repayment history of the country.

3.2.2.2. Combined model for developed/developing countries (Model 2). Note that in the model of the previous subsection, an indicator variable was introduced so as to distinguish between developed and developing countries, mainly because debt information is not systematically available for some of the developed countries. As these missing values are typically replaced by the median of the developing countries, a systematic bias for developed countries is introduced. The coefficient of the indicator variable allows to adjust the rating for developed countries by a constant shift.

[10] Besides the small training set, the difference in performance can be explained by the fact that some rating categories are underrepresented in the full database, and become even more underrepresented when the specific country is removed from the training database, influencing the θ parameters.

ARTICLE IN PRESS

Table 3
Comparison of the leave-one-out (loo) and leave-country-out (lco) cumulative accuracy on 0 to 4 notches difference, respectively

Model nr.	Performance	0 (%)	0–1 (%)	0–2 (%)	0–3 (%)	0–4 (%)	Dev.
1	loo	39.7	69.7	88.5	96.7	99.2	1721
2	loo	42.3	78.1	92.2	98.3	99.0	1602
3	loo	43.6	79.3	92.6	97.6	99.4	1564
4	loo	44.8	79.8	92.8	98.0	99.6	1542
1	lco	32.7	63.2	84.7	95.3	98.6	1917
2	lco	34.1	71.2	88.5	97.3	98.6	1858
3	lco	35.8	72.8	92.6	97.3	99.0	1822
4	lco	36.6	74.4	91.8	97.5	99.0	1805

Referring to Ref. [16], it can be concluded that the difference in deviance is significant.

However, it could also be interesting to let the z-score depend on whether variables or ratios are measured for developed or developing countries. This is done by rewriting[11] the z-score as follows:

$$z = -\beta_1 x_1 - \beta_2 x_2 - \ldots - \beta_n x_n - I(D)\beta'_p x_p - I(U)\beta''_q x_q, \quad (12)$$

with the indicator functions $I(D)$ ($I(U)$) that equal one for developed (developing) countries, and zero otherwise. Hence, this means that the variables $x_1, x_2, \ldots, x_n$ and x_p are used to calculate the score of a developed country, whereas the score of a developing country is calculated using variables $x_1, x_2, \ldots, x_n$ and x_q. The use of different coefficients for the same variable allows to weight that variable in a different way for developed and developing countries. As a result, there are 3 types of variables. Type 0 variables discriminate between all countries (both developed and developing). Type 1 and 2 variables discriminate, respectively, between developed and developing countries only. The optimal set of selected inputs in the new model is reported in the columns labelled Model 2 of Table 2. The corresponding performances on 0–4 notches differences are reported in Table 4. Note that the performance improved when compared to the previous model. Furthermore, when contrasting the new results with the previous ones, it can be seen that many type 0 inputs are the same, mainly because the new model was conceived starting from the previous one, and because type 0 variables are preferable from the readability perspective.

As could be expected, the selected type 2 variables reported in Table 2 are mainly debt variables. The evolution of public debt (% of GDP) is a general indicator for both developed and developing countries. For developed countries, the interest payments (% of current revenue) are an indication for the debt burden (although without debt repayment information). For the developing countries, more debt indicators are available and many are selected, including the debt service and debt stocks (and its trend), as well as the liquidity indicator import cover and debt repayment history via interest arrears and cumulated debt forgiveness. The selected debt variables are in line with the findings in the literature on "default" prediction[12] [22] and explaining external ratings [8,14]. Furthermore, total debt service is less significant compared to total debt stocks as can be seen from the difference in deviance reported in Table 2.

Demographic, economic and market variables are selected as type 0 or as type 1 variables, i.e. to discriminate between all countries or between developed countries. Since the real interest rate is significantly different between developed and developing countries, due to the different macro-economic and financial climate, different weights are used in the rating model. As mentioned earlier, health expenditure remains the most important variable when discriminating between all countries. However, it needs to be noted that higher health expenditure does not necessarily imply better health and thus socio-economic welfare, since it also depends on the distribution thereof. The Gini index is found not to be additionally significant. In some sense, it is surprising that also the mortality rate under −5 is considered as an important discriminating variable between developed countries. On the other hand, it is well known that "development" is strongly associated with improvements in mortality [7]. According to Ref. [12], the child mortality rate can be explained by three factors: cost effectiveness on public spending, the net impact of additional public supply and public sector efficacy. Investment in capital goods, measured via gross capital formation, is a classical indicator for future growth and becomes an important discriminative variable in the model [8]. Gross domestic savings is significant only for developed countries. A positive savings result is, e.g., positive for future growth and the strength of the banking system; while a too high

[11] For notational convenience, only one variable for developing and developed countries is used in Eq. (12). Of course, more variables for the developed and developing part can be introduced in practical models.

[12] In most of these studies, default is labelled as debt service difficulty, debt crisis or default.

Table 4
Analysis of the rating accuracy for the rating spectrum divided into 5 main categories

Ext. rating	Nobs	$y_{ext} - y_{pred}$						
		>2 (%)	=2 (%)	=1 (%)	=0 (%)	=−1 (%)	=−2 (%)	<−2 (%)
Aaa–Aa3	139	1.4	5.7	19.4	64.0	6.4	2.1	0.7
A1–A3	52	11.5	0.0	7.7	28.8	26.9	13.4	11.5
Baa1–Baa3	111	0.0	6.3	7.2	42.3	31.5	12.6	0.0
Ba1–Ba3	108	3.7	8.3	31.4	29.6	14.8	5.5	6.4
B1–CCC	101	10.8	9.9	15.8	45.5	15.8	1.9	0.0

saving may reduce public spending and slow down the economy. Given the small difference in deviance, it is observed that this variable is rather weakly significant.

3.2.2.3. Intrinsically linear model for developed/developing countries (Model 3). We also investigated whether transformations like Eq. (6) (see Section 2) could improve the performance. The following two criteria are considered before using a nonlinear transformation in the model:

1. The model fit needs to improve significantly according to Ref. [16].
2. Each nonlinear transformation has to be meaningful from a financial perspective.

It is preferred to keep the number of nonlinear transformations as low as possible.

The identification of the nonlinearities $f(x_i; \lambda_i)$ and the transformation parameter λ_i is done using a grid search algorithm described in Appendix A. This procedure is applied starting from the identified linear model of the previous paragraph. First, for each variable, the optimal nonlinear transformation is determined. In a next step, the nonlinear transformation with the highest decrease in deviance (if possible) is included. Again, input selection is performed and the next nonlinear transformation is identified. This greedy procedure is stopped when there are no more valid transformations to be included. The univariate nonlinear transformations that were found in this way are visualized in Fig. 6.

The corruption perception index (CPI) classifies the countries in terms of perceived corruption on a scale from 10, the best to 0, the worst. It can be seen from Fig. 6a that an increase with 1 from 2 to 3 is much more important than an increase from 7 to 8. This suggests that as long as a country's CPI is above 5, corruption is considered as "low" and a weak translation is seen towards the country's rating. As a country goes down the CPI scale[13] (lower than 5), the impact on the rating becomes very substantial.

The current account balance (% of GDP), which is used to discriminate between developed countries, sums up all cross-border transactions, including exports and imports of goods and services, net income revenues and net current transfers revenues. A current surplus indicates that the country has a net investor position vis-à-vis the rest of the world. A deficit indicates how much net import of capital from the rest of the world is required. As long as the current account balance for developed countries is positive, little effect is expected on the country's rating (see Fig. 6b). When the current account balance becomes negative, it indicates the country's increasing dependence for external or foreign capital. A current account balance below −3% to −5% is considered as an important deficit. Although some studies seem to agree that this variable is uncorrelated to a country's risk rating [8,19], this ratio is found to be significant here for developed countries only.

The cumulated debt forgiveness/GPD ratio is the (cumulative) amount of the external debt that has been let off by the foreign lenders. A low ratio is not considered to be very important, while a saturation applies when this ratio becomes high (see Fig. 6c). The interpretation is that the first initial debt forgiveness (likely as a result of a country's debt restructuring) for a country is penalized quite strong in the country's rating, while a higher forgiveness does not impact the rating any further. The variable can be interpreted as an indicator variable similar to the indicator variable indicating that the sovereign defaulted in the past [8].

The total debt stocks/current account ratio reflects the outstanding debt of a country compared to its revenues from goods and services. If this ratio becomes

[13] Countries with very low scores (2002 values) are, e.g., Argentina (2.8), Indonesia (1.9) and Venezuela (2.5). Some main European and North-American countries have the following scores: Canada (9.0), Finland (9.7), France (6.3), Germany (7.3), Italy (5.2), Spain (7.1) and U.S.A. (7.7).

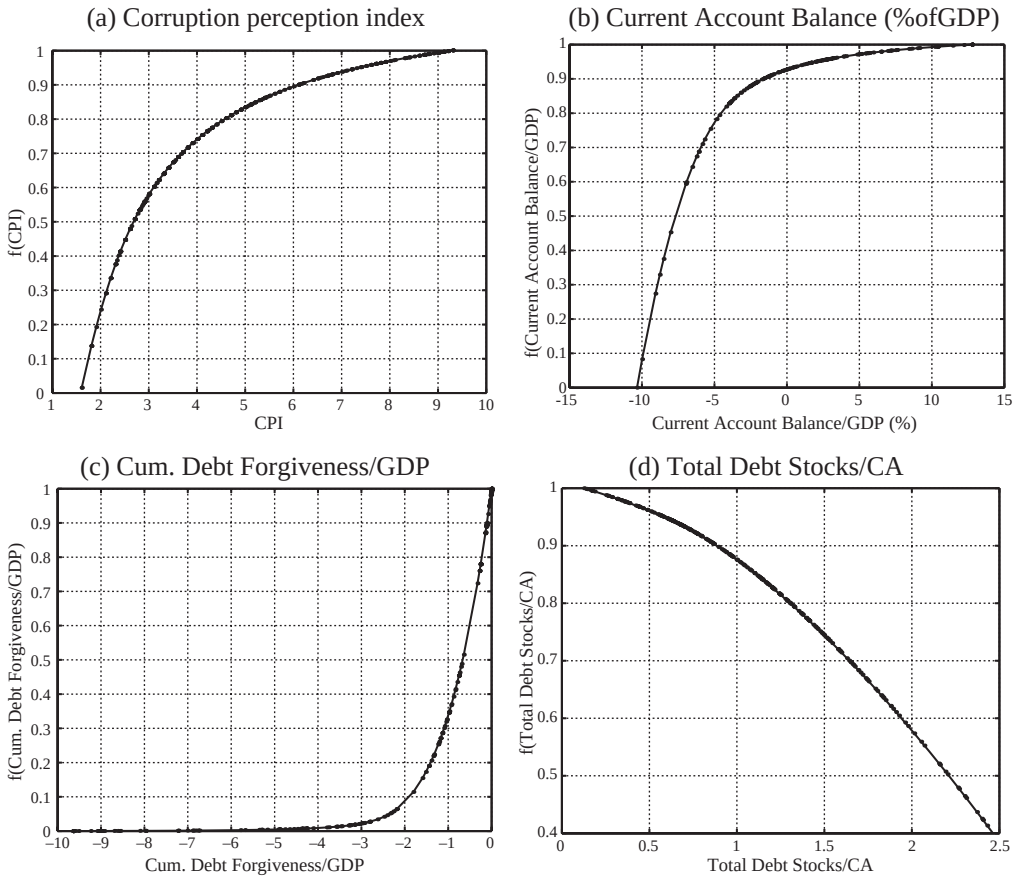


Fig. 6. Visualization of the identified univariate nonlinear transformations. Data points are denoted by the dots.

higher than 1, the country receives in general a higher penalization in its rating (see Fig. 6d) as this is generally considered as a weakness for a country's economy. A debt lower than the current account seems to be indifferent to the country's rating. External debt information is also used in the model of Cantor and Packer and was considered as an important predictor for the risk rating of a country [8]. Nonlinear relations between debt as a percentage of GDP and exports and growth are also reported in Ref. [20].

The column labelled Model 3 of Table 2 depicts the variables and the characteristics of the model estimated with the transformed inputs. The main difference is the removal of the school enrolment variable, which was previously discriminated for both developed and developing countries, but with a p-value close to 1% (Model 2). Likewise, for developed countries, gross domestic savings is no longer significant, while the current account balance is added to the model. For developing countries, the evolution of inflation becomes significant, while also the average level of inflation remains a significant discriminative variable for both developed and developing countries. The resulting 0–4 notches performances are reported in Section 3.2.2. When comparing these performances with the model without transformation, it can be clearly concluded that the performance improved.

3.2.2.4. Nonlinear SVM model for developed/developing countries (Model 4). In this step, the intrinsically linear model is extended with the SVM terms (as discussed in Section 2). We did not use the input subset that was identified using the previous model with the transformed inputs, but started from a set of candidate inputs suggested by the financial analyst. We used an RBF-kernel because of its good generalization capability [3,28]. The kernel parameter σ was selected from a grid $\Sigma = \sqrt{n} \times [0.8, 1, 1.2, 1.5, 2.5]$ using a cross-validation based tuning procedure. For each candidate σ-value, the eigenvalue decomposition of Eq. (17) is solved using Nyström sampling. The elements of the feature vector $\varphi(\boldsymbol{x})$ are then calculated from Eq. (18)

[24]. We start with 20 nonlinear transforms $\varphi_i(\boldsymbol{x})$, $i=\ldots 20$. Backward input selection is then applied to reduce the model complexity.

The selected model uses the following inputs: health expenditure (% of GDP) (T0 value), inflation (average) and mortality rate under −5 (last available, type 1). The resulting 0–4 notches performances are reported in Table 3 and contrasted with the results of the intrinsically linear model. The corresponding model deviances are equal to 1564 and 1542 for the intrinsically linear model without and with SVM terms, respectively, on a leave-one-out basis; and equal to 1822 and 1805 on a leave-country-out basis. Referring to Ref. [16], it can be concluded that the difference in deviance is significant.

3.2.3. Model evaluation

In addition to the general performance analysis reported above, it is also important to analyze how the external and internal ratings are distributed and how the performance varies across the rating classes. Fig. 7 represents the distribution of the assigned ratings and the target external ratings for the intrinsically linear model with SVM terms, for both performance criteria. It can be seen that in all cases, the distributions are very similar. The few mismatches are compensated one notch lower or higher. The mean rating (using numerical coding Aaa=1,...,B3=16,≤CCC=17) is equal to 8.5 (internal rating leave-one-out/leave-country-out) and 8.59 (Moody's long term rating), which is quite close.

Furthermore, the performance was also analyzed for different parts of the rating spectrum Aaa–Aa3, A1–A3, Baa1–Baa3, Ba1–Ba3, B1–CCC. In Table 4 the difference between the external rating y_{ext} and the predicted rating y_{pred} is compared for different values of the difference $y_{\text{ext}}-y_{\text{pred}}$. It is seen that most of the predicted observations are in the 2 notches difference range.

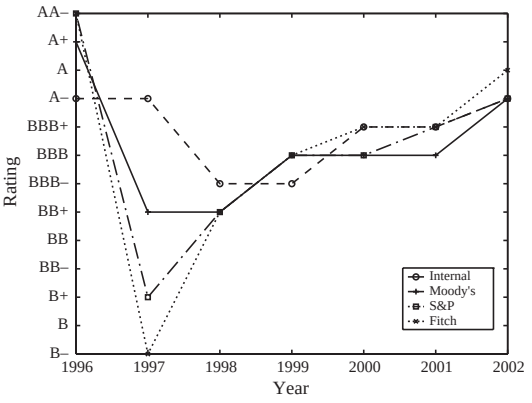


Fig. 8. Evolution of the ratings of South Korea from December 1996 to December 2002.

A gap analysis was performed to analyze the predicted ratings outside this range, revealing that most differences are due to missing data, local specificities like, e.g., Hong Kong and projection analysis. The latter will be included via the scenario-analysis module.

Besides the average rating performance, the obtained model should also be reactive on changes in the sense that a change in the financial and macro-economic situation of the country results into a timely change in the country rating. These rating changes were analyzed from a financial perspective by the financial analysts. The example of the rating evolution of South Korea is depicted here for illustrative reasons only in Fig. 8. The model has been built on external ratings of multiple years (1997–2002) to avoid that the model is too dependent on the year of the cycle it has been built. The prediction accuracy is also analyzed year by year and was found to be stable, yielding, e.g., yearly leave-one-out 2 notches performances ranging from 93.6% to 88.6%.

The model was built using the long term rating from Moody's. On the other hand, it is also interesting to see

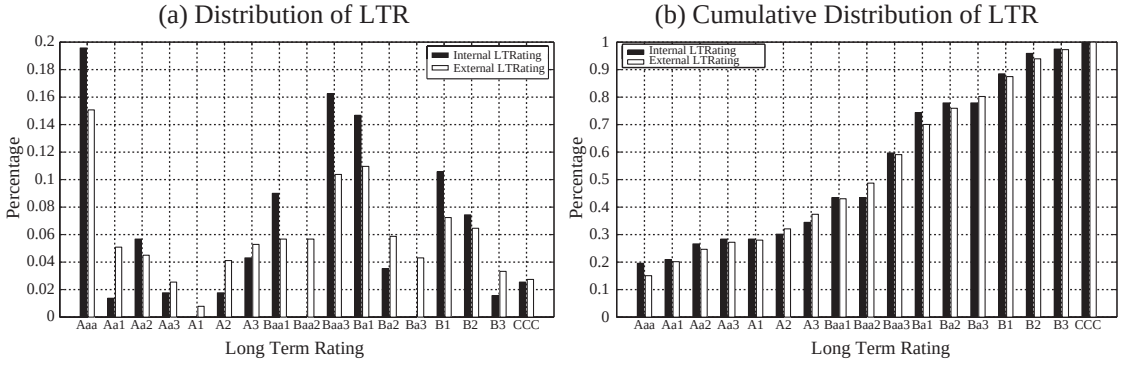


Fig. 7. Distribution and cumulative distribution of the internal and external long term rating (LTR).

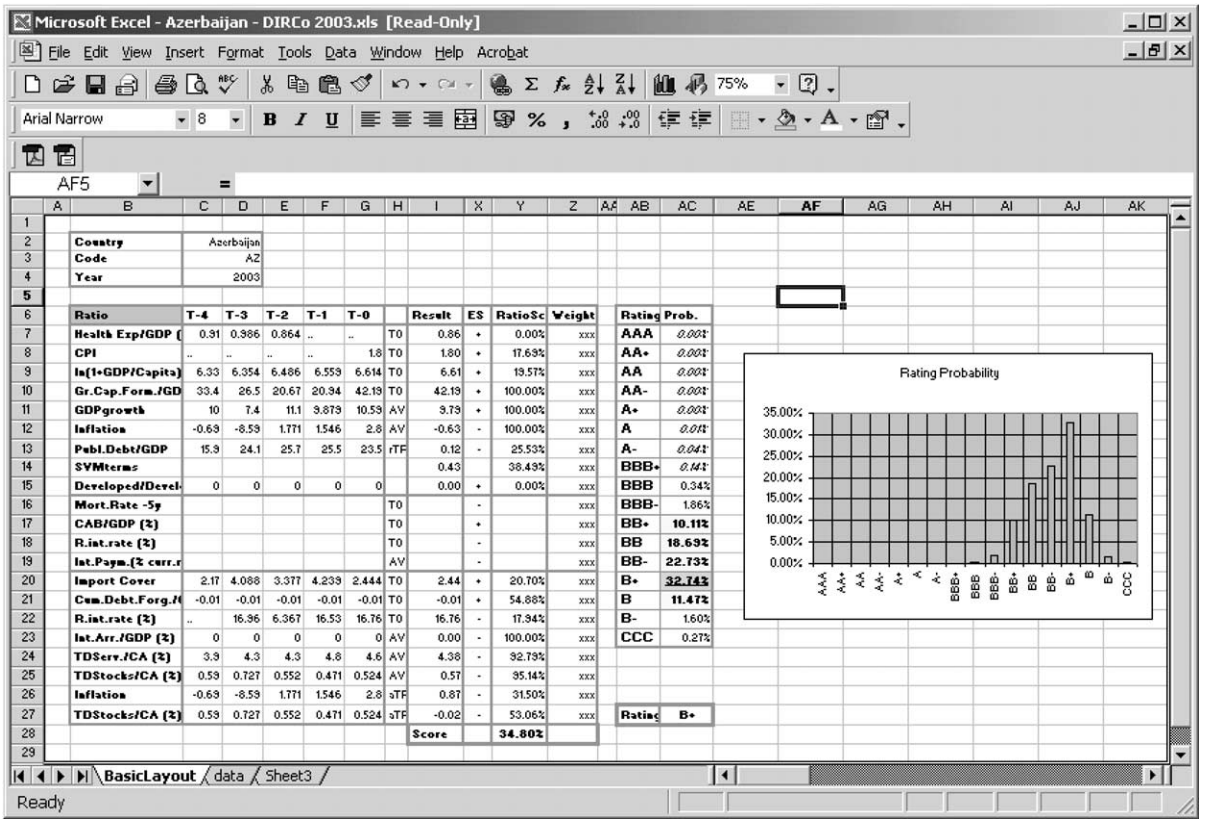


Fig. 9. Example screenshot of the Excel implementation.

how the model performs in the case of split ratings [4]. Therefore, the performance was also compared with the 3 rating agencies Fitch, Moody's and Standard & Poor's. The external rating interval was defined as the ratings in between the lowest and highest external rating. A zero notch difference is obtained when the internal rating is in the external rating interval. A one-notch difference is obtained when the internal rating is one notch outside the interval: one notch higher than the highest external rating or one notch lower than the lowest internal rating. Other rating differences are defined analogously. The obtained performances[14] are 50.68%, 81.02%, 95.11%, 99.22% and 99.80% (leave-country-out) and 56.95%, 84.74%, 96.67%, 99.61%, 99.80% (leave-one-out). These performances are good compared to the well-known pioneer and reference model [8] for Moody's model, which yields a 62% performance on 0–2 notches absolute difference, recognizing that this model was estimated on a much smaller database.

3.2.4. Scorecard development

An important aspect of the model application is a user-friendly and informative graphical-user interface that gives as much information as possible to the financial analysts rating the country. Therefore, the score function is scaled between 0% (bad) and 100% (good) in two steps. First, each of the ratios x_i is scaled into a ratio-score $x_{sc,i}$ between 0% and 100%, taking into account the sign of the coefficient. This yields an interpretable number that can also be viewed as a score that compares the country with the full database population. Secondly, these scaled ratios are used in the score function, where the coefficients are scaled appropriately.

For illustrative purposes, the following transformation is applied to the score function

$$z = |w_1|x_1 - |w_2|f_2(x_2) + f_{\mathrm{SVM}}(x_3, x_4), \tag{13}$$

where the absolute value of the coefficient is taken to indicate the sign of the true coefficient, a positive sign

[14] For the sake of completeness, it is mentioned that the rating agreements on 0, 0–1 and 0–2 notches differences on the considered database are the following: 57.7%, 89.2%, 97.38% (Fitch–Moody's); 59.1%, 94.6%, 99.7% (Fitch–S&P), 49.6%, 88.3%, 98.2% (Moody's–S&P).

indicating better creditworthiness. For each ratio i in the score function, the maximum M_i and minimum m_i are taken, e.g., $M_1=\max(x_1)$, $M_2=\max(f(x_2))$, $M_{\mathrm{SVM}}=\max(f_{\mathrm{SVM}})$. The ratios are then transformed to the ratio-scores as follows $x_1 \mapsto x_{\mathrm{sc},1}=(x_1-m_1)/(M_1-m_1)$, $f(x_2)\mapsto x_{\mathrm{sc},2}=(M_2-f(x_2))/(M_2-m_2)$ and $f_{\mathrm{SVM}}(x_3,x_4) \mapsto f_{\mathrm{sc,SVM}}=(f_{\mathrm{SVM}}(x_3,x_4)-m_{\mathrm{SVM}})/(M_{\mathrm{SVM}}-m_{\mathrm{SVM}})$, where the minimum or maximum is used in the numerator depending on the sign of the coefficients. Observe that the capping of the variables in Step 1 now receives a financial interpretation: above the upper capping, no more points are given/substracted.

Given the transformed ratios, the score function (13) is then translated into

$$z_{\mathrm{sc}} = \frac{1}{W}\big((|w_1|\times(M_1-m_1))x_{\mathrm{sc},1} + (|w_2|\times(M_2-m_2))x_{\mathrm{sc},2} + (M_{\mathrm{SVM}}-m_{\mathrm{SVM}})f_{\mathrm{sc,SVM}}\big),$$

with $W=|w_1|\times(M_1-m_1)+|w_2|\times(M_2-m_2)+(M_{\mathrm{SVM}}-m_{\mathrm{SVM}})$. The relative importance of, e.g., ratio x_1 is given by ratio weight $(|w_1|\times(M_1-m_1))/W$. The ideal counterparty that has 100% on all ratio scores receives value 1, while the worst possible counterparty receives a zero on all ratio-scores and a 0 on the resulting score.

Fig. 9 shows a screenshot of the Excel implementation of the country rating system, with data entry (columns C–G), variable and ratio-score calculation (column I and Y) and weights (column Z). The resulting score and rating are reported in cells Y28 and AC27. The corresponding rating probabilities (column AC and graph) are an indication on how sure the model is on the resulting rating and may help assist the analyst in the final rating decision.

4. Conclusions

The development of internal risk rating systems is becoming increasingly important in the context of the Basel II guidelines. In this paper, a process model to develop an internal rating system for country risk analysis is presented in which the different steps from data collection and preprocessing to model development and model implementation have been described and discussed in detail.

In the database construction and preprocessing step, it was discussed how the country risk data was collected from several types of financial databases. Furthermore, we also elaborated on how to create new more powerful predictors and how to deal with missing values and outliers. In the modelling step, we argued that, ideally, a risk rating system should be both accurate and readable, i.e. user-friendly and easy to understand for the financial expert. In order to achieve both these objectives, a gradual modelling approach was applied. First, an ordinal logistic regression model was formulated and estimated. Next, as debt information is not systematically available for developed countries, the model was extended with indicator variables such that the first part was used by both the developed and developing countries, the second part by the developed countries and the third part by developing countries only. As expected, the latter part of the model mainly consisted of debt variables. This model was then further optimized to an intrinsically linear model where advanced nonlinear transformations of the ratios were considered. Because of the readability requirement, the detected transformations were extensively studied with respect to their financial meaning and implications. Finally, the latter model was augmented in a new, gradual way with kernel based learning capability by adding Support Vector Machine terms to the model formulation. The SVM terms clearly improved the classification performance, although the readability of the model decreased to some extent. The intrinsically linear and SVM models were thoroughly evaluated. It was discussed how a user-friendly, easy to understand scorecard can be developed.

We would like to conclude by saying that the suggested process model is very generic in the sense that it can be easily applied in other risk assessment contexts such as rating corporates, banks, public sector entities or retail. However, the model is only a first step towards a full-fledged mature risk strategy, since other aspects such as loss given default and exposure at default clearly imply new modelling challenges that are interesting to address in future research.

Acknowledgement

All authors would like to thank Daniel Feremans, Daniel Saks, Mark Itterbeek, Frank Lierman (Dexia Bank); Luc Leonard, Eric Hermann (Dexia Group) and Jos De Brabanter (Katholieke Universiteit Leuven) for the many helpful comments. Johan Suykens acknowledges support from K.U.Leuven, IUAP V, GOA-MEFISTO 666 and FWO project G.0407.02.

Appendix A. Estimation of univariate nonlinear transformation

The following type of transformation is considered: $x\mapsto f(x+c,\lambda)$; with location parameter c and transfor-

ARTICLE IN PRESS

mation parameter λ. The location parameter is introduced so as to shift the distribution to the appropriate part of the nonlinear function $f(\cdot, \lambda)$ defined in Eq. (7). The parameters c and λ for a given ratio x_i are inferred from the data as follows. Step 1: The ratio x_i is standardized to zero median and unit variance [5]. Step 2: Given the already identified nonlinearities and the current input set, the additional nonlinear transformation is estimated using a simple grid search mechanism similar to the one applied in Ref. [28]. In this grid, the parameter c varies from -3 to $+3$ and the parameter λ from -2 to $+2$. For each hyperparameter combination (c, λ), the model was estimated and its deviance stored. Step 3: The combination (c, λ) having the lowest deviance is selected. The optimal deviance is compared with the deviance obtained with $\lambda = 1$. When the deviance of the nonlinear model is 10 or lower than the deviance of the model with linear term [16], the nonlinear transformation is applied, given that the cross-validation performance is satisfactory and the transformation is financially meaningful.

Appendix B. Support Vector Machines

For the sake of completeness, the (primal) feature space formulation for SVMs is given. It is illustrated how the corresponding dual optimization problem allows to estimate and evaluate the classifier in terms of the kernel function. The estimation of an explicit expression for the nonlinear mapping is also given.

B.1. Primal–dual formulations

Consider a training set of N data points $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^{N}$, with input data $\boldsymbol{x}_i \in \mathcal{R}^n$ mapped into the feature space $\boldsymbol{\varphi}(\boldsymbol{x}_i) \in \mathcal{R}^{n_\varphi}$ and corresponding binary class labels $y_i \in \{-1, +1\}$. When the data of the two classes are separable (Fig. 10a), one can say that $\boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x}_i) + b \geq +1 (y_i = +1)$ and $\boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x}_i) + b \leq -1 (y_i = -1)$. This set of two inequalities can be combined into one single set as follows

$$y_i\left(\boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x}_i) + b\right) \geq +1, \qquad i = 1, \ldots, N. \tag{14}$$

As can be seen from Fig. 10a, from the multiple solutions possible, the solution with largest margin $2/\|\boldsymbol{w}\|_2$ yields the best generalization.

In most practical, real-life classification problems, the data are non-separable in linear or nonlinear sense, due to the overlap between the two classes (see Fig. 10b). In such cases, one aims at finding a classifier that separates the data as much as possible. The SVM classifier formulation (14) is extended to the non-separable case by introducing slack variables $\xi_i \geq 0$ in order to tolerate misclassifications [29]. The inequalities in Eq. (14) are changed into

$$y_i\left(\boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x}_i) + b\right) \geq 1 - \xi_i, \qquad i = 1, \ldots, N. \tag{15}$$

In the primal weight space, the optimization problem becomes

$$\min_{\boldsymbol{w}, b, \boldsymbol{\xi}} \mathcal{J}_P(\boldsymbol{w}) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + c\sum_{i=1}^{N}\xi_i \text{ such that}$$
$$y_i\left(\boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x}_i) + b\right) \geq 1 - \xi_i \text{ and } \xi_i \geq 0, \quad i = 1, \ldots, N, \tag{16}$$

where c is a positive real constant that determines the trade-off between the large margin term $1/2\boldsymbol{w}^T\boldsymbol{w}$ and error term $\sum_{i=1}^{N}\xi_i$ that aims at minimizing the training set error in the non-separable case.

SVMs are modelled within a context of convex optimization theory [24,29]. The general methodolo-

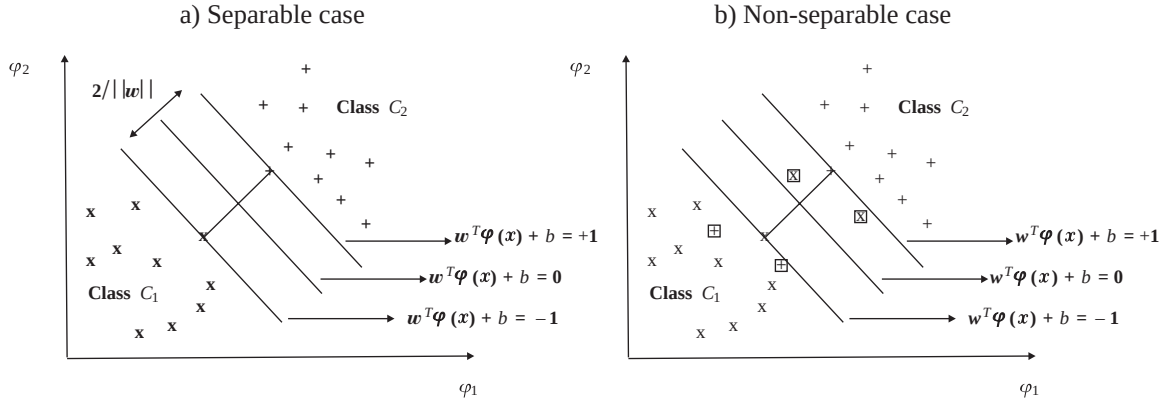


Fig. 10. Illustration of SVM classification in two dimensions (φ_1, φ_2) of the feature space. Left: separable case (margin $2/\|\boldsymbol{w}\|$); right: non-separable case.

gy is to start formulating the problem in the primal weight space as a constrained optimization problem, next formulate the Lagrangian, take the conditions for optimality and finally solve the problem in the dual space of Lagrange multipliers, which are also called support values. The Lagrangian is equal to $\mathcal{L} = 0.5\boldsymbol{w}^T\boldsymbol{w} + c\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i\left(y_i\left(\boldsymbol{w}^T\varphi(\boldsymbol{x}_i)+b\right) - 1 + \xi_i\right) - \sum_{i=1}^{N}v_i\xi_i$, with Lagrange multipliers $\alpha_i \geq 0$, $v_i \geq 0$ $(i=1,\ldots,N)$. The solution is given by the saddle point of the Lagrangian $\max_{\alpha,\nu}\min_{\boldsymbol{w},b,\xi}\mathcal{L}(\boldsymbol{w},b,\boldsymbol{\xi};\boldsymbol{\alpha},\boldsymbol{\nu})$, with conditions for optimality:

$$\frac{\partial\mathcal{L}}{\partial\boldsymbol{w}}\rightarrow\boldsymbol{w}=\sum_{i=1}^{N}\alpha_i y_i\varphi(\boldsymbol{x}_i),\frac{\partial\mathcal{L}}{\partial b}\rightarrow\sum_{i=1}^{N}\alpha_i y_i = 0$$

$$\text{and}\frac{\partial\mathcal{L}}{\partial\xi_i}\rightarrow 0\leq\alpha_i\leq c,\qquad i=1,\ldots,N.$$

Given Eq. (16), this yields the following dual QP-problem

$$\max_{\boldsymbol{a}}\mathcal{J}_D(\boldsymbol{a}) = -\frac{1}{2}\sum_{i,j=1}^{N}y_i y_j\varphi(\boldsymbol{x}_i)^T\varphi(\boldsymbol{x}_j)\alpha_i\alpha_j + \sum_{i=1}^{N}\alpha_i$$

$$= -\frac{1}{2}\boldsymbol{a}^T\mathbf{D}_y\Omega\mathbf{D}_y\boldsymbol{a} + 1^T\boldsymbol{a}$$

$$\text{such that}\sum_{i=1}^{N}\alpha_i y_i = 0\text{ and }0\leq\alpha_i\leq c,\quad i=1,\ldots,N,$$

with the vectors $\boldsymbol{a} = [\alpha_1,\ldots,\alpha_N]^T$, $\boldsymbol{1} = [1,\ldots,1]^T \in\mathcal{R}^N$, $\boldsymbol{y} = [y_1,\ldots,y_N]^T\in\mathcal{R}^N$, the diagonal matrix $\mathbf{D}_y = \mathrm{diag}(\boldsymbol{y})\in\mathcal{R}^{N\times N}$ and the positive (semi-) definite kernel matrix $\Omega\in\mathcal{R}^{N\times N}$:

$$\Omega = \begin{bmatrix} K(\boldsymbol{x}_1,\boldsymbol{x}_1) & K(\boldsymbol{x}_1,\boldsymbol{x}_2) & \ldots & K(\boldsymbol{x}_1,\boldsymbol{x}_N)\\ \vdots & \vdots & \ddots & \vdots\\ K(\boldsymbol{x}_N,\boldsymbol{x}_1) & K(\boldsymbol{x}_N,\boldsymbol{x}_2) & \ldots & K(\boldsymbol{x}_N,\boldsymbol{x}_N)\end{bmatrix}\in\mathcal{R}^{N\times N}.$$

The bias term b is obtained as a by-product of the QP-calculation or from a non-zero support value. More generally, one obtains other SVM formulations, e.g., for least squares and logistic regression using the same methodology [24,29].

B.2. Estimation of nonlinear mapping

Given the data points $\{\boldsymbol{x}_i,\ldots,\boldsymbol{x}_N\}$ and the kernel function K, one can estimate the nonlinear mapping $\varphi(\boldsymbol{x})$ based on the eigenvalue decomposition of the kernel matrix

$$\Omega = \mathbf{U}\Upsilon\mathbf{U}^T \tag{17}$$

with $\mathbf{U} = [\boldsymbol{u}_1,\ldots,\boldsymbol{u}_N]\in\mathcal{R}^{N\times N}$ and $\mathbf{Y} = diag([v_1,\ldots,v_N])\in\mathcal{R}^{N\times N}$. The elements φ_i of the mapping $\varphi=[\varphi_1,\ldots,\varphi_{nf}]^T$ are estimated as follows [24]

$$\varphi_i(\boldsymbol{x}) = \frac{\sqrt{N}}{\sqrt{v_i}}\sum_{k=1}^{N}u_{ki}K(\boldsymbol{x}_k,\boldsymbol{x}),\qquad i=1,\ldots,N, \tag{18}$$

and $\varphi_i(\boldsymbol{x})=0$ for $v_i=0$ or $i\geq N+1$. Using this estimate, it is easy to see that $\varphi(\boldsymbol{x}_i)^T\varphi(\boldsymbol{x}_j)=K(\boldsymbol{x}_i,\boldsymbol{x}_j)$ for $i,j=1,\ldots,N$.

For large data sets, the computational requirements may become too high. The idea of Nyström sampling [24] is to estimate $\varphi(\cdot)$ on a (carefully) selected subsample of size $M\leq N$ from the data $\{\boldsymbol{x}_i\}_{i=1}^{N}$. In fixed size Least Squares Support Vector Machines, the Renyi entropy measure is used to select the sub-sample [24]. In this paper, one observation[15] of each country was selected in the sub-sample. A similar solution as Eq. (18) is obtained with $M\leq N$ non-zero components. The computational and memory requirements reduce from $O(N^3)$ to $O(M^3)$, and from $O(N^2)$ to $O(M^2)$, respectively.

References

[1] E.L. Altman, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, Journal of Finance 23 (1968) 589–609.
[2] B. Baesens, R. Setiono, C. Mues, J. Vanthienen, Using neural network rule extraction and decision tables for credit-risk evaluation, Management Science 49 (3) (2003) 312–329.
[3] B. Baesens, T. Van Gestel, S. Viaene, M. Stepanova, J. Suykens, Bench-marking state of the art classification algorithms for credit scoring, Journal of the Operational Research Society 54 (6) (2003) 627–635.
[4] Basel Committee on Banking Supervision, International convergence of capital measurement and capital standards, BIS (2004), 239 pp.
[5] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
[6] G.E.P. Box, D.R. Cox, An analysis of transformations, Journal of the Royal Statistical Society Series B 26 (1964) 211–243.
[7] J.C. Cadwell, Routes to low mortality in poor countries, Population and Development Review 12 (2) (1986) 171–220.
[8] R. Cantor, F. Packer, Determinants and impacts of sovereign credit ratings. Federal Reserve Bank of New York, Economic Policy Review 2 (1996) 37–55.
[9] L.S. Chakraborty, Public expenditure and human development: an empirical investigation, Wider International Conference on Inequality, Poverty and Human Well-Being, 2003, Helsinki, May 30–31.

[15] For reporting the leave-one-out and leave-country-out performance, the corresponding observation is also left out for the estimation of φ. In this way, it is avoided that country information on the left out observation would still be present in the model when evaluating the model on the country.

ARTICLE IN PRESS

[10] S.P. Chattopadhyay, Neural network approach for assessing country risk for foreign investment, International Journal of Management 14 (2) (1997) 159–167.
[11] R.A. Eisenbeis, Pitfalls in the application of discriminant analysis in business, finance and economics, Journal of Finance 32 (3) (1977) 875–900.
[12] D. Filmer, L. Pritchett, Child mortality and public spending on health: how much does money matter?, Technical Report 1864, World Bank — Country Economics Department, 1997.
[13] D.J. Hand, W.E. Henley, Statistical classification methods in consumer risk: a review, Journal of the Royal Statistical Society, Series A 160 (1997) 523–541.
[14] N.U. Haque, M.S. Kumar, N. Mark, D.J. Mathieson, The economic content of indicators of developing country creditworthiness, IMF Staff Papers 43 (1996) 688–724.
[15] S. Hoti, M. McAleer, An empirical assessment of country risk ratings and association models. Journal of Economic Surveys 18 (4) (2004) 539–588.
[16] H. Jeffreys, Theory of Probability, Oxford University Press, 1961.
[17] P. McCullagh, Regression models for ordinal data, Journal of the Royal Statistical Society, Series B (Methodological) 42 (2) (1980) 109–142.
[18] J.A. Ohlson, Financial ratios and the probabilistic prediction of bankruptcy, Journal of Accounting Research 18 (1980) 109–131.
[19] M. Oral, O. Kettani, J.C. Cosset, M. Daouas, An estimation model for country risk ratings, International Journal of Forecasting 8 (1992).
[20] C. Patillo, H. Poirson, L. Ricii, External debt and growth, IMF Working Paper, vol. 69, IMF, 2002.
[21] S. Piramuthu, H. Ragavan, M.J. Shaw, Using feature construction to improve the performance of neural networks, Management Science 44 (3) (1998) 416–430.
[22] P. Rivoli, T.L. Brewer, Political instability and country risk, Global Finance Journal 8 (1997) 309–321.
[23] A. Sorkin, K. Summers, I. Farquhar, Investing in Health: The Social and Economic Benefits of Health Care Innovation, Elsevier, 2001.
[24] J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, Least Squares Support Vector Machines, World Scientific, Singapore, 2002.
[25] R.J. Taffler, B. Abassi, Country risk: a model for predicting debt servicing problems in developing countries, Journal of the Royal Statistical Society Series A 147 (4) (1984) 541–568.
[26] L.C. Thomas, D.B. Edelman, J.N. Crook, Credit scoring and its applications, SIAM Monographs on Mathematical Modeling and Computation, Philadelphia, U.S. 2002
[27] T. Van Gestel, J.A.K. Suykens, G. Lanckriet, A. Lambrechts, B. De Moor, J. Vandewalle, A Bayesian framework for least squares support vector machine classifiers, Neural Computation 15 (4) (2002) 1115–1147.
[28] T. Van Gestel, J. Suykens, B. Baesens, S. Viaene, J. Vanthienen, G. Dedene, B. De Moor, J. Vandewalle, Benchmarking least squares support vector machine classifiers, Machine Learning 54 (2004) 5–32.
[29] V. Vapnik, Statistical Learning Theory, John Wiley, New York, U.S., 1998.
[30] I.K. Yeo, R.A. Johnson, A new family of power transformations to improve normality or symmetry, Biometrica 87 (2000) 949–959.

Tony Van Gestel obtained his electromechanical engineering degree and his Ph.D. degree in Applied Sciences (subject: Mathematical Modelling for Financial Engineering) in 1997 and 2002 at the Katholieke Universiteit Leuven. His work and research focuses on linear and non-linear mathematical modelling for financial risk management and financial engineering in general. He has co-authored the book "*Least Squares Support Vector Machines* (World Scientific Press)" and has published papers in about 25 journal articles in this field. He also serves as a referee and guest editor for several international journals. He currently works as a senior quantitative analyst at Dexia Group focusing on the development, backtesting, validation and implementation of rating and risk systems for Basel II. In his free time, he continues with his academic research.

Bart Baesens was born in Bruges, Belgium, in February 27, 1975. He received his Ph.D. degree in Applied Economic Sciences from the Katholieke Universiteit Leuven in 2003. He is an assistant professor (lecturer) at the School of Management of the University of Southampton (United Kingdom). He has done extensive research on credit scoring and data mining. His findings have been published in well-known international journals and presented at international top conferences. He also frequently serves as a reviewer and guest editor for several international journals. He regularly tutors, advices and provides consulting support to international financial institutions with respect to their credit risk management and credit scoring policy.

Peter Van Dijcke was born in Opbrakel, Belgium, in March 2, 1966. He received his degree in Commercial Engineering and an MBA degree from the Katholieke Universiteit Leuven, in 1988 and 1989, respectively. From 1989 to 1991, he was a Research Assistent in Managerial Economics at the Katholieke Universiteit Leuven. Between 1991 and 1998, he worked as an Economic Advisor at, respectively, the Belgian Saving Banks Associations, the Belgian Banking Association and the Federation of Coordination Centers (Forum 187). As of 1998, he has been working for Dexia Bank Belgium, first as Senior Economist in the Research Departmant and, since 2004, as Project Leader for the "Company Project". He received a Robert Schuman scholarship in 1992 and the SUERF Marjolin prize for best paper at the Vienna International Colloquium in 2000.

João B.C. Garcia is a Senior Quantitative Analyst at the Credit Methodology team in Dexia Group in Brussels. His current interest includes credit derivatives, structured products and credit risk models for determining economic capital. Prior to this position, he has worked as a Quantitative Analyst in Artesia Banking Corporation in Brussels modelling exotic interest rate derivatives. He is an Electronic Engineer from the Instituto Tecnologico de Aeronautica (ITA-Brazil), holds an M.Sc. in Physics from the UFPe-Brazil and a Ph.D. in Physics from the University of Antwerpen (UIA-Belgium).

Johan A.K. Suykens was born in Willebroek, Belgium, in May 18, 1966. He received a degree in Electro-Mechanical Engineering and his Ph.D. degree in Applied Sciences from the Katholieke Universiteit Leuven, in 1989 and 1995, respectively. In 1996, he has been a Visiting Postdoctoral Researcher at the University of California, Berkeley. He has been a Postdoctoral Researcher with the Fund for Scientific Research FWO Flanders and is currently an Associate Professor with K.U. Leuven. His research interests are mainly in the areas of the theory and application of neural networks and nonlinear systems. He is author of the books "*Artificial Neural Networks for Modelling and Control of Non-linear Systems*" (Kluwer

ARTICLE IN PRESS

Academic Publishers) and "*Least Squares Support Vector Machines*" (World Scientific), co-author of the book "*Cellular Neural Networks, Multi-Scroll Chaos and Synchronization*" (World Scientific) and editor of the books "*Nonlinear Modeling: Advanced Black-Box Techniques*" (Kluwer Academic Publishers) and "*Advances in Learning Theory: Methods, Models and Applications*" (IOS Press). In 1998, he organized an International Workshop on Nonlinear Modelling with Time-series Prediction Competition. He has served as associate editor for the *IEEE Transactions on Circuits and Systems-I* (1997–1999) and since 1998, he is serving as associate editor for the *IEEE Transactions on Neural Networks*. He received an IEEE Signal Processing Society 1999 Best Paper (Senior) Award and several Best Paper Awards at International Conferences. He is a recipient of the International Neural Networks Society INNS 2000 Young Investigator Award for significant contributions in the field of neural networks. He has served as Director and Organizer of a NATO Advanced Study Institute on Learning Theory and Practice taking place (Leuven 2002) and as a program co-chair for the International Joint Conference on Neural Networks IJCNN 2004.

Jan Vanthienen received a degree in Applied Economics and Information Systems in 1979 and his Ph.D. degree in Applied Economics in 1986, both from Katholieke Universiteit Leuven. He is currently full professor of information systems at Katholieke Universiteit Leuven, Department of Decision Sciences and Information Management. His current research interests include information and knowledge management, business intelligence and business rules, information systems analysis and design. He is a founding member of the Leuven Institute for Research in Information Systems (LIRIS), and a member of the ACM and the IEEE Computer Society. He has published more than 100 refereed full papers in reviewed international journals and conference proceedings. He is chairholder of the Pricewaterhouse Coopers Chair on E-Business at K.U. Leuven. In the past, he was co-chair of the European Conference on Verification and Validation of Knowledge Based Systems (EuroVaV 97).