A Support Vector Machine Approach to Credit Scoring

Dr. Ir. Tony Van Gestel(1), Bart Baesens(2), Dr. Ir. Joao Garcia(1), Peter Van Dijcke(3)

(1) Credit Methodology, Global Market Risk, Dexia Group, {tony.vangestel,joaobatista.crispinianogarcia}@dexia.com
(2) Dept. of Applied Economic Sciences, K.U.Leuven, bart.baesens@kuleuven.ac.be
(3) Research, Dexia Bank Belgium, peter.vandijcke@dexia.be

Abstract:

Driven by the need to allocate capital in a profitable way and by the recently suggested Basel II regulations, financial institutions are being more and more obliged to build credit scoring models assessing the risk of default of their clients. Many techniques have been suggested to tackle this problem. Support Vector Machines (SVMs) is a promising new technique that has recently emanated from different domains such as applied statistics, neural networks and machine learning. In this paper, we experiment with least squares support vector machines (LS-SVMs), a recently modified version of SVMs, and report significantly better results when contrasted with the classical techniques.

Keywords:

Basel II, Internal Rating Based System, credit scoring, Support Vector Machines

I. Introduction

One of the major goals of the framework recently put forward by the Basel Committee on Banking Supervision [5] is to use risk based approaches to allocate and charge bank capital. In order to minimize capital requirements and exposure, financial may choose to use the Internal Rating Based (IRB) approach to calculate the risk of their credit portfolios. Essentially, the IRB approach implies that a bank gives a score to every credit in its portfolio, which reflects its view on the default probability of the underlying issuer. This default assessment is then evaluated together with the bank's internal risk system so as to allocate economic capital in case of default of the issuer. Those internal risk computations are also used by the management in order to evaluate whether the taken credit exposures are desirable and acceptable, given the economic capital that they consume.

As the internal credit scoring system becomes a key building block of the capital allocation strategy of a financial institution, several modelling techniques have been and are being developed to perform credit scoring. Beaver [6] investigated the use of a single financial ratio to predict bankruptcy. Altman [2] applied Fisher Discriminant Analysis (FDA) to find an optimal linear combination of financial ratios to discriminate between bankrupt and non-bankrupt firms. Logistic regression was first studied by Ohlson [11] and has become, together with FDA, an often used benchmark technique when evaluating the performance of more recent and advanced classification techniques [4], like k-nearest neighbors, decision trees, neural networks and Support Vector Machines (SVMs). Especially the latter two have already been successfully validated and implemented as illustrated by the reports of rating agencies like Moody's [10] and S&P [7].

In this paper, we report on the use of Least Squares SVMs [12] (LS-SVMs) for credit rating of banks[1]. As the number of defaults is rather low, we focus on learning the rating scores from given training data of the rating agencies. As we will be dealing with multiclass problems (>2 classes) in which an order exists among the classes considered, we contrast the test set performance of the LS-SVMs with that of Ordinary Least Squares (OLS) regression, Ordinal Logistic Regression (OLR) and Multilayer Perceptrons (MLPs). For this purpose, the data set is split up into two sets: 1) one used to train and design the system, generating the model parameters; 2) one test set used for out of sample validation, testing the goodness of fit of the model.

This paper is organized as follows. Section 2 gives a brief review of the credit rating methodologies used here and of the LS-SVMs in particular. The dataset is described in Section 3 and the empirical test set results are compared and discussed in Section 4.

II. Methodology

The basic underpinnings of linear and logistic regression for bankruptcy prediction and credit rating are reviewed first. Then the LS-SVM formulation is discussed. Mathematical details can be found in the references.

---

[1] A similar research has been conducted for countries of which the main results are available upon request.

II.1. Linear least squares and logistic regression

Altman used a linear combination $z = w_1 x_1 + w_2 x_2 + \ldots + w_n x_n$ of n financial ratios $x_1, \ldots, x_n$ to discriminate between solvent and non-solvent firms. Starting from an initial set of candidate financial ratios, Altman selected the following 5 financial variables: Working Capital/Total Assets, Retained Earnings/Total Assets, Market Value of Equity/Total Debt and Sales/Total Assets [2]. The optimal linear combination is obtained by constructing a representative training data set and by applying Fisher Discriminant Analysis (FDA), which is closely related to other statistical techniques like Linear Discriminant Analysis, Canonical Correlation Analysis and Ordinary Least Squares regression. An intuitive explanation of FDA is depicted in Figure 1. Firms with a z-value above a threshold are classified as solvent; otherwise they are classified as non-solvent. Instead of making a black and white decision, one may also relate a default probability to the z-value. Indeed, high values correspond to firms with a good financial strength, while low z values are more likely to indicate financial distress; thus, the lower the z-value, the larger the default probability. Note that this probability is often translated into alphabetical ratings.

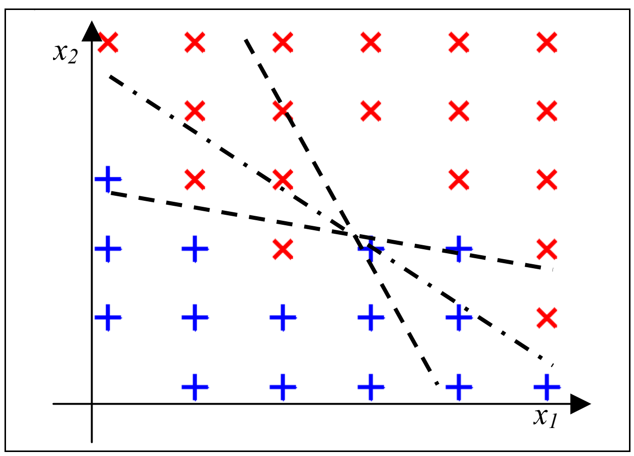


Figure 1: Graphical illustration of Fisher Discriminant Analysis: given a set of training data points, FDA analysis aims at finding a linear separating hyperplane (linear boundary in the two-dimensional case) that separates both classes. It is seen that the dashed-dotted line yields a better separation than the two dashed lines.

In logistic regression [11], one directly optimizes the probability to a given data set instead of solving a least squares problem. It can be considered as an alternative technique to estimate the linear combination of ratios and is known to be more robust to the assumptions made in FDA and OLS.

When the number of well-documented default firms is rather low, it becomes difficult to retrieve a representative training set database from which the binary classifier can be estimated. Instead, one constructs a database with the financial strength ratings provided by the rating agencies and builds a rating system based on these data. In Ordinary Least Squares regression for credit scoring [9] one aims at finding a linear combination $z=w_1 x_1 + w_2 x_2 + \ldots + w_n x_n + b$ of the input variables x_i, $(i=1,\ldots n)$, such that the latent variable z is as close as possible to the target variable t corresponding to the ratings, which are nominally coded as t=1 (A+), t=2 (A), t=3 (A-), …, t=15 (E-). In the two class case, the above formulation corresponds to FDA.

The binary logistic regression formulation is extended to the multiclass categorization problem in which the classes vary in an ordinal way from very good credibility to very weak financial strength. In the ordinal logistic regression (OLR) model [1], a formula for the cumulative probability is presented and, instead of minimizing a least squares cost function, one optimizes directly the probability, resulting in the parameter estimates.

II.2. Least Squares Support Vector Machines

Neural networks can be interpreted as mathematical representations inspired by the functioning of the human brain. The Multilayer Perceptron (MLP) neural network is a popular neural network for both regression and classification and has often been used for bankruptcy prediction and credit scoring in general, see, e.g., the references in [3]. Although there exist good training algorithms, like Bayesian inference, to design the MLP, there are a number of drawbacks like the choice of the architecture of the MLP and the existence of multiple local minima, which implies that the results may not be reproducible.

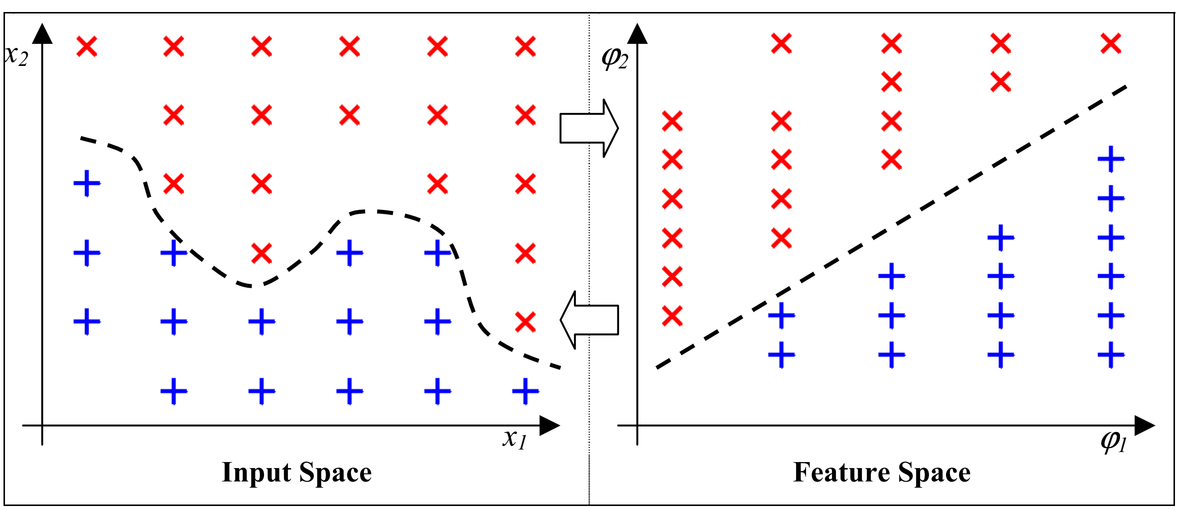


Figure 2:Graphical illustration of nonlinear classification with LS-SVMs: given a two-class problem in the input space (x_1 and x_2), the data points are first mapped to a higher dimensional feature space (denoted by φ_1 and φ_2 in the figure for convenience, but typically of much higher dimensionality). In this high dimensional feature space, a linear classifier is constructed using, e.g., FDA. This linear classifier in the feature space corresponds to a nonlinear classifier in the input space, for which a practical expression is obtained in terms of the kernel function (that implicitly defines the nonlinear mapping).

More recently, a new type of learning methodology emerged, called Support Vector Machines [15] . Basically, this technique can be understood as follows: in a first step, the inputs are mapped or transformed in a nonlinear way to a high dimensional feature space, in which a linear classifier is constructed (like, e.g., FDA) in the second step. Besides the importance of using a regularization term to improve out of sample accuracy, a key element is that the nonlinear mapping is only implicitly defined in terms of a kernel function, also known in applied statistics. Hence, the SVMs and kernel based algorithms in general can be understood as applying a linear technique in a feature space that is obtained via a nonlinear preprocessing step (see Figure 2). An important advantage is that the solution follows from a convex optimisation problem yielding a unique solution, solving the reproducibility problem of MLPs. In this paper, we applied Least Squares Support Vector Machines [12], which correspond to a kernel version of Fisher Discriminant Analysis for binary classification (two class problems) [13]. In order to solve the multi-class problem, one uses an idea from information theory and statistics [8]. The multi-class problem is represented by a set of paired comparisons: an LS-SVM classifier between each pair of rating categories is constructed and the evaluation is done by a voting system in which each LS-SVM gives his vote and the rating with the highest number of votes is selected. More details on LS-SVMs can be found in [13] and [14] .

III. Dataset Description

The rating problems concern the scoring of banks into the credit rating categories A+, A, A-, B+, …, E+, E and E-. The bank dataset was retrieved from BankScope and consists of 8 year data (1995-2002) of Moody's financial strength rating for 831 financial institutions.

The dependent variable or output y is the financial strength rating on May of the next year and is ordinally coded via A+ = 1, A = 2, A+ = 3, …, E- = 15. For the independent variables or explanatory inputs, we used 36 predefined BankScope ratios, 14 calculated variables and three size variables (see appendix 1). For each of these input variables, also the information of the previous year is used to detect possible trends in the evolution, see also Figure 3 for a schematic representation. A nominally coded region variable with 5 regional values is also included and coded by using 4 dummy variables. The resulting database consisted of 3972 observations, where each observation consists of the rating, a unique bank identifier, the year and the 110 candidate explanatory variables.

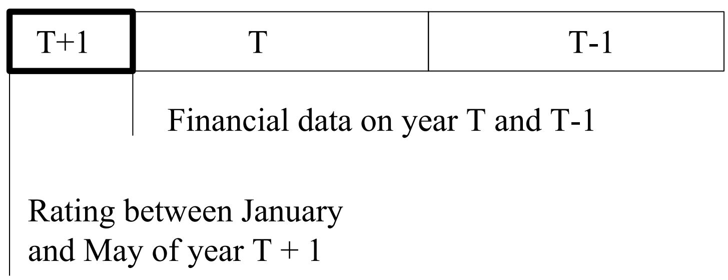


Figure 3: Prediction of financial strength of the rating in January-May of the year T+1 given the financial data of year T and T-1.

In the data cleaning and data preprocessing stage, both inputs with too many missing values and observations with too many missing inputs were removed from the database. All other missing inputs were replaced by the median. This resulted in a clean database of 3599 observations and 79 candidate inputs. The resulting scale variables were preprocessed as follows:

1. Ratios were normalized to zero median and unit variance (measured by the interquartile range (IQR): $\hat{s}$ = IQR/(2x0.6745)). Positive and negative outliers outside the 3$\hat{s}$-range were set to these limits in the same way as in the Winsorized mean calculation.
2. The log-transform[2] was applied to positive size related variables like, e.g., total assets.

The numbers of observations per rating category are reported in Figure 4, from which it is seen that the number of observations in the minus categories is rather limited. Also the number of banks in the highest and lowest categories is low, which may cause problems for some of the models as their main focus will be on classifying the middle rating categories correctly.

The four models (OLS, OLR, MLP and LS-SVM) were designed on the training set containing 2/3 of the database. The remaining 1/3 of the data was used for the out of sample test set: this data is not used during the training and, hence, is new for all models allowing for a fair apple-to-apple comparison. In order to make the training and test set representative, they were selected

[2] Two subsequent Yeo-Johnson transformations (Biometrica, 87:954-959, 2000) were applied to size related variables with both positive and negative values (e.g., total profit).

using stratified sampling such that per rating category 2/3 of the data is used for training and 1/3 for out of sample testing.

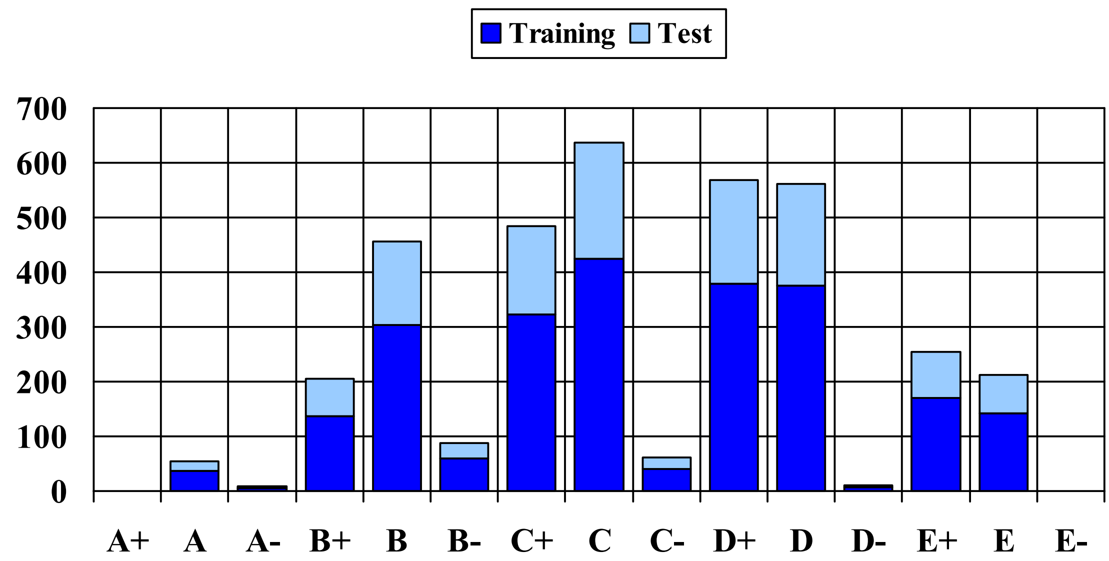


Figure 4: Number of observations per rating category in training and hold out test set. The split up into training and test set is done via stratified sampling, allowing for a simple and straightforward comparison of the evaluated models. Also observe that the minus ratings are systematically underrepresented.

IV. Evaluation

The main purpose of the underlying rating based system was to make accurate decisions for each type of rating category[3] (A+, …, E-). For the evaluation of the rating models, the following concerns are taken into account:

(1) The number of differences between the rating of the system and of the rating agency should be small.

(2) In a quantitative stage, it could be considered that the analyst is allowed to adjust the rating by one or two notches based upon his personal experience. Hence, when interpreting an "error" as a different rating, the error becomes more important with increasing notch difference. Therefore, the x-notch difference accuracy is the percentage of correct ratings when allowing up to x notches differences between the ratings.

---

[3] Additionally cost functions could be introduced in a later stage in order to adapt scoring precision of the model.

(3) The system should perform in a similar way on all categories, i.e., significant underperformance in some categories is not allowed to be compensated by better performance in other categories.

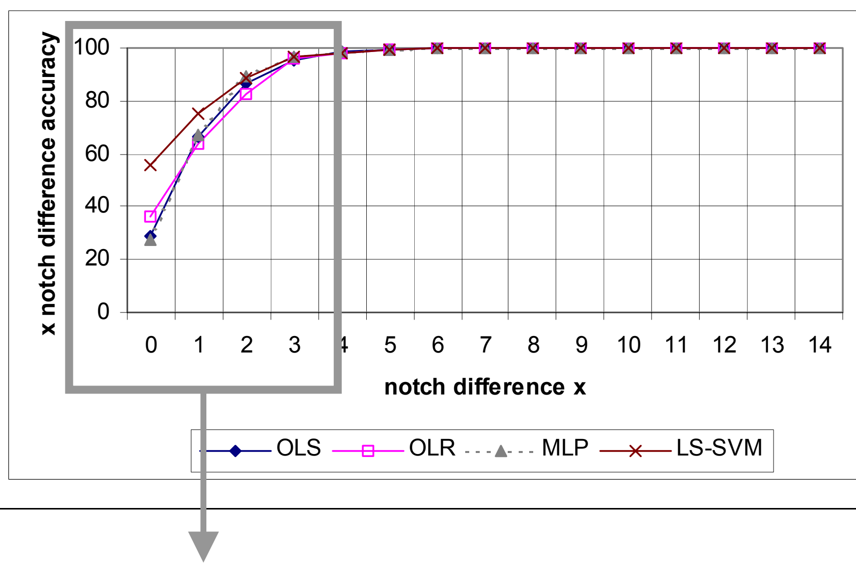
(4) The systems should not significantly produce systematically higher or lower ratings.

The four methodologies compared are Ordinary Least Squares (OLS), Ordinal Logistic Regression (OLR), the Multilayer Perceptron (MLP) and LS-SVM. All models were designed on the training set only and input selection is performed via backward input selection. The four test set confusion matrices are reported in appendix B. The OLS, OLR, MLP and LS-SVM achieve a zero notch difference accuracy of respectively 28.87%, 36.57%, 27.53% and 54.48% (see Figure 5a and 5b). The LS-SVM yields the best zero-notches-difference accuracy. When the rating problem is considered as a multi-class categorization problem, the McNemar test can be applied to detect significant differences in classification performance that are not due to luck. The *p*-value of the H0 hypothesis that the LS-SVM classifier performs equally well as the OLR classifier is equal to 1.3^{E-22}. This low value means that when assuming that the LS-SVM would not improve upon the OLR classifier, the probability of observing the reported performances is almost zero, meaning that the assumption of equal performances is not correct.

The x-notch accuracies are reported for increasing values of x in Figure 5. In this Figure, it is seen that the LS-SVM approach yields better performance on all x-notch accuracies. Allowing for a one notch difference, the following accuracies are obtained 66.36% (OLS), 63.51% (OLR), 67.44% (MLP) and 75.06% (LS-SVM), see Figure 5a and 5b. Besides good accuracy, it is important to know how the models perform in each of the aggregated rating categories A={A+,A,A-} to E = {E+,E,E-}. The zero-notch performances per category are reported in Table 1, from which it is observed that the OLS, OLR and MLP approaches perform poor in the outer aggregated categories A and E. Comparing in more detail the best linear model (OLR) and best nonlinear model (LS-SVM), we report the percentages of correct, lower and higher ratings for the two models in Figure 6a and 6b (i.e., model rating equal to, higher than or lower than the actual rating). These figures also clearly indicate the OLR approach systematically gives lower ratings to the A-category and higher ratings to the E-category.

During the modelling phase, the bank's key identifier and rating year were logged in all the subsequent steps. It allowed us to have a closer look at the ratings of the LS-SVM. These observations can be used when implementing the model in the more advanced stage. Some interesting conclusions are: 1) banks with sound financial ratios indicating good financial strength but with a low FitchIbca support rating are systematically rated higher by the system; 2) banks with poor ratios but with a healthy mother company are rated lower by the system; 3) misclassifications are also due to borderline cases in which the model gives a one or two year ahead forecast of a rating change.

Implementing the rating methodology also for scoring countries, similar results were obtained: the LS-SVM methodology outperformed the linear models. On a stratified test set of 113 countries, the zero notch difference accuracies for the OLS, OLR and LS-SVM are respectively 23.01%, 24.78% and 55.75%. At two notches difference, the performances increase to 76.11% (OLS), 72.57% (OLR) and 84.07% (LS-SVM), while at four notches difference, one obtains 93.81% (OLS), 92.04% (OLR) and 98.23% (LS-SVM), respectively. More detailed results on countries are available upon request.



a)

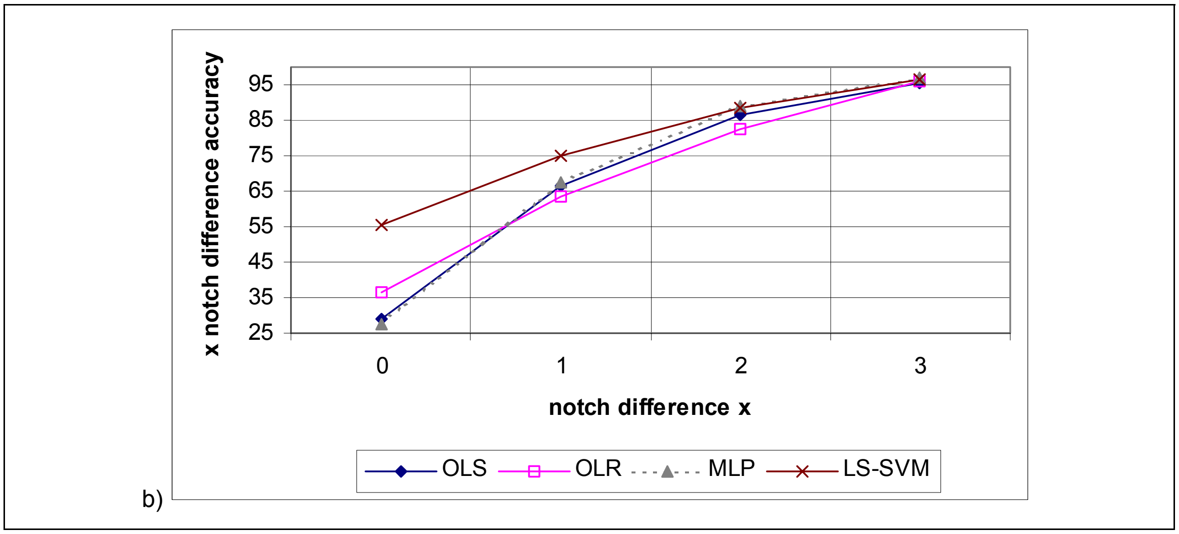


Figure 5: Performance of the four models (OLS, OLR, MLP and LS-SVM) for an increasing number of notch differences. The LS-SVM yields the best accuracy at zero notches and stays in the top of the performances when allowing for more notches difference between the actual rating and the model based rating.

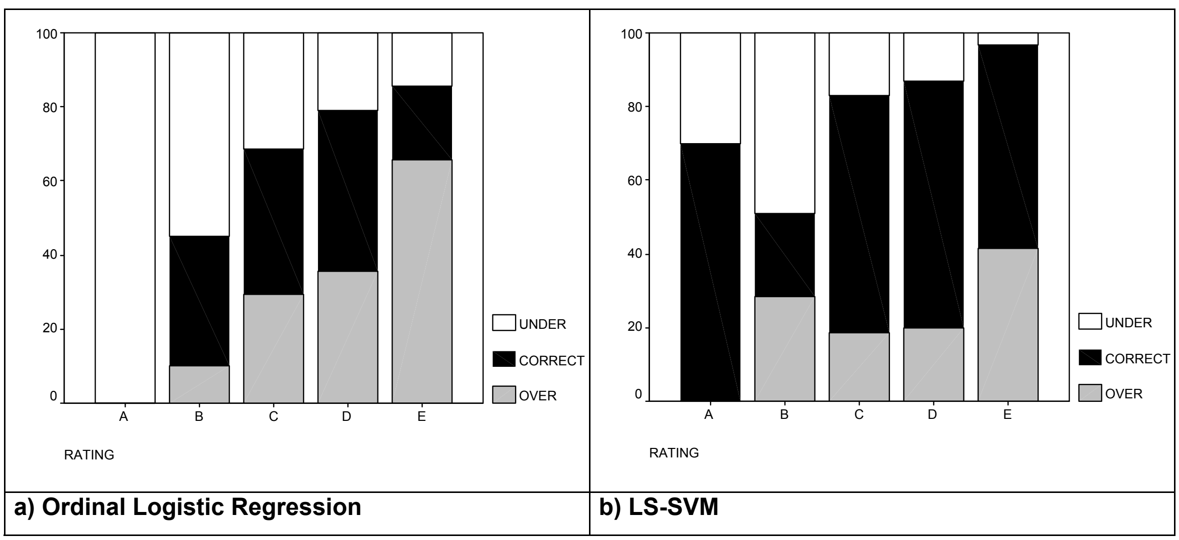


Figure 6: Percentages of lower, correct and higher ratings, aggregated per category A, ..., E. Observe that the Ordinal Logistic Regression model fails at achieving good results in the A and E categories, while the LS-SVM yields a more uniform performance.

	OLS	**OLR**	**MLP**	**LS-SVM**
A	00.00%	00.00%	34.14%	70.00%
B	23.69%	34.94%	34.14%	22.48%
C	33.58%	39.19%	31.04%	64.37%

D	33.51%	43.54%	27.70%	67.28%
E	17.53%	20.13%	11.03%	55.19%

Table 1: Zero-notch-difference for each of the four methods reported per aggregated rating category. The LS-SVM classifier is the only one yielding a good accuracy in category A. Also for the other categories, except B, much better results are obtained.

V. Conclusion

As credit scoring is an important tool in the internal rating based system required by the Basel II regulations, we compared in this paper classical linear rating methods with state-of-the-art SVM techniques, which is a recent technique already being referred to by rating agencies themselves. The test set results clearly indicate the SVM methodology yields significantly better results on an out of sample test set. A key finding is that the good performance is consistent on all considered rating categories.

Acknowledgements

This article is to a large extent derived from a research project on credit scoring modelling for Dexia. The authors would like to express their deep gratitude to E. Hermann (Head of Global Risk Management at Dexia Group) and L. Léonard (Head of Credit Modlling, Dexia Group) for initiating and supporting this research project. Moreover, the authors wish to thank M. Itterbeek, D. Sachs, C. Meessen, G. Kindt, L. Balthazar (Dexia Bank Belgium), D. Soulier (Dexia Group), J. Jaespart (Dexia BIL), A. Fadil (Dexia Crédit Local), J. Zou and M. Deverdun (PWC) for their collaboration in the project and J. De Brabanter, M. Espinoza, J. Suykens, B. De Moor (K.U.Leuven, ESAT-SCD), J. Vanthienen and M. Willekens (K.U.Leuven - FETEW) for the many helpful discussions. T. Van Gestel is on leave from the K.U.Leuven, faculty of applied sciences, ESAT-SCD, where he is a postdoctoral researcher with the FSR-Flanders. The scientific responsibility is assumed by its authors.

Appendices

A. List of candidate explanatory variables

	Ratio
1	Total operating income / number of employees
2	Net interest income / total operating income
3	Net commission income / total operating income
4	Net trading income / total operating income
5	Other operating income / total operating income
6	Personnel expenses / overhead
7	Net trading income / Net commission income
8	Loans/total assets
9	Customer & ST funding/total asset
10	customer depostits/total assets
11	Operational ROAE
12	Operational ROAA
13	Loan loss provisions/operating income
14	Non interest revenue / Total revenue (%)
15	Loan Loss Reserve / Gross Loans
16	Loan Loss Prov / Net Int Rev
17	Loan Loss Res / Non Perf Loans
18	Non Perf Loans / Gross Loans
19	NCO / Average Gross Loans
20	NCO / Net Inc Bef Ln Lss Prov
21	Tier 1 Ratio
22	Total Capital Ratio
23	Equity / Total Assets
24	Equity / Net Loans
25	Equity / Cust & ST Funding
26	Equity / Liabilities
27	Cap Funds / Tot Assets
28	Cap Funds / Net Loans
29	Cap Funds / Cust & ST Funding
30	Capital Funds / Liabilities
31	Subord Debt / Cap Funds
32	Net Interest Margin
33	Net Int Rev / Avg Assets
34	Oth Op Inc / Avg Assets
35	Non Int Exp / Avg Assets
36	Pre-Tax Op Inc / Avg Assets
37	Non Op Items & Taxes / Avg Ast
38	Return on Average Assets (ROAA)
39	Return on Average Equity (ROAE)
40	Dividend Pay-Out
41	Inc Net Of Dist / Avg Equity
42	Non Op Items / Net Income
43	Cost to Income Ratio
44	Recurring Earning Power
45	Interbank Ratio
46	Net Loans / Total Assets
47	Net Loans / Customer & ST Funding
48	Net Loans / Tot Dep & Bor
49	Liquid Assets / Cust & ST Funding
50	Liquid Assets / Tot Dep & Bor
51	Total Assets
52	Total Equity
53	Post Tax Profit

54	Region

B. Confusion matrices

<u>**Note**</u>**:** the diagonal represents the number of correct rating classifications by the model, values to the right of the diagonal imply the number of lower ratings by the model compared to the actual rating, left from the diagonal are higher classifications by the model.

Table 2: OLS Confusion Matrix

TRUE \ PREDICTED	A+	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	E+	E	E-
A+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	6	6	2	3	1	0	0	0	0	0	0	0
A-	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
B+	0	0	1	10	27	15	10	3	2	0	0	0	0	0	0
B	0	0	0	7	35	57	40	9	4	0	0	0	0	0	0
B-	0	0	0	0	2	14	11	2	0	0	0	0	0	0	0
C+	0	0	0	1	10	39	69	21	15	6	0	0	0	0	0
C	0	0	0	0	1	15	80	59	14	32	10	1	0	0	0
C-	0	0	0	0	0	2	8	5	4	1	0	0	0	0	0
D+	0	0	0	0	0	7	17	20	30	62	43	8	2	0	0
D	0	0	0	0	0	1	6	8	23	47	64	29	7	2	0
D-	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
E+	0	0	0	0	0	0	1	5	6	14	21	19	14	2	2
E	0	0	0	0	0	0	0	3	4	5	18	15	7	13	5
E-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 3: OLR Confusion Matrix

TRUE \ PREDICTED	A+	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	E+	E	E-
A+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	8	5	0	2	3	0	0	0	0	0	0	0
A-	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
B+	0	1	0	11	39	0	10	7	0	0	0	0	0	0	0
B	0	0	0	10	76	0	42	23	0	1	0	0	0	0	0
B-	0	0	0	1	13	0	8	7	0	0	0	0	0	0	0
C+	0	0	0	1	43	0	49	55	0	11	2	0	0	0	0
C	0	0	0	0	13	0	41	105	0	38	15	0	0	0	0
C-	0	0	0	0	1	0	6	11	0	2	0	0	0	0	0
D+	0	0	0	0	6	0	12	40	0	66	61	0	2	2	0
D	0	0	0	0	1	0	2	17	0	54	99	0	5	9	0
D-	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0
E+	0	0	0	0	0	0	0	7	0	14	35	0	6	22	0
E	0	0	0	0	0	0	0	6	0	5	30	0	4	25	0
E-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 4: MLP Confusion Matrix

	PREDICTED

TRUE	A+	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	E+	E	E-
A+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	9	3	2	2	2	0	0	0	0	0	0	0
A-	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
B+	0	0	1	11	29	14	10	3	0	0	0	0	0	0	0
B	0	0	1	4	61	53	23	6	4	0	0	0	0	0	0
B-	0	0	0	0	5	13	9	2	0	0	0	0	0	0	0
C+	0	0	0	3	10	51	56	21	16	3	1	0	0	0	0
C	0	0	0	0	3	28	48	60	28	34	10	1	0	0	0
C-	0	0	0	0	0	2	6	4	6	1	0	1	0	0	0
D+	0	0	0	0	2	3	8	29	28	51	47	17	4	0	0
D	0	0	0	0	1	1	2	8	13	62	51	35	13	1	0
D-	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0
E+	0	0	0	0	0	0	2	2	2	13	22	30	11	2	0
E	0	0	0	0	0	0	3	0	2	3	16	21	19	6	0
E-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 5: LS-SVM Confusion Matrix

	PREDICTED														
TRUE	A+	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	E+	E	E-
A+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	14	0	4	0	0	0	0	0	0	0	0	0	0	0
A-	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
B+	0	7	0	50	0	0	5	5	0	1	0	0	0	0	0
B	0	2	2	54	0	12	58	17	0	5	1	0	1	0	0
B-	0	0	0	6	0	6	10	6	0	1	0	0	0	0	0
C+	0	0	0	13	0	3	99	34	0	11	1	0	0	0	0
C	0	1	0	5	0	1	37	151	1	12	4	0	0	0	0
C-	0	0	0	2	0	0	3	8	3	4	0	0	0	0	0
D+	0	0	0	1	0	0	7	20	2	131	24	0	3	1	0
D	0	0	0	1	0	0	2	9	1	32	122	0	13	7	0
D-	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0
E+	0	0	0	0	0	0	0	4	0	12	26	0	37	5	0
E	0	0	0	0	0	0	0	4	0	4	9	0	5	48	0
E-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

References

[1] P.D. Allison, *Logistic Regression using the SAS system*, SAS Institute, 1999.

[2] E.I. Altman. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 23:589-609, 1968.

[3] A.F. Atiya. Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results. *IEEE Transactions on Neural Networks* (Special Issue on Neural Networks in Financial Engineering), 12, 929-935, 2002.

[4] B. Baesens, T. Van Gestel, S. Viaene, M. Stepanova, J.A.K. Suykens and J. Vandewalle. Benchmarking state of the art classification algorithms for credit scoring. *Journal of the Operational Research Society*, 2003 (in press).

[5] Bank of International Setllements. Quantitative impact study 3, October 2002.

[6] W.H. Beaver. Financial ratios as predictors of failure. Empirical Research in Accounting, Selected Studies, supplement to the Journal of Accouning, 5:71-111, 1966.

[7] C. Friedman, CreditModel Technical White Paper, Standard&Poor's, New York, 17-Sep-2002.

[8] T. Hastie, R. Tibshirani. Classification by Pairwise Coupling. *The Annals of Statistics*, 26, 451-471, 1998.

[9] J. Horrigan. The determination of long-term credit standing with financial ratios. *Journal of Accounting Research* (Supplement), 4:44:62, 1966.

[10] B. Khandani, A.E. Kocagil, and L. Carty. Risk Calc PublicTM – Europe: Rating Methodology. Moody's Investors Service, Global Credit Research, March 2001. http://www.moodysqra.com/us/research/crm/64793.pdf.

[11] J.A. Ohlson. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18:109-131, 1980.

[12] J.A.K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor and J. Vandewalle. *Least Squares Support Vector Machines*. World Scientific, Singapore, November 2002.

[13] T. Van Gestel, J:A.K. Suykens, B. Baesens, S. Viaene; J. Vanthienen, G. Dedene, B. De Moor and J.Vandewalle. Benchmarking Least Squares Support Vector Machine Classifiers. *Machine Learning*, 2003. (In press.)

[14] T. Van Gestel, J.A.K. Suykens, D. Baestaens, A. Lambrechts, G. Lanckriet, B. Vandaele, B. De Moor, J. Vandewalle. Financial time series prediction using Least Squares Support Vector Machines within the evidence framework. *IEEE Transactions on Neural Networks* (Special Issue on Neural Networks in Financial Engineering), 12, 809-821, 2002.

[15] V. Vapnik. *Statistical Learning Theory*. New York, Wiley, 1998.