



HAL
open science

Forecast bankruptcy using a blend of clustering and MARS model - Case of US banks

Zeineb Affes, Rania Hentati-Kaffel

► **To cite this version:**

Zeineb Affes, Rania Hentati-Kaffel. Forecast bankruptcy using a blend of clustering and MARS model - Case of US banks. 2016. halshs-01314553

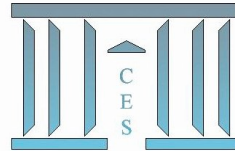
HAL Id: halshs-01314553

<https://shs.hal.science/halshs-01314553>

Submitted on 11 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**Forecast bankruptcy using a blend of clustering and
MARS model – Case of US banks**

Zeineb AFFES, Rania HENTATI-KAFFEL

2016.26



Forecast bankruptcy using a blend of clustering and MARS model - Case of US banks

Zeineb Affes* Rania Hentati-Kaffel[†]

March 30, 2016

Abstract

In this paper, we compare the performance of two non-parametric methods of classification, Regression Trees (CART) and the newly Multivariate Adaptive Regression Splines (MARS) models, in forecasting bankruptcy. Models are implemented on a large universe of US banks over a complete market cycle and running under a K-Fold Cross validation. A hybrid model which combines K-means clustering and MARS is tested as well. Our findings highlight that i) Either in training or testing sample, MARS provides, in average, better correct classification rate than CART model ii) Hybrid approach significantly enhances the classification accuracy rate for both the training and the testing samples iii) MARS prediction underperforms when the misclassification rate is adopted as a criteria iv) Results proves that Non-parametric models are more suitable for bank failure prediction than the corresponding Logit model.

Keywords: Bankruptcy prediction, MARS, CART, K-means, Early-Warning System
JEL: C14, C25, C38, C53, G17, G21, G28, G33

1 Introduction

Since 2009, default rate in banks has grown dramatically. More than 500 banks failed in the US since late 2008 and this risk begins to impact both banks of larger sizes that until now have always been rescued by public authorities "too big to fail". Prevention against systemic risk, the failure of the banking system, becomes an ineluctable concern

*Contact author. Centre d'Economie de la Sorbonne, Université Paris1 Panthéon-Sorbonne, Maison des Sciences Economiques, 106-112 Boulevard de l'Hôpital 75013 Paris, France. Email: zeineb.affes@univ-paris1.fr

[†]Contact author. Centre d'Economie de la Sorbonne, Université Paris1 Panthéon-Sorbonne, Maison des Sciences Economiques, 106-112 Boulevard de l'Hôpital 75013 Paris, France. Phone: 00 33 1 44 07 87 45, Email: rania.kaffel@univ-paris1.fr

[‡]This work was achieved through the Laboratory of Excellence on Financial Regulation (Labex ReFi) supported by PRES heSam under the reference ANR-10-LABX-0095.

and the need of new forecasting tools is of major importance to not only regulators but also academics.

In this sense, by insuring a monitoring system, The Federal Deposit Insurance Corporation (FDIC) try to provide an early warning to investors in order to draw their attention to those banks that have a great default probability ratio. The system employed is called "Capital Adequacy, Asset Quality, Management, Earnings and Liquidity (CAMEL). This latter is used by all the US bank regulatory agency. The FDIC developed also a Statistical CAMELS Off-site Rating system (SCOR) to perform the bank's stability evaluation. Collier et al. (2003) have examined the performance of this model over the 1986 to 2002 period and pointed the limitations of this model despite of the usefulness of SCOR which is based only on financial ratios. Cole and Gunther (1995) proved the same results and reported that these CAMELS ratings decay rapidly.

Predicting bank bankruptcy has reached a specific interest in financial literature. Thus, numerous models have been developed since the early 70s. All these models proposed could be considered as classification methods in a multidimensional space defined by a set of specific variables.

The literature has been studied by many authors using non-parametric and parametric methods. With regard to the later, Beaver (1966) was one of the first researchers who focused on univariate analysis to study bankruptcy prediction. He tested the power of financial ratios to classify and predict bankrupt firms. Cash flow and debt ratios appeared to be the important predictors of bankruptcy. Altman (1968) used Multivariate Discriminant Analysis (MDA) to develop a five factor model to calculate the well-known "Z-score" and predict bankruptcy of manufacturing firms.

Examples of statistical method, well known in the literature, include the logistic regression (logit) (Ohlson (1980)), Demirgüç-Kunt and Detragiache (1997)), Probit analysis (Zmijewski (1984) and Hanweck et al. (1977)) and factor analysis methods to estimate the probabilities of default and to predict firm failure. West et al. (1985) demonstrated that that the combination of factor analysis and Logit estimation was promising in evaluating banks' condition. The factors identified by the Logit model as important descriptive variables for the banks' operations are similar to those used for CAMELS ratings.

Martin (1977) used both Logit and MDA statistical methods to predict bank failures. Results obtained proved that the two models have similar classifications in terms of identifying failures and non failures of banks. In the same sense, Jones and Hensher (2004) presented mixed Logit model for firm distress prediction and compared it with Multinomial Logit Models (MNL). They concluded that mixed Logit obtained substantially better predictive accuracy than Multinomial Logit models. Wiginton (1980) finds the logit model results superior to discriminant analysis for consumer credit scoring.

Some other papers use non-statistical methods to set up a model for default prediction. For example, Kolari et al. (2002) use a Trait Recognition Model (TRA) (a kind of the image recognition algorithm). Empirical results recognized that these methods perform better than the statistical approach. Lanine and Vander Vennet (2006) showed that TRA approach outperforms Logit in predicting failures among Russian commercial

banks.

Other non-statistical methods include Intelligence techniques such as induction of classification trees and Neural Networks methods (NM). NM procedures follow a process similar to the human brain and contain mathematical and algorithmic elements that mimic the biological neural networks of the human nervous system (see Odom and Sharda (1990), Lenard et al. (1995)), Zhang et al. (1999), McKee and Greenstein (2000), Anandarajan et al. (2001)). Boyacioglu et al. (2009) tested various neural networks techniques and multivariate statistical methods to the problem of predicting bank failures in Turkey and presented a comparison of the classification performances of the techniques tested. They used similar financial ratios to those used in CAMELS ratings.

More recently, to solve classification and financial distress problems, Decision Trees (DT) and regression tree (CART) method was applied. Chen (2011) compared empirically DT with Logit for Taiwan firms, and found that DT got higher accuracy than Logit in short run (less than one year), while Logit performed better in long run (above one and half year). Iturriaga and Sanz (2015) developed a hybrid neural network model to study the bankruptcy of U.S banks by combining a Multilayer perception (MLP) network and Self-Organizing Maps (SOM). They found that the MLP-SOM can detect 96.15% of the failures in the period between May 2012 and December 2013 and outperforms traditional models of bankruptcy forecast.

De Andrés et al. (2011) and Sánchez-Lasheras et al. (2012) propose a new approach for the forecasting of firms' bankruptcy. They combines MARS model with i) fuzzy clustering and ii) Self Organized Map (SOM). They found that these hybrids models outperforms a single classification models i) LDA, NN (Feed-forward neural networks) and single MARS ii) NN (back propagation neural networks) and MARS in terms of correct classification and of the identification of the companies that go bankrupt.

In this paper, we aim to model the relationship between ten financial variables and default probability of US bank by using the so-called non-parametric or flexible models. The main purpose is to highlight and to test the accuracy of non-parametric methods of classification and to increase its prediction ability and to reduce misclassification problem by proposing a blend of k-means and MARS model. We propose, for the first time in the bank failure literature, the use of MARS and also a hybrid model that combines K-means and MARS model.

Methodologically, in addition to Regression Trees (CART) we propose to validate the Multivariate Adaptive Regression Splines model which gained an increasing interest in financial literature. MARS was first proposed by Friedman (1991). The main advantage of this model is the capacity to explore the complex nonlinear relationships between response variable and various predictor variables.

However, in the empirical part we contribute to the existing literature by implementing all these models on a large universe of US banks over the period spanning 2008 to 2013, under a K-fold Cross validation. according to the size of our data set we apply a 10-fold cross validation to separate our data set into training and validation sets. In order to evaluate models fitting, we consider the confusion matrix both for the training and the testing samples. Also, we use the Receiver Operating Characteristic Curves

(ROC) for evaluating classification success. Finally, we evaluate the performance of each model according to the Area under the ROC Curve.

Our main findings highlights the promising functionality of MARS model and suggest that: i) Either in training or in the testing sample, MARS provide better correct classification than CART model in average (96,06%-94,37% versus 94,76%-94,07%) ii) Hybrid approach enhanced the classification accuracy by 3% for the training sample and 2,55% for the testing one iii) Relying on misclassification rate to test prediction power, MARS underperformed, especially in 2008 and 2009 iv) According to the AUC of the Receiver Operating Characteristic Curve we observe a tiny difference in the training and the testing sample results of MARS except for 2013 .

The paper is organized as follows. Section 2 presents the methodology and data used. Section 3 describes results of CART and MARS models. In Section 4, we analyze used models predictability. Finally, Section 5 concludes this paper.

2 Review methodology

Empirical validation in this paper is based on a large panel of US banks. We collect data from BankScope for active banks (AB) and FDIC for non-active (NAB) ones from 2008 to 2013.

We extract all ratios to build 10 financial variables detailed as follow:

Categories CAMEL	Variables	Definition
Capital Adequacy	EQTA	Total Equity/Total Assets
	EQTL	Total Equity/Total Loans
Assets Quality	NPLTA	Non Performing Loans/Total Assets
	NPLGL	Non Performing Loans/Gross Loans
	LLRTA	Loan Loss Reserves/Total Assets
	LLRGL	Loan Loss Reserves/Gross Loans
Earnings Ability	ROA	Net Income/Total Assets
	ROE	Net Income/Total Equity
Liquidity	TLTD	Total loans /Total customer Deposits
	TDTA	Total Customer Deposits/Total Assets

The choice of these ten ratios was conducted and justified by an abundant literature (Sinkey (1979), Martin (1977), Thomson (1991), Barr et al. (1994), Pantalone et al. (1987), Godlewski et al. (2003)).

We maintain the same rule of bank statue selection as in Affes and Hentati-Kaffel (2016). Thus number of (NAB) banks was 411 failed banks over the entire period 2008-2013 and 836 active banks each year. In our calculation, we use standardized values of predictor's variables.

However, it was proved that classification tends to favor active banks (AB) which represent the majority class. This means that the original database has a highly skewed distribution. To create homogeneous groups we apply k-fold cross validation.

We apply a 10-fold cross validation to separate our data set into training and validation sets. 10-fold is the most widely used number of fold in cross validation.

The procedure for each model is the same and summarized as follow:

1. For each of 10 experiments, we use 9 folds for training and the remaining one for testing ,
2. We reiterate 10 times for each 10-fold cross validation experiments,
3. We select parameters of the best model and then we minimize the cross validation error rate.

In order to evaluate the classification suitability of models, we established the confusion matrix for the training and the testing samples. Also, we use the Receiver Operating Characteristic Curves (ROC) for evaluating classification success.

The ROC curve presents the possible distributions of scores for the banks. We determine the optimal cut-off value that maximize the sensitivity (failed banks correctly classified) and the specificity (non-bankrupt banks correctly classified) and then classify the banks as a potential default bank when the score is lower than the cut-off or as healthy if the score is higher than the cut-off value.

3 Running Classifications methods

3.1 Multivariate Adaptive Regression Splines (MARS) implementation

Introduced by Jerome Friedman in 1991, Multivariate Adaptive Regression Splines is a form of stepwise linear regression which can model non-linearity between variables.

MARS is based on three parameters:

1. the maximum number of basis functions (term),
2. the smooth parameter (called also the penalty parameter),
3. and the maximum number of iteration between variables (see Andalib and Atry (2009)).

In MARS, the basis function (term) is fitted to segregate independent variable intervals by using recursive splits. In this model all possible splitting points are estimated with a linear spline (also called piecewise polynomials). The best splitting point (knot positions) is the one for which the model extensions minimize a squared error criterion. Knot is the point at which two polynomial pieces connect. The best splitting point is the one for which the model extension minimize a squared error criterion. Knots uses two-sided truncated power functions as spline basis functions, described in Eqs [1] and [2]

$$h^+(\boldsymbol{x}; t) = [+(\boldsymbol{x} - t)]_+ \quad (1)$$

$$h^-(\boldsymbol{x}; t) = [-(\boldsymbol{x} - t)]_+ \quad (2)$$

where $[q]_+ = \max\{0, q\}$ and t is a univariate knot. MARS is represented as a combination of piecewise linear or hinge functions. These letters have a knot or hinge at t , are zero on the one side of the knot and are linear on the other side.

The MARS fit a linear model in basis functions $\{h_m(\boldsymbol{x})\}_{m=1}^M$:

$$\hat{f}_M(\boldsymbol{x}; \beta) = \beta_0 + \sum_{m=1}^M \beta_m h_m(\boldsymbol{x}), \quad (3)$$

where $h_m(\boldsymbol{x})$ is a basis function of the form described below in Eqs [1] and [2], M is the number of linearly independent basis functions, and β_m is the unknown coefficient for the m th basis function.

As mentioned above, a knot is the point in a range at which the slope of the curve changes. Both the number of the knots and their placement are unknown at the beginning of the process. A stepwise procedure is used to find the best points to place the spline knots. In its most general form, each value of the independent variable is tested as a possible point for the placement of a knot. The model initially developed is overfit (forward phase). A statistical criterion (generalized cross-validation) that tests for a significant impact on a goodness of fit measure is used to remove knots. Only those that have a significant impact on the regression are retained (backward phase).

Another complication occurs when working with an interaction. The relationships between a predictor variable and the target variable β_m may depend on the value of a second variable.

MARS can perform regressions on binary variables. When the dependent variable is binary, MARS is run in binary mode. In binary mode, the dependent variable is converted into a 0 (non-failed banks) or a 1 (failed bank). Ordinary Least Square (OLS) regression is then performed, it regresses the binary variable on the predictor variables.

Below we provide a detailed explanation of MARS output for both 2008 and 2013 ¹.

For 2013, first of all, we begin our analysis by detecting interaction between variables. Generalized Cross Validation (GCV) gives the amount of degradation in the model when a ratio is deleted. A model with minimum (GCV) should be chosen. In this sense Mean-Squared Error (MSE) and (GCV) are given in MARS output (Table 1) and values of (GCV) and (MSE) are similar at the value 0.00699; therefore, we can conclude that an optimal model is reached for six basis functions.

Table 2 exhibits the importance of each ratio in the model. TDTA, TLTD, EQTA, LLRGL and ROE ratios were the most important variables with 100, 89.29, 88.19, 87.84, and 87 percent importance, respectively. NPLTA and ROA percent were less

¹Results for the rest of 2009, 2010, 2011 and 2012 are detailed in Appendix A

important(61.80 and 49.89 percent respectively), whereas percent of EQTL, NPLGL, LLRTA are null.

The ANOVA function 2 has the greatest effect on the model with a (GCV) score of 0.01053, meaning that the interaction between the most important variables (TDTA and TLTD) impacts significantly the target variable. Also, the contributions of the variable ROE and the interaction between EQTA and LLRGL were considerable (GCV score on both ANOVA function 1 and 6 is equal to 0.0104). Moreover, the effects of the ANOVA function 3, 4 and 5 are quite good (see Table 3).

Basis functions (BF) in 2013 are as follow:

$$\begin{aligned} BF1 &= \max(0, TDTA - 0.843717); \\ BF8 &= \max(0, -0.728978 - EQTA); \\ BF10 &= \max(0, ROE - 0.0800314); \\ BF28 &= \max(0, TLTD + 0.755844) * BF1; \\ BF59 &= \max(0, NPLTA - 1.16217) * BF1; \\ BF63 &= \max(0, LLRGL + 1.61568) * BF1; \\ BF70 &= \max(0, ROA + 0.73126) * BF1; \\ BF88 &= \max(0, LLRGL - 1.15542) * BF8; \end{aligned}$$

Where max (a, b) returns the maximum of a and b.

The above basis functions prove the non-linear relationships between the dependent and independent variables (Figure 1).

The final model is expressed as follow:

$$\begin{aligned} Y &= 0.003665 + 0.466474 * BF10 + 3.07491 * BF28 - 0.246509 * BF59 \\ &\quad - 0.224633 * BF63 - 2.15719 * BF70 + 0.241695 * BF88 \end{aligned} \quad (4)$$

It appears that for example from BF1, on variable TDTA (liquidity proxy), data is grouped into two sets: the first one is assigned 0 for all TDTA values that are below a threshold (e.g., $c = 0.843717$) and the second set contains the elevation values that are more than 0.843717. The BF1 does not appear in the final model but it contributes in the construction of others basis function (BF28, BF59 and BF70).

Indeed, the BF28 is defined as a combination between TLTD and TDTA. This basis function has a positive effect on the target variable only when the value of TLTD exceeds - 0.755 and the value of TDTA is greater than 0.843717. The basis function (BF59) and (BF70) which are a combination of (BF1) with respectively NPLTA and ROA, have a negative impact on the output. In other word, a value of TDTA greater than 0.843717 multiplied by value of NPLTA exceeding 1.16 or value of ROA great than -0.7312, affect negatively the target 'Y'. Moreover, an increase in the value of ROE above 0.08 (in BF10) will increase the variable 'Y'. The negative effect of (BF63) appears only when the value of LLRGL at least -1.615 and BF10 is positive. The BF88 is a function of the variable LLRGL with a knot of 1.155, multiplied by the basis function BF8. It means that when the value of LLRGL is greater than 1.155 and the value of EQTA exceeds -0.728, we note a positive impact on the target variable.

The positive sensitivity to BF10 and BF28 suggests that variables ROE, TLTD and TDTA have a positive effect on the output. With a high level of liquidity and a better

profitability, the score of the bank augments (see Table 4).

Results depend on considered year. In 2008, basis functions are as follow:

$$\begin{aligned} BF2 &= \max(0, -0.471622 - EQTL); \\ BF3 &= \max(0, NPLGL - 1.33738) * BF2; \\ BF6 &= \max(0, -0.879754 - EQTA) * BF2; \\ BF7 &= \max(0, NPLGL - 0.70157) * BF2; \end{aligned}$$

And the final model is expressed by:

$$Y = 0.00703772 + 1.09848 * BF2 - 4.41337 * BF3 - 1.3883 * BF6 + 4.57587 * BF7 \quad (5)$$

We note that BF2 is zero for value of EQTL greater or equal to -0.471622. A positive sign for the estimated beta factor of BF2 indicates an increase of the output variable. We also note the presence of interaction between predictor variables, which means that the effect of a predictor on the target variable may depend on the value of another predictor. We see in the definition of the BF6 that the effect of the variable EQTA on the output variable depends on the value of the ratio EQTL. The effect of this interaction can be explained as follow. If the value of EQTA is lower than -0.879 and the value of EQTL is below -0.47, it has a negative impact on the target variable. The basis functions BF3 and BF6 account for the nonlinear effect between the target and the predictor variables. Indeed, when the value of NPLGL is above 1.337 at the same time, the EQTL is lower than -0.47, then the output variable will decrease. On the other hand, a value of NPLGL greater than 0.7 and EQTL below -0.47, have a positive effect on the target variable.

3.2 CART Algorithm to build Tree Classifiers

The CART (Classification and Regression Trees) algorithm proposed by Breiman et al. (1984) is widely used statistical procedure. This procedure produces classification and regression models via tree-based structure. It is based on a hierarchy of univariate binary decisions and operates by selecting the best variable for splitting the data into two groups at the root node. CART is a form of binary recursive partitioning in which partitions can be split into sub-partitions. This classifier assigns a predicted class membership obtained under a specific measurement (x_1, x_2, \dots, x_k) . Let's X the measurement space of all possible values of x. Tree's classifiers are constructed by making repetitive splits of X and the subsequently obtained subsets of X. As consequence, a hierarchical structure is formed. In Finance feature, Frydman et al. (1985) were the first who employed decision trees to forecast default. After that, many research highlighted the accuracy of this method to predict bankruptcy (Carter and Catlett (1987), Gepp et al. (2010)).

To build a tree by CART, the procedure should specify a number of parameters: i) the splitter that will allow to visualize the left branch if the splitter's variable < value split ii) the competitor identifier variable. In our bank failure problem, the dependent variable is either bankrupt or non-bankrupt, so classification tree is suitable for our case.

Table 5 identifies the node competitors in order of improvement and Table 6 exhibits the importance of each ratio in the building of CART tree. In 2013 EQTA, EQTL, ROA, TDTA, LLRGL and NPLGL ratios were the most significant variables with 100, 98.60,

89.80, 67.07, 65.32, and 62.34 percent importance, respectively. TLTD and NPLTA were less significant (1.22 and 0.71 percent respectively), whereas percent of ROA is null.

We implement CART on our data. Figure 2 shows a classification tree that has been constructed from the data in 2013. The variable EQTA produces the greatest “separation” in the target and we consider it in as a first node of the Tree. Thus the value of $EQTA = -1.15$ represents the split. This tree is split at each branch by a decision boundary (left is yes, right is no).

The tree shows a classification of banks based on the ratio EQTA. Thus, if a bank has a *ratio* ≤ -1.15 , then the bank is considered non-active (NAB) and healthy (AB) otherwise.

Classification must normally converge to the following clusters: among the 851 banks surveyed in 2013, 98.2% are in (AB) group (836 banks) and only 15 banks (1.8%) are in the (NAB) group.

According to the left branch of the tree ($EQTA \leq -1,15$), among the 851 banks, 36 are classified in (NAB) (class1). However, only 15 banks were actually NAB, yielding to a misclassification of 21 banks. According to the right branch of the tree ($EQTA > -1,15^2$), all banks (851) were well classified.

In the second level, a second distinction is based on the target value of -0.55 for the ratio ROA. For banks having a common ROA ratio greater than - 0.55. Among the 36 (NAB) only 9 are (AB) and for banks having a common $ROAratio \leq -0.55$, 27 are considered (NAB). However, if we check the accuracy of CART classification, we find 15 banks have actually gone bankrupt (correctly classified by CART) and 12 were misclassified because they still active.

By simply cascading down the decision branches of these tree, we see at the level of node 3 that (NAB) class which additionally have a ratio of $TLTD > - 1.2$ contains 24 banks. (15 are really (NAB) and 9 are misclassified). The left branch of this split counts 3 (AB) banks which are all correctly classified.

For This 24 banks, as their ratio (TLTD) is less than or greater than the target value of 0.04 allows us to descend into the next node and make an even finer distinction to add another subgroup. If $TLTD > 0.04$ (0.7849) we find that 2 (AB) banks are classified in Class 0 (correctly classified). If $TLTD \leq 0.04$ (0.7849), 22 banks will be classified in the group of failing banks, and among these 22 banks, 15 are actually faulty and 7 active banks will be declared as failed (miss classification).

In the last node, according to the target -0.16, only one bank is in the class 0 (if $LLRGL \leq -0.16$) else 21 are (NAB) (6 banks are misclassified).

CART classification allows regulator to provide a lot of information about banks that can potentially fail in the coming years. Indeed, in our case we have checked that banks that were misclassified by CART actually default in 2015 and 2014.

It also allows the regulator to set target values that should be used to detect these suspected banks. For 2013, for example we can consider that banks with simultaneously an $EQTA \leq -1.15$, $ROA \leq -0.55$, $TLTD > -1.20$ and $LLRGL > -0.16$ should be classified in the group of failing banks (miss error classification type II)

² $EQTA > -0,15$ correspond in non-normalized value to $EQTA > 0,0562$

4 Models accuracy and prediction results

The primary goal to the following section is to compare the prediction estimates and the classification performance between decision tree modeling and MARS methods, taken solely or combined in a hybrid model.

Hereafter, we focus on two problems:

- i) Selecting the best model with good sensitivity and specificity rates
- ii) Find the appropriate cut-off points or criteria for performance, separating wrights from false rates results. Setting up these values must be proposed carefully. In these sense we propose both the confusion matrix analysis under the selection of the best cut-off delivered from the Receiver-Operating characteristic (ROC) curves and also the Area under (ROC)Curve.

4.1 MARS vs CART

Based on the confusion matrix, in Tables 7, 8, 10 we record several rates to quantify sensitivity and accuracy rates and correctly and incorrectly classification (errors type I and II).

In 2013, in Testing Sample (TS), MARS model was able to correctly classify 97.18% of the banks. Only one failed bank was misfiled (type I error: 6.67%). By cons, 23 (AB) banks were considered by the model as failing banks (type II error 2.75%). For the years 2012 and 2011, the model's ability to properly classify banks is justified both in the "Testing Sample" and "Training Sample, (TRS)".

However, for the years 2008, 2009 and 2010, MARS underperformed other models in terms of sensitivity. A slight improvement in the sensitivity rate was found as we moved far from the year of the economic crisis. Therefore, type I error rate is quite high for the year 2008. Thus, the error term I was lower for the years 2009 and 2010. In addition, percentage of non-failed banks correctly predicted was 93.54%, 89.06% and 92.58% respectively for the years 2008, 2009 and 2010.

In terms of correct classification rate, MARS model performs better in the (TNS) than in (TS) (98.82% against 97.18%). However, we observed that the error type 1 is higher in (TNS) than in (TS) (13.33% against 6.67%). Moreover, 0.96% of the non failed banks in the (TNS) were classified in the group of the failed banks. We also note that 97.25% of the non failed banks in the (TS) are classified in the group of the (NFB). The same results were observed for the years 2011 and 2012. To summarize, we conclude that MARS model has a good predictive performance in term of ability to reduce the error type I and also by generating the best signal to trace suspicious banks among the non failed banks.

For the period 2008-2010 (both in (TNS) and (TS)), the performance of MARS was slightly lower in terms of the correct classification, sensitivity and specificity. Indeed, results exhibits higher rate of the type I error rates.

Table 8 highlighted also that decision tree modeling (CART) produces a high level of correct classification in 2013. We observe a small difference between the training and the testing sample (99.29% against 98.94%). 99.29% of the banks in the (TNS)

are correctly classified and 98.94% of the banks in (TS) are classified in their adequate groups). It appears that CART for the (TNS) classify correctly all the failed banks (sensitivity 100%) and for the testing only one inactive bank was predicted as non failed bank (error type I 6.67%). We also notice that 0.72% actives banks of the training and 0.96% non failed banks of the testing sample are predicted as failed banks.

We observed the same main results in term of sensitivity for the years 2011 and 2012. Actually, in the learning sample, all failed banks are classified as bankrupt ones (error type I: 0.00%) and when we test CART model we obtain a low rates of error type I (4.11% and 2.7% for 2011 and 2012 respectively).

However, results obtained in 2008, 2009 and 2010 show that CART does not procure a high correct classification rate. In fact, for 2009 we note a correct classification rate about 89% in both training and testing sample. Moreover for this period, CART model provides a high misclassification rate. Indeed, we notice an average of error type II about 8.83% and an average of error type I about 7.94% in the testing sample.

4.2 Hybrid model accuracy

In order to improve the results of both models CART and MARS, we propose to build a hybrid model based on the classification model K-means and MARS.

Clustering is a method of grouping (Anderberg (2014); Hartigan (1975); Jain and Dubes (1988)) a set of objects into groups according to criteria predefined similarities between objects. Most clustering methods are based on a distance measure between two objects. Technically, clustering can be regarded as a minimization problem.

Let X the matrix of dimension data (N, n) :

$$X = \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \cdot & \dots & \cdot \\ \cdot & \dots & \cdot \\ x_{N1} & \dots & x_{Nn} \end{pmatrix},$$

N corresponds to the number of banks, n the number of years and x_{ij} ratios variables.

From $N \times n$ dimensional data set K-means algorithms allocates each data point to one of c clusters to minimize the within-cluster sum of squares:

$$\sum_{i=1}^c \sum_{k \in A_i} \|x_k - v_i\|^2 \quad (6)$$

where A_i is banks in the cluster i and v_i is the mean for these banks group over cluster i . This equation denotes actually a distance norm. In K-means clustering v_i is called the cluster prototypes, i.e. the cluster centers:

$$v_i = \frac{\sum_{x_k \in A_i} x_k}{N_i} \quad (7)$$

where N_i is the number of banks in A_i .

In our paper Z-score standardization is applied to find clusters and the number of cluster solutions is Two.

In Figure 3 we run principal components analysis on the data set and display a plot of the first two principal dimensions.

Also, In table 9 we provide results of classification based only on K-means. What we propose is to classify data set under K-means and then we apply MARS model.

Results are as follows (Table 10): In 2013, the hybrid model provides a satisfactory rate of correct classification but we notice a slight gap between the training and the testing sample (98% against 96%). The model classifies correctly all the bankrupt banks both in training and testing sample (sensitivity: 100%). On the other hand, we note a misclassification rate of the actives banks in the testing sample about 4.20% against 2.10% in the training. For the others years, we observe mainly the same results in term of correct classification rate. As it can be seen there is a difference between the accuracy of the model in the training and the testing sample (for 2008: correct rate was 99.20% in (TNS) and 95.98% in (TS)). Moreover, the hybrid model provides a low error type I in both training and testing. However, we notice that the misclassification of the actives banks is more important in the testing than in the learning sample.

Finally, according to the Area under Curve (ROC)(Table 12), we conclude that the MARS model provides a better accuracy results than CART model. For example for 2013 the AUC of MARS is greater than CART in the testing sample (99.08% against 96.41%). While for 2008, CART outperforms MARS with an AUC of 92.19% against 88.94%. However, in the training sample the results of CART models show a better performance than MARS only for the years 2008 and 2013 (95.00% against 86.56% for 2008 and 99.64% against 89.86% for 2013).

We also observe a slightly difference in the training and the testing sample results of MARS except for 2013 where we note a big difference (89.86% for training against 99.08% for testing).

The hybrid model K-MARS outperforms all the other models CART and MARS in both training and testing samples with an AUC around 99.5%.

To sum up, MARS model provides better results than CART in terms of average correct classification rate (94.37% against 94.07%) and in terms of average AUC (95.81% versus 94.53%) for the testing sample. The combined model K-means and MARS is the best model in terms of performance and accuracy in both training and testing sample in terms of average correct classification rate (98.84% vs 96.92%) and average AUC (99.93% 99.43%)(Figures 4, 5 and 6).

5 Conclusion

In this paper, we developed a blend model based on two non-parametric classification models to study the bankruptcy of US banks. We provide a comparative approach between CART, MARS and K-means-MARS. Our main objective is to predict bank defaults some time before the bankruptcy occurs, and to build an early warning system based on CAMEL's ratios.

We based our empirical validation on a large panel of US banks gathered from both Bankscope and from the Federal Deposit Insurance Corporation.

The main contributions of our paper with regard to the literature which deals with the topic of predicting failure are two fold:

-Methodological and conceptual: First, we propose for the first time a hybrid model that combines K-means and MARS model. We provide a comparative framework not to non-parametric models but also to parametric models Logit and CDA (Affes and Hentati-Kaffel (2016)).

-Empirical validation: Our study focuses on a large sample: small and large US banks are considered. Also, we do not limit ourselves to the study of commercial banks. The study allows to scrutinize the behavior of banks over a period of six years which include both type of periods, stress (2008) and recovery (thereafter).

Our results reached tow targets: it showed the interest of all the non-parametric models tested and it justified their adequacies when it comes to conduct both a classification procedure and forecasting of bank failures. After, our hybrid model shows a high discriminant power and was able to differentiate correctly wealthy and distressed banks. The hybrid model k-means MARS has a better classification capability than the others models CART and MARS in both training and validation samples in terms of average correct classification and misclassification rates. Our main findings according to the accuracy in term of classification suggest that:

First, our hybrid model shows a high discriminant power and is able to differentiate correctly wealthy and distressed banks. The hybrid model k-means MARS has a better classification capability than the others models CART and MARS in both training and validation samples in terms of average correct classification and misclassification rates. The model enhanced the classification accuracy by 3% for the training sample and 2, 55% for the testing one.

Second, either in training or in the testing sample, MARS provide better correct classification than CART model in average (96.06%-94.37% versus 94.76%-94.07%) (see Table 11).

In term of prediction, we proved that MARS underperform according to the misclassification rate notably in 2008 and 2009. Also, according to the Receiver Operating Characteristic Curve we observe a slightly difference in the training and the testing sample results of MARS except for 2013. According to the Area under Curve (ROC), we conclude that the MARS model provides a better accuracy results than CART model. For example for 2013 the AUC of MARS is greater than CART in the testing sample (99.64% against 96.41%). While for 2008, CART outperforms MARS with an AUC of 92.19% against 88.94%. As mentioned in the introduction, the ultimate goal of this paper is to provide regulators and investors with an early warning model of banking difficulties. From the results obtained in our paper, to assess the fragility of banks and identify the most vulnerable institutions risks, supervisors must conduct an analysis on financial ratios and variables.

It is true that the results differ from one year to another, but we manage to detect a general behavior for all distressed banks. For example, based on CART classification, our

study shows that among the 10 used ratios, the most important predictors are: (EQTA) and (EQTL) proxies of Capital Adequacy. Also, we note that the ratios (NPLTA) and (NPLGL) of Asset Quality proxy are much more important than the other two components (LLRTA) and (LLRGL). The (ROA) variable is of importance (except for 2008). The CART model allows to advise banking regulators to develop capital and liquidity rules to allow banks to withstand future downturns in the global financial system. For example, in the study of 2008 TREE (year of the economic crisis), three alert rules can be applied: all banks have both the following characteristics should be made a control and close monitoring because they present a high risk of bankruptcy in the years to follow: i) $NPLGL > 0,75$ and $EQTA > -0.64$ and $EQTL < -0.31$ ii) $NPLGL > 0.75$ and $EQTA < -0.64$ iii) $NPLGL < 0.75$ and $ROE < -0.79$.

The results obtained in this study thus help supervisory authorities to be able to closely monitor and to statistically analyze information from documents and data provided by targeted bank based on threshold variables obtained from the CART analysis.

According to MARS the most important variables are the components of the Capital Adequacy. (ROE), (TLTD) and (TDTA) variables are important only in 2013. We note that, with respect to parametric models (see Affes and Hentati-Kaffel (2016)) the Asset Quality is the most important component to explain the financial conditions of banks. We note also that the Capital Adequacy (EQTA and EQTL) and the Liquidity (TLTD and TDTA) have an importance in detecting bank failure.

We can conclude that MARS model is a useful tool to identify in advance financial institutions in stress and so will be deserved with a special attention by supervisors. For example, in 2009, among the 60 AB predicted as NAB by MARS, 51 banks really will go bankrupt 4 years later: 44 banks bankrupted in 2010, 3 in 2011, 3 in 2012 and 1 in 2013. Finally, we believe that further extensions can be developed by including more financial variables and macroeconomic variables.

References

- Affes, Z. and Hentati-Kaffel, R. (2016). Predicting US banks bankruptcy: logit versus Canonical Discriminant analysis. Documents de travail du Centre d'Economie de la Sorbonne 2016.16 - ISSN : 1955-611X.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4):589–609.
- Anandarajan, M., Lee, P., and Anandarajan, A. (2001). Bankruptcy prediction of financially stressed firms: An examination of the predictive accuracy of artificial neural networks. *Intelligent Systems in Accounting, Finance and Management*, 10(2):69–81.
- Andalib, A. and Atry, F. (2009). Multi-step ahead forecasts for electricity prices using narx: a new approach, a critical analysis of one-step ahead forecasts. *Energy Conversion and Management*, 50(3):739–747.

- Anderberg, M. R. (2014). *Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks*, volume 19. Academic press.
- Barr, R. S., Siems, T. F., et al. (1994). Predicting bank failure using dea to quantify management quality. Technical report, Federal Reserve Bank of Dallas.
- Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of accounting research*, pages 71–111.
- Boyacioglu, M. A., Kara, Y., and Baykan, Ö. K. (2009). Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (sdif) transferred banks in turkey. *Expert Systems with Applications*, 36(2):3355–3366.
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and regression trees*. CRC press.
- Carter, C. and Catlett, J. (1987). Assessing credit card applications using machine learning. *IEEE expert*, 2(3):71–79.
- Chen, M.-Y. (2011). Predicting corporate financial distress based on integration of decision tree classification and logistic regression. *Expert Systems with Applications*, 38(9):11261–11272.
- Cole, R. A. and Gunther, J. (1995). A camel rating’s shelf life. *Available at SSRN 1293504*.
- Collier, C., Forbush, S., Nuxoll, D., and O’Keefe, J. P. (2003). The scor system of off-site monitoring: its objectives, functioning, and performance. *FDIC Banking Review Series*, 15(3).
- De Andrés, J., Lorca, P., de Cos Juez, F. J., and Sánchez-Lasheras, F. (2011). Bankruptcy forecasting: A hybrid approach using fuzzy c-means clustering and multivariate adaptive regression splines (mars). *Expert Systems with Applications*, 38(3):1866–1875.
- Demirgüç-Kunt, A. and Detragiache, E. (1997). *The determinants of banking crises-evidence from developing and developed countries*, volume 106. World Bank Publications.
- Friedman, J. H. (1991). Multivariate adaptive regression splines. *The annals of statistics*, pages 1–67.
- Frydman, H., Altman, E. I., and KAO, D.-L. (1985). Introducing recursive partitioning for financial classification: the case of financial distress. *The Journal of Finance*, 40(1):269–291.
- Gepp, A., Kumar, K., and Bhattacharya, S. (2010). Business failure prediction using decision trees. *Journal of forecasting*, 29(6):536–555.

- Godlewski, C. J. et al. (2003). Modélisation de la prévision de la défaillance bancaire une application aux banques des pays émergents. *Document de travail LARGE, Université Robert Schuman Strasbourg*.
- Hanweck, G. A. et al. (1977). Predicting bank failure. Technical report, Board of Governors of the Federal Reserve System (US).
- Hartigan, J. A. (1975). Clustering algorithms.
- Iturriaga, F. J. L. and Sanz, I. P. (2015). Bankruptcy visualization and prediction using neural networks: A study of us commercial banks. *Expert Systems with Applications*, 42(6):2857–2869.
- Jain, A. K. and Dubes, R. C. (1988). *Algorithms for clustering data*. Prentice-Hall, Inc.
- Jones, S. and Hensher, D. A. (2004). Predicting firm financial distress: A mixed logit model. *The Accounting Review*, 79(4):1011–1038.
- Kolari, J., Glennon, D., Shin, H., and Caputo, M. (2002). Predicting large us commercial bank failures. *Journal of Economics and Business*, 54(4):361–387.
- Lanine, G. and Vander Vennet, R. (2006). Failure prediction in the russian bank sector with logit and trait recognition models. *Expert Systems with Applications*, 30(3):463–478.
- Lenard, M. J., Alam, P., and Madey, G. R. (1995). The application of neural networks and a qualitative response model to the auditor’s going concern uncertainty decision*. *Decision Sciences*, 26(2):209–227.
- Martin, D. (1977). Early warning of bank failure: A logit regression approach. *Journal of banking & finance*, 1(3):249–276.
- McKee, T. E. and Greenstein, M. (2000). Predicting bankruptcy using recursive partitioning and a realistically proportioned data set. *Journal of forecasting*, 19(3):219–230.
- Odom, M. D. and Sharda, R. (1990). A neural network model for bankruptcy prediction. In *1990 IJCNN International Joint Conference on neural networks*, pages 163–168.
- Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, pages 109–131.
- Pantalone, C. C., Platt, M. B., et al. (1987). Predicting commercial bank failure since deregulation. *New England Economic Review*, (Jul):37–47.
- Sánchez-Lasheras, F., de Andrés, J., Lorca, P., and de Cos Juez, F. J. (2012). A hybrid device for the solution of sampling bias problems in the forecasting of firms’ bankruptcy. *Expert Systems with Applications*, 39(8):7512–7523.

- Sinkey, J. F. (1979). *Problem and failed institutions in the commercial banking industry*. Wiley Online Library.
- Thomson, J. B. (1991). Predicting bank failures in the 1980s. *Economic Review-Federal Reserve Bank of Cleveland*, 27(1):9.
- West, M., Harrison, P. J., and Migon, H. S. (1985). Dynamic generalized linear models and bayesian forecasting. *Journal of the American Statistical Association*, 80(389):73–83.
- Wiginton, J. C. (1980). A note on the comparison of logit and discriminant models of consumer credit behavior. *Journal of Financial and Quantitative Analysis*, 15(03):757–770.
- Zhang, G., Hu, M. Y., Patuwo, B. E., and Indro, D. C. (1999). Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European journal of operational research*, 116(1):16–32.
- Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting research*, pages 59–82.

Appendix A: Basis function

$$\begin{aligned}
 BF2 &= \max(0, -0.408772 - EQTA); \\
 BF3 &= \max(0, ROA + 0.365441) * BF2; \\
 BF9 &= \max(0, EQTL + 0.683567) * BF2; \\
 BF11 &= \max(0, EQTL + 0.758831) * BF2; \\
 BF13 &= \max(0, EQTL + 0.666082) * BF2; \\
 BF15 &= \max(0, EQTL + 0.600936) * BF2; \\
 BF24 &= \max(0, 4.57365 - NPLTA); \\
 BF28 &= \max(0, -0.600855 - EQTA) * BF24; \\
 BF30 &= \max(0, -0.535146 - EQTL) * BF24; \\
 BF38 &= \max(0, 3.58056 - NPLTA); \\
 BF46 &= \max(0, 2.22526 - NPLGL) * BF38; \\
 BF58 &= \max(0, 1.09343 - TDTA); \\
 BF91 &= \max(0, NPLGL - 1.47036) * BF58.
 \end{aligned}$$

Basis functions For 2009

The final model is expressed as follows:

$$\begin{aligned}
 Y = & 0.919013 + 1.43553 * BF3 + 50.2014 * BF9 - 9.96475 * BF11 - 52.2465 * BF13 + \\
 & 12.2264 * BF15 - 0.252718 * BF24 + 0.446248 * BF28 - 0.778493 * BF30 + \\
 & 0.0330096 * BF46 - 0.231863 * BF91
 \end{aligned}
 \tag{8}$$

The basis function BF2 does not appear in the model but it contributes in the construction of others basis function (BF3, BF9, BF11, BF13 and BF15). The value of the BF2 is positive when the value of EQTA is inferior than -0.4087. The BF3 has a positive effect on the target value when the value of ROA is bigger than -0.365, multiplied by BF2. A negative impact on the output variable appears for values of EQTL above -0.7588, -0.666 and -0.6009, multiplied by BF2. We note that the variable NPLTA (BF24) affects negatively the target for value weaker than 4.57. Moreover, we note the presence of an interaction between the BF24 and two others variables (EQTA and EQTL). The positive effect of the BF28 appears when EQTA is lower than -0.66 and for a value of EQTL inferior than -0.535 the BF30 impacts negatively the output variable. The BF46 has a slight positive effect on the target only when NPLTA is lower than 3.58 and NPLGL isles than 2.225. For a value of TDTA below 1.093 and for value of NPLGL above 1.47, the BF91 appears and impact negatively the target variable.

$$\begin{aligned}
 &BF2 = \max(0, -0.468008 - EQTL); \\
 &BF3 = \max(0, ROA + 1.97633) * BF2; \\
 \text{Basis functions For 2010} \quad &BF4 = \max(0, -1.97633 - ROA) * BF2; \\
 &BF6 = \max(0, -1.54086 - EQTA) * BF2; \\
 &BF14 = \max(0, -0.0634867 - ROA); \\
 &BF16 = \max(0, -0.744265 - EQTL) * BF14;
 \end{aligned}$$

The final Model is expressed as below:

$$\begin{aligned}
 Y = 0.00323519 + 0.766835 * BF2 + 0.875269 * BF3 - 1.13558 * BF4 - 1.51414 * BF6 + \\
 1.584 * BF16
 \end{aligned}
 \tag{9}$$

We note that the basis function BF14 is not involved in the model but it's used to compute the BF16, meaning that there is an interaction between ROA and EQTL. We note that the positive effect of the BF16 on the model appears only when the value of EQTL is less than -0.744 and ROA is inferior than -0.063. The basis function BF3 and BF4 define a piecewise linear function with a knot or threshold of -1.9733, multiplied by the BF2. According to the definition of these basis functions, we conclude that the relationship between the target and the predictor variables is nonlinear. In fact, a value of ROA greater than -1.97 and a value EQTL lower than -0.468, have a positive effect on the output variable. But if the ROA and the EQTL are both below -1.97 and -0.468 respectively, it has a negative impact on the target. As mentioned earlier, the EQTL impact positively the output variable for a value below -0.468.

$$\begin{aligned}
& BF2 = \max(0, -0.63523 - EQTL); \\
& BF4 = \max(0, 0.467418 - NPLGL) * BF2; \\
& BF5 = \max(0, EQTA + 1.70669) * BF2; \\
& BF15 = \max(0, ROA + 1.34293) * BF2; \\
\text{Basis functions For 2011 } & BF25 = \max(0, ROA + 2.00727) * BF2; \\
& BF31 = \max(0, ROA + 1.15584); \\
& BF35 = \max(0, LLRGL - 2.29337); \\
& BF37 = \max(0, LLRGL - 1.67056); \\
& BF54 = \max(0, -0.856318 - EQTA) * BF31;
\end{aligned}$$

The final model is expressed as:

$$\begin{aligned}
Y = -2.19129e - 005 + 1.58255 * BF2 - 2.77362 * BF4 - 5.5853 * BF5 - 6.9801 * BF15 + \\
2.63976 * BF25 - 0.499781 * BF35 + 0.426016 * BF37 + 1.4204 * BF54
\end{aligned} \tag{10}$$

As it can be seen, the BF2 appears only when the value of EQTL is lower than -0.635. The BF2 was used in the construction of other basis functions (BF4, BF5 and BF15). In fact, the BF4 is defined as an interaction between NPLGL and EQTL. We also note that the BF4, BF5 and BF15 have a negative impact on the target variable for values of NPLGL inferior than 0.467, EQTA upper than -1.7068 and ROA superior than -1.3429, multiplied by BF2. However, a value of ROA upper than -2.007 multiplied by BF2 affects positively the variable 'Y'. When the value of LLRGL is bigger than 2.293, we note a negative effect while if the value of LLRGL is above 1.6705 we remark a positive impact on the target value.

$$\begin{aligned}
& BF11 = \max(0, EQTA + 2.18488); \\
& BF13 = \max(0, EQTA + 1.77674); \\
\text{Basis functions For 2012 } & BF14 = \max(0, -1.77674 - EQTA); \\
& BF15 = \max(0, EQTL + 0.874739) * BF14; \\
& BF18 = \max(0, 1.20033 - NPLTA) * BF14;
\end{aligned}$$

The final model is expressed as:

$$Y = 0.832567 - 2.03375 * BF11 + 2.03275 * BF13 + 30.847 * BF15 - 1.14783 * BF18 \tag{11}$$

The BF13 and BF14 together define a piecewise linear function of EQTA with a knot of -1.776. When the EQTA is bigger than -1.776, we note a positive effect on the target variable, but for a value of EQTA above -2.184, the BF11 impact negatively the output variable. It appears that the basis function BF14 is not associated to the model but it's used to compute the BF15 and BF18. We note that the interaction between EQTA and EQTL (BF15) has a positive effect on the target variables for a value of EQTL upper than -0.874 and for EQTA lower than -1.776. Moreover, the negative impact of the BF18 appears only when NPLTA is below 1.2 and EQTA lower than -1.779.

List of Tables

1	MARS models: N Basis Functions and Error Measures	21
2	Variable importance: MARS Model	22
3	ANOVA Decomposition	23
4	Final Model (After Backward Stepwise Elimination)	24
5	Root Node Competitors in Order of Improvement	25
6	Variables Importance: CART Model	26
7	MARS accuracy and prediction results	26
8	CART accuracy and prediction results	27
9	K-means accuracy and prediction results	27
10	K-means MARS accuracy and prediction results	28
11	Results in average	28
12	Area Under Curve (AUC) results	28

List of Figures

1	Relationship between variables based on MARS output (2013)	20
2	Classification Tree diagram	29
3	Principal Component Analysis PCA	32
4	ROC curve MARS	33
5	ROC curve CART	34
6	ROC curve K-means MARS	35

Figure 1: Relationship between variables based on MARS output (2013)

variables/variables MARS 2013.png

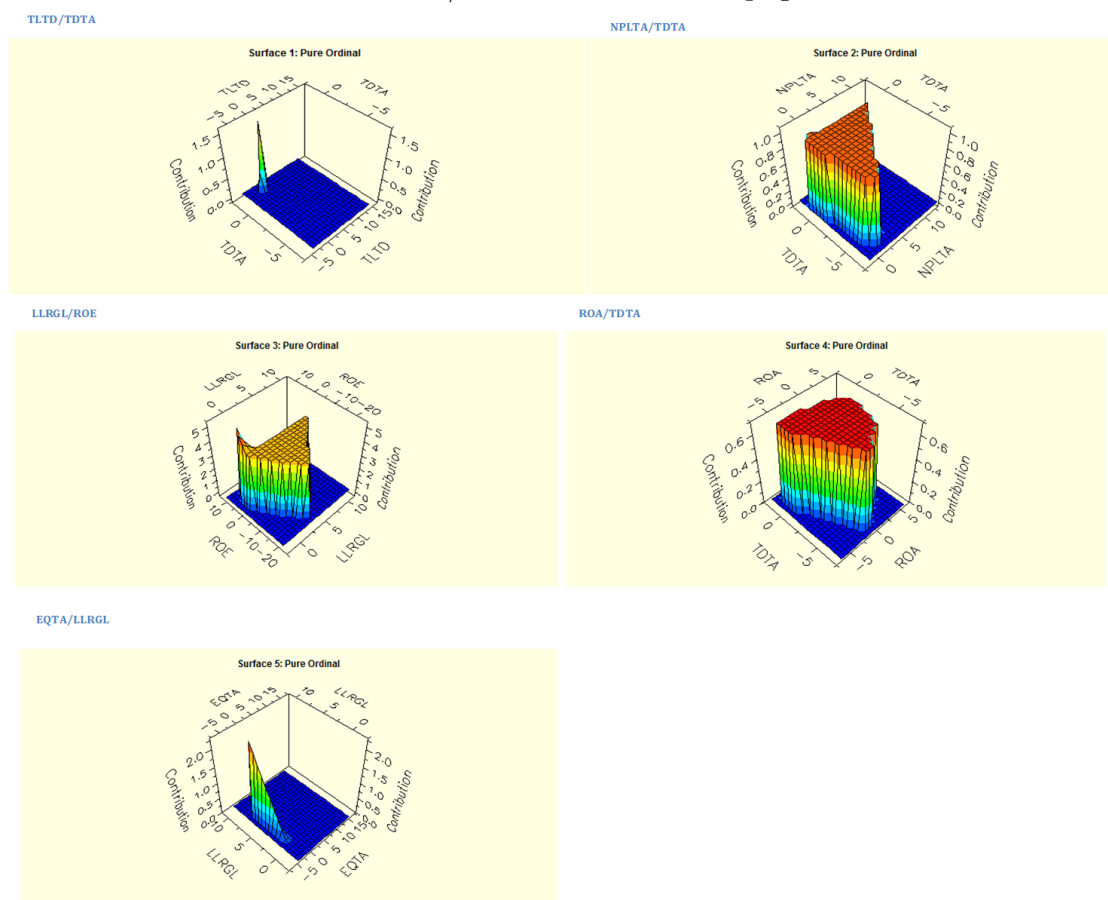


Table 1: MARS models: N Basis Functions and Error Measures

	Model	N,BFs	GCV	MSE	TEST,MSE	Adj,MSE	ROC
2008	1	2	0,01657	0,01639	0,01853	0,01635	0,84412
	2	2	0,01657	0,01639	0,0185	0,01635	0,84412
	3	2	0,01657	0,01639	0,01898	0,01635	0,84412
	4	2	0,01657	0,01639	0,01846	0,01635	0,84412
	5	2	0,01657	0,01639	0,01826	0,01635	0,84412
	6	1	0,01851	0,01839	0,01913	0,01837	0,84349
	7*	4	0,01536	0,01505	0,01713	0,01499	0,86556
	8	2	0,01657	0,01639	0,01915	0,01635	0,84412
	9	2	0,01657	0,01639	0,01966	0,01635	0,84412
	10	1	0,01851	0,01839	0,01901	0,01837	0,84349
	Model	N,BFs	GCV	MSE	TEST,MSE	Adj,MSE	ROC
2009	1	4	0.04720	0.04623	0.05040	0.04604	0.94741
	2	8	0.04269	0.04101	0.08550	0.04071	0.95010
	3	5	0.04620	0.04503	0.04809	0.04481	0.95427
	4	4	0.04720	0.04623	0.05069	0.04604	0.94741
	5	4	0.04720	0.04623	0.04708	0.04604	0.94741
	6	5	0.04620	0.04503	0.04985	0.04481	0.95427
	7*	10	0.04037	0.03841	0.05016	0.03806	0.94940
	8	4	0.04720	0.04623	0.04950	0.04604	0.94741
	9	5	0.04620	0.04503	0.04848	0.04481	0.95427
	10	5	0.04620	0.04503	0.04980	0.04481	0.95427
	Model	N,BFs	GCV	MSE	TEST,MSE	Adj,MSE	ROC
2010	1*	5	0.02992	0.02918	0.03297	0.02902	0.96544
	2	5	0.02992	0.02918	0.03429	0.02902	0.96544
	3	4	0.03241	0.03175	0.03332	0.03160	0.96566
	4	3	0.03673	0.03615	0.03824	0.03602	0.96438
	5	3	0.03673	0.03615	0.03458	0.03602	0.96438
	6	4	0.03241	0.03175	0.03591	0.03160	0.96566
	7	4	0.03241	0.03175	0.03374	0.03160	0.96566
	8	5	0.02992	0.02918	0.03414	0.02902	0.96544
	9	5	0.02992	0.02918	0.03305	0.02902	0.96544
	10	5	0.02992	0.02918	0.03240	0.02902	0.96544
	Model	N,BFs	GCV	MSE	TEST,MSE	Adj,MSE	ROC
2011	1	4	0.01834	0.01789	0.02799	0.01780	0.98175
	2	5	0.01835	0.01780	0.02445	0.01769	0.98828
	3	9	0.01556	0.01475	0.02991	0.01460	0.99012
	4	5	0.01835	0.01780	0.03001	0.01769	0.98828
	5*	8	0.01548	0.01476	0.02301	0.01462	0.99014
	6	7	0.01613	0.01547	0.02265	0.01534	0.99000
	7	7	0.01613	0.01547	0.02471	0.01534	0.99000
	8	5	0.01835	0.01780	0.02830	0.01769	0.98828
	9	2	0.02122	0.02094	0.02587	0.02088	0.98010
	10	3	0.02048	0.02009	0.02666	0.02001	0.98047
	Model	N,BFs	GCV	MSE	TEST,MSE	Adj,MSE	ROC
2012	1	3	0.00844	0.00826	0.01209	0.00823	0.99639
	2	3	0.00844	0.00826	0.01421	0.00823	0.99639
	3*	4	0.00642	0.00625	0.01208	0.00622	0.99808
	4	2	0.01310	0.01292	0.01792	0.01287	0.99423
	5	2	0.01310	0.01292	0.01415	0.01287	0.99423
	6	2	0.01310	0.01292	0.01622	0.01287	0.99423
	7	5	0.00645	0.00624	0.01439	0.00620	0.99793
	8	3	0.00844	0.00826	0.01383	0.00823	0.99639
	9	5	0.00645	0.00624	0.01422	0.00620	0.99793
	10	4	0.00642	0.00625	0.01626	0.00622	0.99808
	Model	N,BFs	GCV	MSE	TEST,MSE	Adj,MSE	ROC
2013	1	2	0.01030	0.01013	0.01635	0.01009	0.82855
	2	2	0.01030	0.01013	0.01085	0.01009	0.82855
	3	3	0.00977	0.00954	0.01193	0.00950	0.92831
	4	1	0.01140	0.01129	0.01280	0.01127	0.79689
	5	2	0.01030	0.01013	0.01434	0.01009	0.82855
	6	1	0.01140	0.01129	0.01568	0.01127	0.79689
	7	1	0.01140	0.01129	0.01401	0.01127	0.79689
	8	1	0.01140	0.01129	0.01514	0.01127	0.79689
	9	1	0.01140	0.01129	0.01404	0.01127	0.79689
	10*	6	0.00699	0.00669	0.01502	0.00663	0.89844

Table 2: Variable importance: MARS Model

	EQTA	EQTL	LLRGL	LLRTA	NPLGL	NPLTA	ROA	ROE	TDTA	TLTD
2008	29,61	100,00	0,00	0,00	54,80	0,00	0,00	0,00	0,00	0,00
2009	51,42	32,78	0,00	0,00	21,78	100,00	26,00	0,00	15,00	0,00
2010	20,70	100,00	0,00	0,00	0,00	0,00	31,70	0,00	0,00	0,00
2011	25,00	100,00	23,50	0,00	28,62	0,00	33,15	0,00	0,00	0,00
2012	100,00	36,61	0,00	0,00	0,00	26,10	0,00	0,00	0,00	0,00
2013	88,19	0,00	87,84	0,00	0,00	61,80	49,89	87,00	100,00	89,29

Table 3: ANOVA Decomposition

	fun	std. dev,	gcv	#bsfns	variable
2008	1	0,05534	0,01598	1	EQTL
	2	0,12941	0,01943	2	EQTL
	3	0,0773	0,01655	1	NPLGL EQTA EQTL
2009	1	0,24338	0,04726	1	NPLTA
	2	0,06391	0,04368	1	EQTA ROA
	3	0,1394	0,04581	4	EQTA EQTL
	4	0,51074	0,04796	1	EQTA NPLTA
	5	0,31745	0,04354	1	EQTL NPLTA
	6	0,13021	0,04207	1	NPLTA NPLGL
	7	0,03988	0,04147	1	NPLGL TDTA
2010	1	0,11914	0,03241	1	EQTL
	2	0,24052	0,03701	3	EQTL ROA
	3	0,1355	0,03295	1	EQTA EQTL
2011	1	0,16593	0,02435	1	EQTL
	2	0,0466	0,01713	2	LLRGL
	3	0,0526	0,01793	1	EQTL NPLGL
	4	0,04383	0,01704	1	EQTA EQTL
	5	0,17821	0,01877	2	EQTL ROA
	6	0,20181	0,01613	1	EQTA ROA
2012	1	0,13739	0,01941	2	EQTA
	2	0,07162	0,01039	1	EQTA EQTL
	3	0,0487	0,00844	1	EQTA NPLTA
2013	1	0,13421	0,0104	1	ROE
	2	0,07546	0,01053	1	TLTD TDTA
	3	0,08686	0,00869	1	NPLTA TDTA
	4	0,14404	0,00954	1	LLRGL ROE
	5	0,04186	0,0081	1	ROA TDTA
	6	0,15955	0,01044	1	EQTA LLRGL

Table 4: Final Model (After Backward Stepwise Elimination)

	Basis Fun	Coefficient	Variable	Parent	Knot
2008	0	0,00704			
	2	1,09848	EQTL		-0,47162
	3	-4,41337	NPLGL	EQTL	1,33738
	6	-1,3883	EQTA	EQTL	-0,87975
	7	4,57587	NPLGL	EQTL	0,70157
2009	0	0,91901			
	3	1,43553	ROA	EQTA	-0,36544
	9	50,20139	EQTL	EQTA	-0,68357
	11	-9,96475	EQTL	EQTA	-0,75883
	13	-52,24647	EQTL	EQTA	-0,66608
	15	12,22639	EQTL	EQTA	-0,60094
	24	-0,25272	NPLTA		4,57365
	28	0,44625	EQTA	NPLTA	-0,60085
	30	-0,77849	EQTL	NPLTA	-0,53515
	46	0,03301	NPLGL	NPLTA	2,22526
	91	-0,23186	NPLGL	TDTA	1,47036
2010	0	0,00324			
	2	0,76683	EQTL		-0,46801
	3	0,87527	ROA	EQTL	-1,97633
	4	-1,13558	ROA	EQTL	-1,97633
	6	-1,51414	EQTA	EQTL	-1,54086
	16	1,584	EQTL	ROA	-0,74427
2011	0	-0,00002			
	2	1,58255	EQTL		-0,63523
	4	-2,77362	NPLGL	EQTL	0,46742
	5	-5,5853	EQTA	EQTL	-1,70669
	15	-6,9801	ROA	EQTL	-1,34293
	25	2,63976	ROA	EQTL	-2,00727
	35	-0,49978	LLRGL		2,29337
	37	0,42602	LLRGL		1,67056
54	1,4204	EQTA	ROA	-0,85632	
2012	0	0,83257			
	11	-2,03375	EQTA		-2,18488
	13	2,03275	EQTA		-1,77674
	15	30,84698	EQTL	EQTA	-0,87474
	18	-1,14783	NPLTA	EQTA	1,20033
2013	0	0,00367			
	10	0,46647	ROE		0,08003
	28	3,07491	TLTD	TDTA	-0,75584
	59	-0,24651	NPLTA	TDTA	1,16217
	63	-0,22463	LLRGL	ROE	-1,61568
	70	-2,15719	ROA	TDTA	-0,73126
	88	0,24169	LLRGL	EQTA	1,15542

Table 5: Root Node Competitors in Order of Improvement

	Competitor	Split	Improvement	Left Node	Right Node
2008	NPLGL	0,7528253	0,3094293	1084	160
	EQTA	-0,7376655	0,3068252	164	1080
	NPLTA	0,7843314	0,2875027	1083	161
	EQTL	-0,3783329	0,266294	230	1014
	ROA	-0,1995939	0,2482743	301	943
	ROE	-0,04969738	0,2384608	243	1001
	LLRTA	0,3922257	0,2367943	998	246
	LLRGL	0,0615043	0,2217104	891	353
	TDTA	0,4067549	0,0895273	774	470
	TLTD	0,4128059	0,0473563	898	346
2009	EQTA	-0,7851334	0,3129933	227	983
	EQTL	-0,4295057	0,3014897	262	948
	NPLGL	-0,03432735	0,2247864	791	419
	NPLTA	-0,1403783	0,2100617	751	459
	ROA	0,09301518	0,204073	434	776
	LLRGL	0,215884	0,1628053	865	345
	LLRTA	0,3042838	0,1568755	908	302
	TDTA	0,7302819	0,1429197	1030	180
	ROE	-0,03321227	0,1046313	105	1105
	TLTD	0,2019615	0,0291358	707	503
2010	EQTL	-0,6625983	0,3710726	175	904
	EQTA	-0,961138	0,3700976	194	885
	NPLGL	0,1238613	0,2731861	782	297
	NPLTA	-0,004726889	0,2706729	746	333
	ROA	0,2448697	0,248093	404	675
	LLRTA	0,1336472	0,2020354	766	313
	LLRGL	0,2257459	0,1898762	775	304
	TDTA	0,663318	0,1647155	904	175
	ROE	0,05435728	0,1381864	245	834
	TLTD	0,4184276	0,0274782	822	257
2011	EQTA	-1,067187	0,4279733	127	830
	EQTL	-0,5830956	0,4220076	133	824
	NPLGL	0,4071394	0,3508578	762	195
	ROA	-0,04254466	0,3269411	258	699
	NPLTA	0,3651701	0,3246399	762	195
	LLRGL	0,341674	0,2672399	736	221
	LLRTA	0,2904416	0,2521601	731	226
	ROE	-0,08649313	0,2409426	117	840
	TDTA	0,7509308	0,2022613	856	101
	TLTD	0,2879991	0,0637466	632	325
2012	EQTA	-1,40872	0,4725714	57	827
	EQTL	-0,5921312	0,4725714	57	827
	NPLTA	0,8683394	0,3938451	781	103
	ROA	-0,397359	0,3892372	139	745
	NPLGL	0,8980433	0,3694974	785	99
	ROE	0,0184353	0,3132174	57	827
	LLRGL	0,2599867	0,2839672	655	229
	LLRTA	0,5624378	0,2621549	723	161
	TDTA	0,8841329	0,2619614	838	46
	TLTD	0,0367059	0,1030475	450	434
2013	EQTA	-1,152309	0,4754959	36	815
	EQTL	-0,5505671	0,466474	44	807
	ROA	-0,6195569	0,4176729	90	761
	TDTA	0,9668738	0,3185172	826	25
	LLRGL	1,029036	0,3082259	769	82
	NPLGL	0,3425343	0,2937366	694	157
	NPLTA	0,3336544	0,2928898	693	158
	ROE	0,01182485	0,2906014	51	800
	LLRTA	0,7617488	0,279998	739	112
	TLTD	-0,03057846	0,1813366	406	445

Table 6: Variables Importance: CART Model

	EQTA	EQTL	LLRGL	LLRTA	NPLGL	NPLTA	ROA	ROE	TDTA	TLTD
2008	100,00	86,14	73,31	75,73	83,49	76,37	10,54	11,66	2,72	4,42
2009	100,00	92,86	49,55	0,00	70,11	72,38	58,44	0,00	0,00	0,00
2010	100,00	99,85	0,64	53,41	77,41	76,57	68,29	5,14	1,80	0,01
2011	100,00	99,67	64,82	2,84	84,51	78,27	77,22	1,94	0,00	0,23
2012	100,00	100,00	0,00	0,00	77,98	83,34	82,37	65,29	0,00	0,00
2013	100,00	98,59	65,40	0,24	62,35	0,72	89,80	0,00	67,07	1,23

Table 7: MARS accuracy and prediction results

MARS	2008		2009		2010	
	Learn	Test	Learn	Test	Learn	Test
Failed banks correctly predicted	28	31	115	115	115	115
Non-failed banks correctly predicted	1137	1129	1019	961	898	886
Failed banks incorrectly predicted	9	6	16	16	7	7
Non-failed banks incorrectly predicted	70	78	60	118	59	71
Incorrectly predicted total	79	84	76	134	66	78
Correctly predicted total	1165	1160	1134	1076	1013	1001
% of failed banks correctly predicted (Sensitivity)	75,68%	83,78%	87,79%	87,79%	94,26%	94,26%
% of error type I	24,32%	16,22%	12,21%	12,21%	5,74%	5,74%
% of non-failed banks correctly predicted (specificity)	94,20%	93,54%	94,44%	89,06%	93,83%	92,58%
% of error type II	5,80%	6,46%	5,56%	10,94%	6,17%	7,42%
% of total incorrectly predicted (error rate)	6,35%	6,75%	6,28%	11,07%	6,12%	7,23%
% of total correctly predicted (correct classification rate)	93,65%	93,25%	93,72%	88,93%	93,88%	92,77%
MARS	2011		2012		2013	
	Learn	Test	Learn	Test	Learn	Test
Failed banks correctly predicted	72	71	32	32	13	14
Non-failed banks correctly predicted	855	848	847	835	828	813
Failed banks incorrectly predicted	1	2	1	1	2	1
Non-failed banks incorrectly predicted	29	36	4	16	8	23
Incorrectly predicted total	30	38	5	17	10	24
Correctly predicted total	927	919	879	867	841	827
% of failed banks correctly predicted (Sensitivity)	98,63%	97,26%	96,97%	96,97%	86,67%	93,33%
% of error type I	1,37%	2,74%	3,03%	3,03%	13,33%	6,67%
% of non-failed banks correctly predicted (specificity)	96,72%	95,93%	99,53%	98,12%	99,04%	97,25%
% of error type II	3,28%	4,07%	0,47%	1,88%	0,96%	2,75%
% of total incorrectly predicted (error rate)	3,13%	3,97%	0,57%	1,92%	1,18%	2,82%
% of total correctly predicted (correct classification rate)	96,87%	96,03%	99,43%	98,08%	98,82%	97,18%

Table 8: CART accuracy and prediction results

CART	2008		2009		2010	
	Learn	Test	Learn	Test	Learn	Test
Failed banks correctly predicted	35	34	122	119	118	114
Non-failed banks correctly predicted	1120	1097	955	959	903	897
Failed banks incorrectly predicted	2	3	9	12	4	8
Non-failed banks incorrectly predicted	87	110	124	120	54	60
Incorrectly predicted total	89	113	133	132	58	68
Correctly predicted total	1155	1131	1077	1078	1021	1011
% of failed banks correctly predicted (Sensitivity)	94,59%	91,89%	93,13%	90,84%	96,72%	93,44%
% of error type I	5,41%	8,11%	6,87%	9,16%	3,28%	6,56%
% of non-failed banks correctly predicted (specificity)	92,79%	90,89%	88,51%	88,88%	94,36%	93,73%
% of error type II	7,21%	9,11%	11,49%	11,12%	5,64%	6,27%
% of total incorrectly predicted (error rate)	7,15%	9,08%	10,99%	10,91%	5,38%	6,30%
% of total correctly predicted (correct classification rate)	92,85%	90,92%	89,01%	89,09%	94,62%	93,70%

CART	2011		2012		2013	
	Learn	Test	Learn	Test	Learn	Test
Failed banks correctly predicted	73	70	33	32	15	14
Non-failed banks correctly predicted	841	834	827	828	830	828
Failed banks incorrectly predicted	0	3	0	1	0	1
Non-failed banks incorrectly predicted	43	50	24	23	6	8
Incorrectly predicted total	43	53	24	24	6	9
Correctly predicted total	914	904	860	860	845	842
% of failed banks correctly predicted (Sensitivity)	100,00%	95,89%	100,00%	96,97%	100,00%	93,33%
% of error type I	0,00%	4,11%	0,00%	2,70%	0,00%	6,67%
% of non-failed banks correctly predicted (specificity)	95,14%	94,34%	97,18%	97,30%	99,28%	99,04%
% of error type II	4,86%	5,66%	2,82%	2,70%	0,72%	0,96%
% of total incorrectly predicted (error rate)	4,49%	5,54%	2,71%	2,71%	0,71%	1,06%
% of total correctly predicted (correct classification rate)	95,51%	94,46%	97,29%	97,29%	99,29%	98,94%

Table 9: K-means accuracy and prediction results

K-means clustering	2008	2009	2010	2011	2012	2013
Failed banks correctly predicted	33	109	116	73	33	12
Non-failed banks correctly predicted	1077	895	822	795	801	806
Failed banks incorrectly predicted	4	22	6	0	0	3
Non-failed banks incorrectly predicted	130	184	135	89	50	30
Incorrectly predicted total	134	206	141	89	50	33
Correctly predicted total	1110	1004	938	868	834	818
% of failed banks correctly predicted (Sensitivity)	89,19%	83,21%	95,08%	100,00%	100,00%	80,00%
% of error type I	10,81%	16,79%	4,92%	0,00%	0,00%	20,00%
% of non-failed banks correctly predicted (Specificity)	89,23%	82,95%	85,89%	89,93%	94,12%	96,41%
% of error type II	10,77%	17,05%	14,11%	10,07%	5,88%	3,59%
% of total incorrectly predicted (error rate)	10,77%	17,02%	13,07%	9,30%	5,66%	3,88%
% of total correctly predicted (correct classification rate)	89,23%	82,98%	86,93%	90,70%	94,34%	96,12%

Table 10: K-means MARS accuracy and prediction results

K-means MARS	2008		2009		2010	
	Learn	Test	Learn	Test	Learn	Test
Failed banks correctly predicted	163	161	292	291	249	245
Non-failed banks correctly predicted	1071	1033	898	888	824	808
Failed banks incorrectly predicted	0	2	1	2	2	6
Non-failed banks incorrectly predicted	10	48	19	29	4	20
Incorrectly predicted total	10	50	20	31	6	26
Correctly predicted total	1234	1194	1190	1179	1073	1053
% of failed banks correctly predicted (Sensitivity)	100,00%	98,77%	99,66%	99,32%	99,20%	97,61%
% of error type I	0,00%	1,23%	0,34%	0,68%	0,80%	2,39%
% of non-failed banks correctly predicted (specificity)	99,07%	95,56%	97,93%	96,84%	99,52%	97,58%
% of error type II	0,93%	4,44%	2,07%	3,16%	0,48%	2,42%
% of total incorrectly predicted (error rate)	0,80%	4,02%	1,65%	2,56%	0,56%	2,41%
% of total correctly predicted (correct classification rate)	99,20%	95,98%	98,35%	97,44%	99,44%	97,59%

K-means MARS	2011		2012		2013	
	Learn	Test	Learn	Test	Learn	Test
Failed banks correctly predicted	162	158	83	82	42	42
Non-failed banks correctly predicted	786	781	792	770	792	775
Failed banks incorrectly predicted	0	4	0	1	0	0
Non-failed banks incorrectly predicted	9	14	9	31	17	34
Incorrectly predicted total	9	18	9	32	17	34
Correctly predicted total	948	939	875	852	834	817
% of failed banks correctly predicted (Sensitivity)	100,00%	97,53%	100,00%	98,80%	100,00%	100,00%
% of error type I	0,00%	2,47%	0,00%	1,20%	0,00%	0,00%
% of non-failed banks correctly predicted (specificity)	98,87%	98,24%	98,88%	96,13%	97,90%	95,80%
% of error type II	1,13%	1,76%	1,12%	3,87%	2,10%	4,20%
% of total incorrectly predicted (error rate)	0,94%	1,88%	1,02%	3,62%	2,00%	4,00%
% of total correctly predicted (correct classification rate)	99,06%	98,12%	98,98%	96,38%	98,00%	96,00%

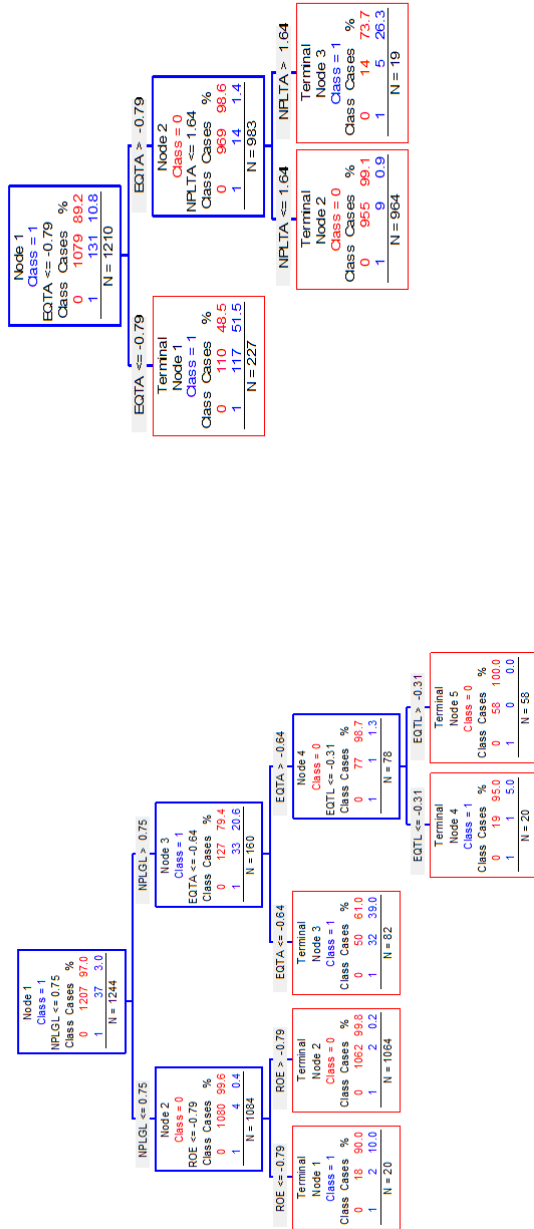
Table 11: Results in average

CART	Correct Classification	94,76%	94,07%
MARS	Correct Classification	96,06%	94,37%
K-MARS	Correct Classification	98,84%	96,92%
CART	Error Type I	2,59%	6,22%
MARS	Error Type I	10,00%	7,77%
K-MARS	Error Type I	0,19%	1,33%
CART	Error type II	5,46%	5,97%
MARS	Error type II	3,71%	5,59%
K-MARS	Error type II	1,31%	3,31%

Table 12: Area Under Curve (AUC) results

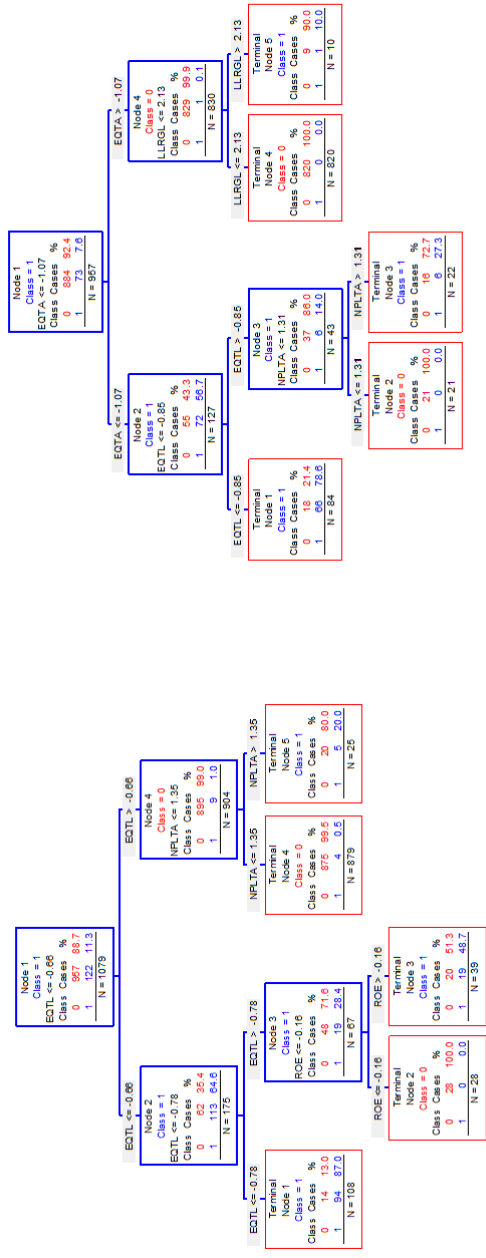
ROC (Area Under Curve)	MARS		CART		K-MARS	
	Training	Testing	Training	Testing	Training	Testing
2008	86,56%	88,94%	95,00%	92,19%	99,99%	99,26%
2009	95,03%	94,16%	91,20%	90,25%	99,94%	99,47%
2010	96,54%	96,27%	97,17%	94,80%	99,98%	99,41%
2011	99,00%	97,10%	98,78%	96,33%	99,96%	99,26%
2012	99,81%	99,28%	98,59%	97,20%	99,97%	99,46%
2013	89,86%	99,08%	99,64%	96,41%	99,72%	99,71%
Average	94,47%	95,81%	96,73%	94,53%	99,93%	99,43%

Figure 2: Classification Tree diagram



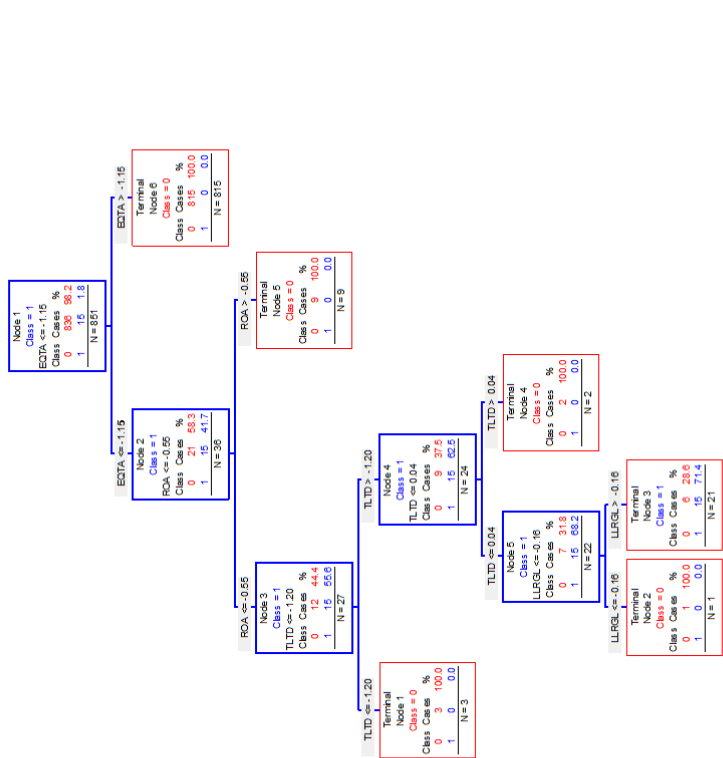
Classification Tree diagram in 2009

Classification Tree diagram in 2008



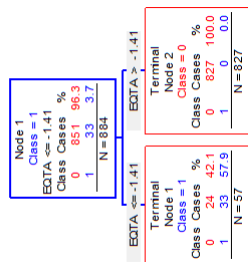
Classification Tree diagram in 2011

Classification Tree diagram in 2010



Classification Tree diagram in 2013

Classification Tree diagram in 2013



Classification Tree diagram in 2012

Figure 3: Principal Component Analysis PCA

variables/PCA.png

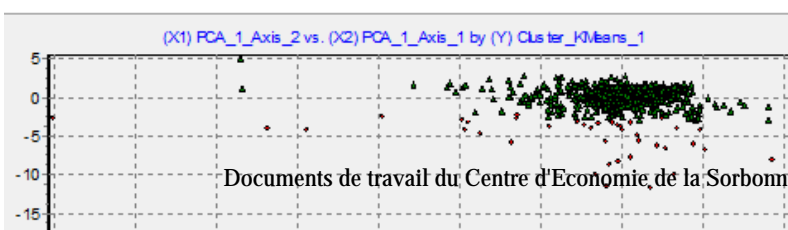
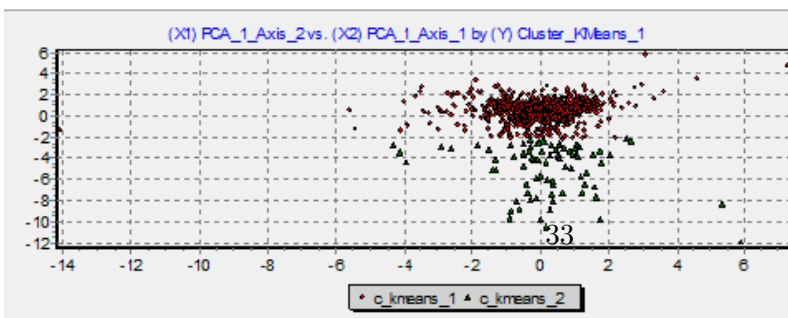
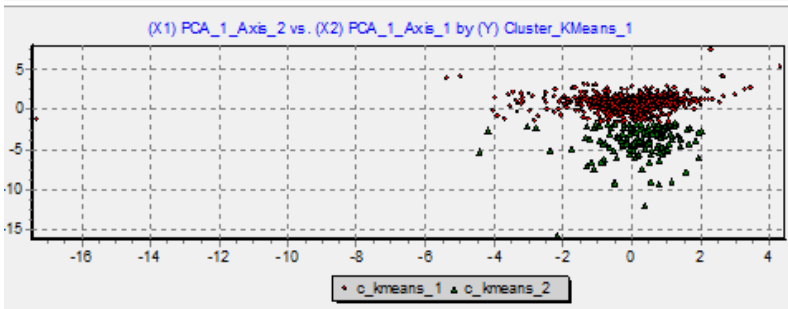
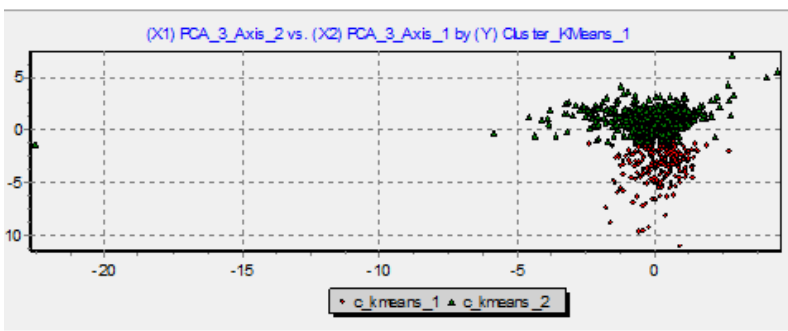
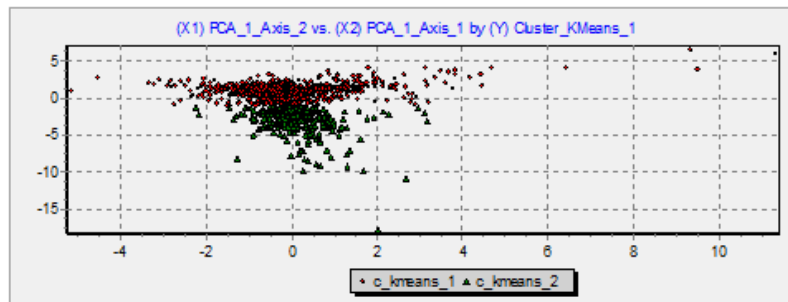
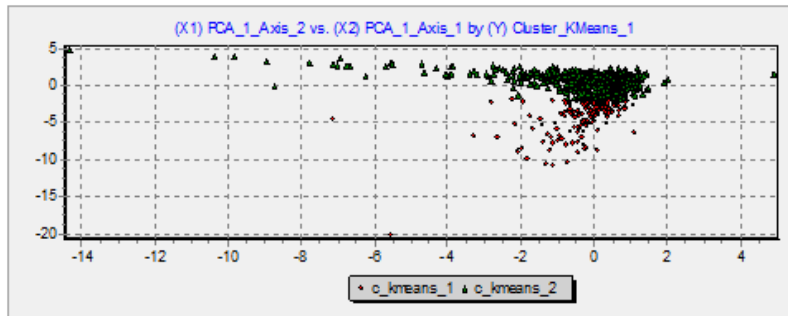
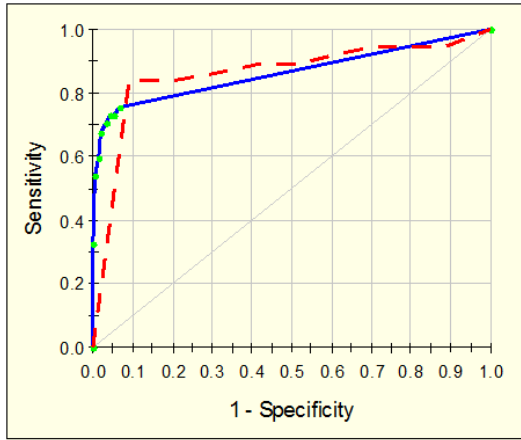
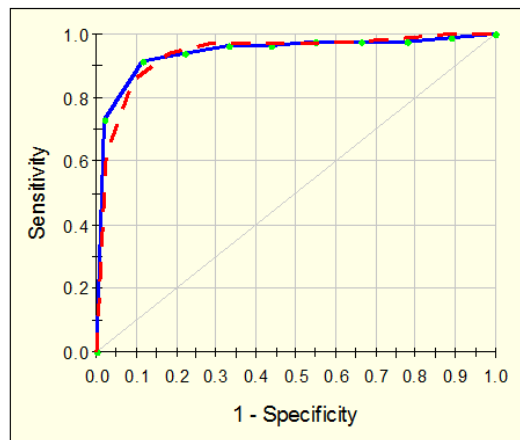


Figure 4: ROC curve MARS

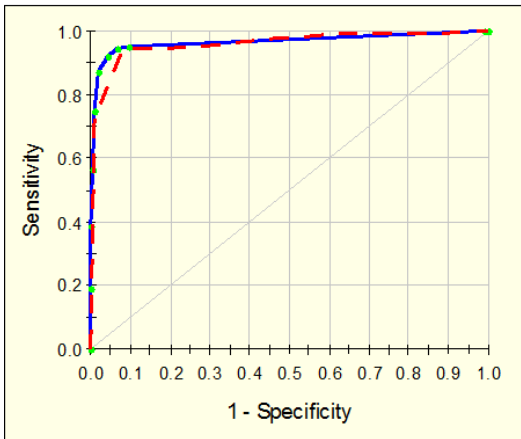
variables/ROCcurve MARS.png



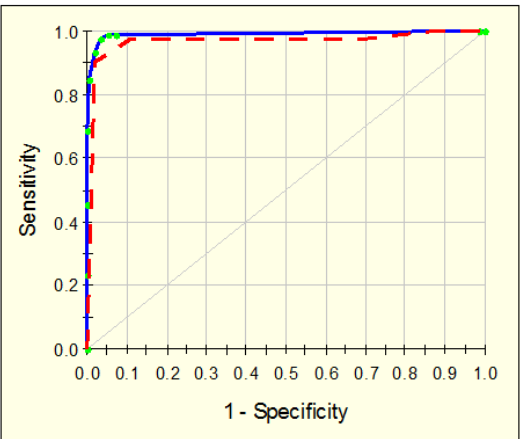
MARS ROC curve 2008



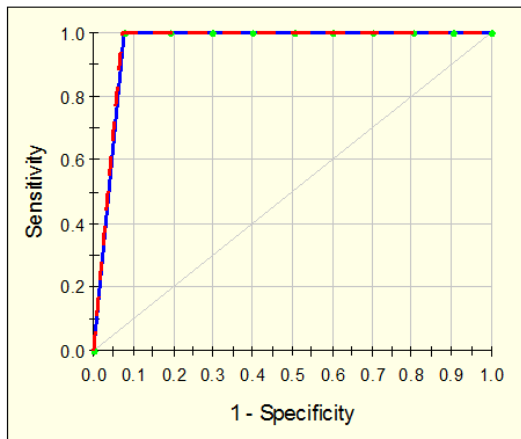
MARS ROC curve 2009



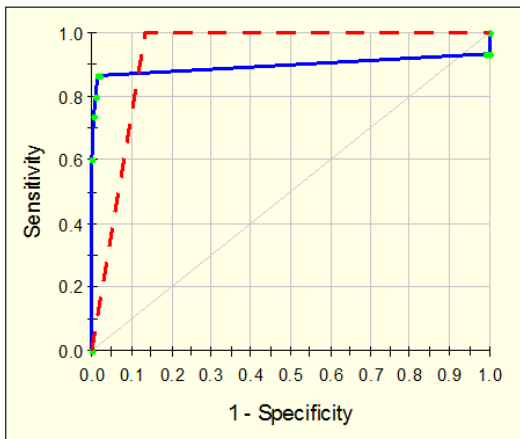
MARS ROC curve 2010



MARS ROC curve 2011



MARS ROC curve 2012



MARS ROC curve 2013

Figure 5: ROC curve CART

variables/ROCcurve CART.png

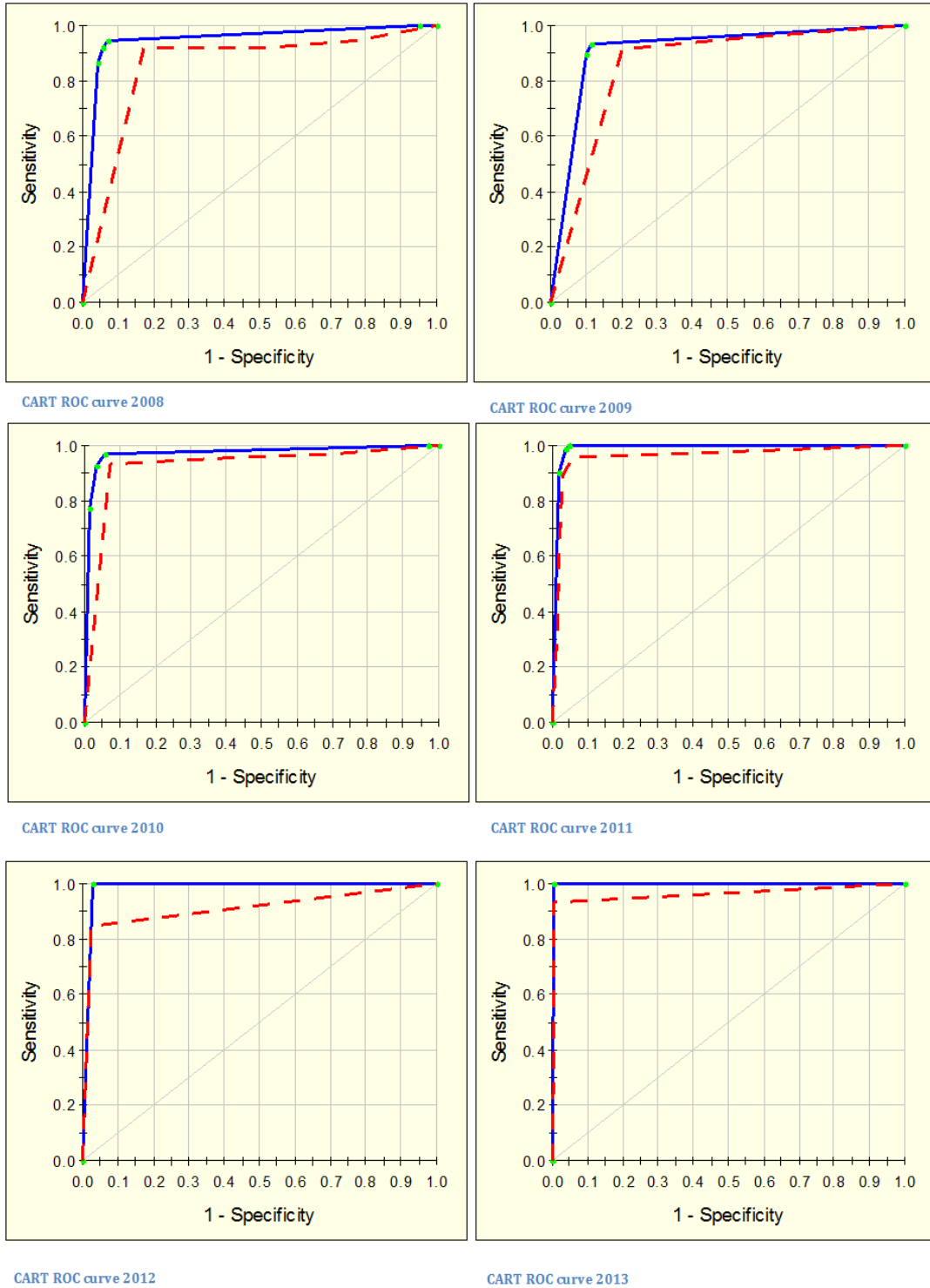


Figure 6: ROC curve K-means MARS

variables/ROCcurve K-MARS.png

