



Risk of refunding default in micro-finance institution by Bayesian networks: Case of Tunisia

Mohamed Wajdi Triki^a

Younes Boujelbene^b

^a PhD in Finance and Accounting Methods for FSEG Sfax, Tunisia.
mohamedwajditriki@yahoo.com

^b Professor, F. S. E. G. Sfax-Tunisia Younes.Boujelbene@yahoo.fr

Keywords

Microfinance, Bayesian network, credit scoring, risk of repayment.

Jel Classification

C14, G1.

Abstract

The objective of this paper is twofold: measuring credit of institution microstructure and studying Enda inter-arab Tunisia by bayesian networks. After the data gathering characterizing of the customers requiring of the micro loans, this approach consists initially with the samples collected, then the setting in works about it of various network architectures and combinations of functions of activation and training and comparison between the results got and the results of the current methods used. To address this problem we will try to create a graph that will be used to develop our credit scoring using Bayesian networks as a method. After, we will bring out the variables that affect the credit worthiness of the beneficiaries of microcredit. Therefore this article will be divided so the first part is the theoretical side of the key variables that affect the rate of reimbursement and the second part a description of the variables, the research methodology and the main results.

I- Introduction

This risk of default of repayment which are facing Microfinance institution (MFI) is all the more worrying because the micro-finance has the special feature of wanting to accomplish a social mission by remaining self-sufficient financially. This presupposes the establishment of mechanisms with the objective of reducing the risk of failure of borrowers. Currently, Microfinance institution: MFI's have developed strategies that allow you to minimize the cost of transactions and the risks related to microcredit (*Lanha, 2002; Mayoukou, 2003; Montalieu, 2002*). To grant or non-credit, certain factors are usually taken into account by the Microfinance institution (MFI) in order to minimize the risk of repayment (*Honlonkou, Acclassato and Quenum Prize, 2006*). It is a part of the factors related to the borrowers (age, marital status, residential proximity and the sector of activities including experience in the area ...) and on the other hand, those linked to the institution (object of the credit, the amount of the credit, the guarantees, and the duration of credit...) (Gool et al; 2011).

This present work has for object to put in place an algorithm of credit scoring which will be of the type **score for granting (or acceptance)**, type defined as the **score of calculated risk** for a client (new or old) who is seeking a credit.

II- Summary of theoretical literature

We focus ourselves in what follows on three principal determinants of refunding in the specific cases of an institution of microfinance: IMF (Rhee, S, G., 2008; Roy, D., 2006 and Redis., 2005): factors related to its characteristics, those related to its environment, and finally those related to the characteristics of its micro-borrowers.

The analysis of the literature having for objective to identify the causes of unpaid (Anderson, R., 2007; Honlonkou and al., 2006) shows that the insufficiency of the amounts of credit to finance the projects is a decisive cause of a bad performance of refunding. In the same way, found that the coefficient of the amount of the loans is significant and negative. This result was also confirmed by M. Labie, and M. Mees., (2005). Indeed, the negative sign is theoretically explained by the fact that the amount of the loans increases the profit associated with the moral risk. However, V. Hartarska, and D. Nasdolnyak., (2007) showed that the majority of the not refunded loans at the maturity were completely refunded a year later. In this context, the moral risk is interpreted as the choice of a project with a longer maturity than that of the loan rather than the choice of a riskier project (Bellucci, A., Borisov, A., Zazzaro, A., 2010). The negative sign relating to the amount of the loan can also be associated with the obstacles which the micro-borrower can face to refund a higher amount over a given period (usually a year) (Arminger et al ; 1997). It may be that for a given maturity, the loans of significant size do not go in par with the requirements of the borrowers and are not appropriate to the local economy (Basel Committee on Banking Supervision., 2010).

For a particular borrower and a given duration of loan, it is shown (Bhagavatula, and all., 2010; Bedecarrats, F., Angora, R.W., 2009; Lhériaü., 2005, p.23; 24) that, the probability of refunding decrease with the size of the loan. The speed of the evolution of the probability of no refunding with the size of the loan changes according to the initial equipment's of the micro-borrowers and the costs which they associate with the strategies of the moral risk and the strategic defect. Thus, the IMF (Institution of Microfinance) cannot reach a rate of perfect refunding on the basis of the several inciting mechanism of its methodology of loan (Salazar,2008). The IMF (Institution of Microfinance) will have to lay down a new objective as regards the performance of refunding (Bhatt and Tang, 2002). With an aim of not exceeding the new target threshold of defect, the IMF will grant higher loans to the slightly risky borrowers (Brennan, J. M, and W. N. Torous., 2009). The main objective of this work is to develop a statistical model that can allow distinguishing the good borrowers from bad. One of the first steps is therefore to define what we mean by good and bad borrowers.

A borrower is considered to be good if he repaid (or has always repaid) correctly its loan and has never been late in paying for thirty (30) days or more.

A bad borrower is a borrower who has experienced at least once a delay in the repayment of its loan for 30 days or more. It is worth mentioning that these definitions arising from the discussions with the credit officers and the team of the credit department of the institution.

III- Constitution of the sample and methodology of research

We are now, the implementation of bayesian networks in a practical application of credit scoring on the clients of Enda inter-Arab. Enda inter-Arab is a non-profit, micro-credit institution based in Tunisia. The mission of the institution is to contribute to the improvement of the incomes and quality of life of low-income Tunisians through a socially responsible and environmentally responsible institution designing innovative financial systems for all. Enda inter-arab offers flexible, diversified products adapted to the needs of micro-entrepreneurs, thanks to a policy of proximity and listening. Ranging from 150 to 5000 dinars, microcredit products are intended for self-employed workers, who are vulnerable in terms of access to financial and human capital, training and supervision.

We will present essentially the methodological aspects in trying to answer the two basic questions: why and how to use the Bayesian networks (Cooper and Herskovits, 1992).

Depending on the type of application, the practical use of a bayesian network can be envisaged in the same way as that of other models: neural network, expert system, decision tree, the model of data analysis (linear regression), shaft failures and logic model (Jenssen, 1997). The choice of method involves different criteria, such as the ease, cost and the delay in implementing a solution.

Outside of any theoretical consideration, the following aspects of bayesian networks make them, in many cases, preferable to other models. In fact, the Bayesian networks allow you to gather and to merge the knowledge of various kinds in the same model (Daly and Atiken, 2011).

In addition, the graphical representation of a bayesian network is explicit, intuitive and understandable by a non-specialist, which makes it easy to both the validation of the model, its any developments and especially its use. Typically, a decision maker is much more inclined to rely on a model which he includes the operation that to trust a black box.

Finally, a bayesian network is versatile: can be used of the same model to assess, predict, diagnose, or optimize the decisions, which helps to make the effort to construct the Bayesian network.

Now we are going to study in more detail the various aspects of the use of bayesian networks as a model of credit scoring.

III. 1. Overview of bayesian networks

A network is Bayesian graphical model probabilise. It is defined by:

-A graph oriented acyclic $G = (X, E)$, where $X = \{X_1; X_2; \dots; X_n\}$ is a set of variables (the nodes of the graph) and E a set of arcs. We note $\Theta = \{\theta_1; \theta_2; \dots; \theta_n\}$ the set of probability distributions such that:

$$\theta_{i} = P(X_i / Pa(X_i))$$

Or $Pa(X_i)$ is the set of nodes, connected to X_i by arcs of end X_i (the parent nodes of X_i). Then we say that $B(G; \Theta)$ is a Bayesian Network if and only if:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n \theta_i \text{ (Theorem of Bayes)}$$

This decomposition of the joint law of probabilities in a product of local terms is at the origin of the attraction generated by the Bayesian Networks. It is of the "compaction" of this act of joint probabilities that is does a number of algorithms for the calculation in a complex system probabilise. These algorithms will allow a typical use of bayesian networks: the inference.

The distributions of probabilities associated with each of the variables in the model can be either continuous or discrete. In addition, a bayesian network can both contain variables continuous and discrete. The parameters of the discrete variables can be summarized and represented by tables of probabilities conditioned to all possible combinations of the states of the variables "parent".

Each variable is a node of the graph, and takes its values in a discrete set or continuous. The graph is always directed and acyclic. The directed arcs represent a link of direct dependence (most of the time it is causation) .

Thus an arc ranging from the variable X to the Y variable will express that Y depends directly on X . The parameters express the weights given to these relationships and are the conditional

probabilities of variables knowing their parents (example: $P(Y|X)$). It is possible to achieve the classifiers thanks to bayesian networks.

Because of this, the probabilities or the scores that will appear by the following represents the scores of reimbursement or the score of creditworthiness of a borrower based on these characteristics. Subsequently, the scores obtained will classify the good borrowers from bad.

III. 2. Bayesian Networks conditional

A Bayesian network allows to model several types of nodes.

In the context of processes, we are mainly in the presence of two types of nodes: a node representing a discrete variable that is called discrete node and a node representing a continuous variable that we named continuous node.

The bayesian networks deal with parametric models. However, in the discrete case, a node multinomial allows to model all functions of probability density of a discrete variable.

In effect, a binary variable (for example True-false) can represent thanks to a discrete node (therefore multinomial) of dimension 2 (with 2 different modalities). For the continuous nodes, it is logically possible to be able to represent any functions of probability density of a continuous variable. But, at the present time, the engines of inference does not know how to process only one probability density function: that of the normal law multi-varied of dimension p .

The construction of a bayesian network is carried out in three essential steps. Each of the three steps may involve a compendium of expertise, through written questionnaires, individual interviews or even of brainstorming sessions.

III. 2.1 . Identification of the variables and their spaces of states

The first stage of construction of the Bayesian network is the only one for which the human intervention is absolutely essential. It is to determine the set of variables X_i , categorical or numeric, which characterize the system.

As in any modeling work, a compromise between the accuracy of the representation and the utility of the model must be found, by means of a discussion between the experts and the modeler.

When the variables are identified, it is then necessary to specify the space of states of each variable X_i , i.e. the set of its possible values.

The majority of the software for bayesian networks do treaty that models to discrete variables, having a finite number of possible values. If this is the case, it is imperative to discretize the ranges of variation of continuous variables.

This limitation is sometimes embarrassing in practice, because of discretisations too thin can lead to tables of probabilities of large size, of nature to saturate the memory of the computer.

III. 2.2 . Definition of the structure of the Bayesian network

The second step is to identify the links between variables, i.e. to answer the question: for what couples (i, j) the variable X_i influence the variable X_j ?

In most applications, this step is carried out by the polling of experts. In this case, the iterations are often necessary to achieve a description of consensual interactions between the variables X_i . However, experience shows that the graphical representation of the Bayesian network is in this step a media dialog extremely valuable.

A Bayesian network must not have a circuit oriented or loop. However, the number and the complexity of the dependencies identified by the experts sometimes assume that the modeling by a graph without circuit is impossible. Then it is important to keep in mind that, whatever the stochastic dependencies between the random variables discrete, there is always a representation by network baye- hers of their joint act. This theoretical result is fundamental and clearly shows the power of modeling of bayesian networks.

When you dispose of a sufficient quantity of data return of experience concerning the variables X_i , the structure of the Bayesian network can also be programed automatically by the Bayesian network, provided of course that the software uses either with adequate functionality.

III. 2.3 . Law of joint probability of variables

The last step of construction of the Bayesian network is to populate the tables of probabilities associated with the different variables.

In a first time, the knowledge of the experts concerning the laws of probability variables is incorporated into the model.

Actually, two cases arise depending on the position of a variable X_i in the Bayesian network:

"The variable X_i has no variable kinship: the experts must clarify the law of marginal probability of X_i .

"The variable X_i has variables kinship: the experts must express the dependence of X_i as a function of parent variables, either by means of conditional probabilities, either by a deterministic equation (that the software then convert in probabilities). The collection of laws of probabilities from experts is a delicate stage of the construction process of the Bayesian network. Typically, the experts are reluctant to quantify the plausibility of an event that they have never observed. However, a thorough discussion with the experts, sometimes leading to a reformulation more precise variables, allows in many cases the obtaining of qualitative assessments. Thus, when an event is clearly defined, the experts are generally better able to express if the latter is likely, unlikely, highly unlikely, etc. It is then possible to use a conversion table of qualitative judgments in terms of probabilities. The case of total absence of information concerning the law of probability of a variable X_i can be met. The pragmatic solution is then to assign to X_i a.

Law of probability arbitrary, for example a uniform law. When the construction of the Bayesian network is completed, the study of the sensitivity of the model to this bill allows you to decide whether or not to devote greater resources to the study of the variable X_i .

Almost all the commercial software for bayesian networks allows the automatic learning of tables of probabilities from data. Therefore, in a second time, any comments of the Xi can be incorporated in the model, in order to refine the probabilities introduced by the experts. Here we seek to estimate the probability distributions from available data. In the case where all the variables are observed, the method the simplest and most used is the statistical estimation which is to estimate the likelihood of an event in the database. This approach is called maximum likelihood which gives us:

$$P(X_i = x_k / Pa(X_i) = x_j) = \frac{N_{i,j,k}}{\sum_1^k N_{i,j,k}} = \frac{N_{i,j,k}}{N_{i,j}}$$

Or $N_{i,j,k}$ is the number of events in the basis of the data for which the variable X_i is in the state x_k and its parents in the configuration j .

III. 3. Constitution of the sample, data collection and definition of variables

The study relates to the whole of 19521 borrowers of Enda inter-arab which are distributed as follows, 73.125 per cent who are women and 26.875 per cent who are men.

From the exploitation of the information provided by this database for 2011, we have been able to obtain the very detailed information on: the personal and family circumstances of the recipient as well as its status, the type and the size of the activity being undertaken, the information on the credit (the date of release, its object, its amount, its duration, the amount repaid..), ability to repay, etc. ..

In considering that the conditions for the exercise of the activity within Enda inter-arab and the characteristics of the borrower may have an influence on the rate of the unpaid during the repayments of credits granted, we have constructed a graph of the study below in order to highlight the fundamental determinants of the rate of unpaid.

The dependent variable in our study is the degree of solvency of borrowers from the Microfinance institution (MFI) note (X9). In the database that we have at our disposal, it is measured in terms of number of days of delay of reimbursement. We define this dependent variable in dummy variable. It takes the value 1 if the clients of Enda inter-arab have redeemed at maturity (therefore not unpaid) and 0 if there was at least a failure occurred in the repayment of the credit (unpaid). As regards the explanatory variables, AGE refers to the age of the borrower expressed in number of years. The sixth variable refers to the marital status of the borrower (married or single). The residential proximity or geographic of the borrower (distant, near) represented by the second variable. The fifth variable refers to the activity for which the credit has been requested (small commerce, agriculture, service delivery, other). The amount of credit granted is designated by the seventh factor. The duration of credit represents the average duration of repayment of the credit.

To test the classifier on the variables chosen in order to see the performance of classification of solvency, it was obtained as accuracy of classification equal to 95.1076 per cent in 27.173762 seconds. This result shows the importance of taking account of the connections between the attributes to increase the power of classification of our model.

To avoid having a numerator equal to 0 and therefore a probability equal to 0, in the case where the number of attributes which have a certain value would be 0, it has been used to estimate the implicit and not frequentiste.

IV- Results and Discussion

We will first introduce the descriptive statistics, then, the analysis of the main results obtained.

IV. 1. The descriptive statistics

The characteristics of the sample are contained in annex 2, in the form of a table summarizing the descriptive statistics and cross-tables between the different variables and the delay in reimbursement.

As shown in table 1 , our sample has the following characteristics:

- ***The age***

The average age of our sample is 40 years, the more small borrower has 21 years old, and the eldest is 71 years old . Consultation of the table of low beam of the age with the delay of repayment allows us to see that the clients less than 30 years old and those over 60 years of age are not efficient in term of repayment at maturity.

- ***The number of monthly payments***

The monthly repayments can be the number of a single maturity, as they can achieve 36 deadlines. For this extreme case (36 monthly payments), the rate of repayment at maturity does not exceed 20 %. The average number of 8 monthly payments and 64% of the loans are for amounts between 800 and 1200 TND.

- ***The amount of the loan***

The amount of the loan granted varies between 150 and dt 5000 dt; the average amount observed is of 773 dt, the category of loans of amount less than or equal to 1500 dt recorded a perfect performance of reimbursement with a rate of 100 %, then that is observed at the level of the highest amounts, a rate of significant delay in the order of 50% and 25 %, for the appropriations for amounts higher than 2000dt.

Table 1: descriptive statistics of the sample of factors related to the institution

| | Age | Amount of credit | Maturity |
|-------------|------------|-------------------------|-----------------|
| Min | 21 | 150 | 1 |
| Max | 71 | 5000 | 36 |
| Mean | 40.65 | 773.89 | 10.88 |

- **The level of instruction**

66% Of clients of Enda inter-arab present in our database are level of secondary study, among the other customers found illiterate, as there is of higher education diplomas.

- **The sector of activity**

The borrowers in our sample operate in the trade sector with a rate of 61.81 %, while 19.19 per cent are farmers and 12.91 per cent are working in the area of services.

- **The reimbursement rate to the deadline**

The reimbursement rate to the maturity of our sample represents 86.08 per cent of the total reimbursements, the dependent variable (delay) is 0 for a total of 2745 borrowers among the 19721 borrowers in the sample.

- **The kind**

With a rate of feminization of the sample of 95.11 % very close to the rate of feminization of the mother population " Enda inter-arab ", the women represent a rate of reimbursement to the maturity higher than that of men, this rate is of the order of 76.74 per cent for the female category and of the order of 23, 26% for the male gender.

- **The family situation**

75% Of the members of our sample are married, 25% are single. While 87.67 per cent of the married repay without delay, this rate is of the order of 48% for the single.

Table 2: descriptive statistics of the sample of factors related to the borrower

| | Sex | Sector of activities | | Level of education | |
|--------------|---------|----------------------|---------|--------------------|---------|
| Man | 4.87 % | Farmer | 19.19 % | Illiterate | 9.23 % |
| Woman | 95.11 % | Trade | 61.81 % | Tray | 86.91 % |
| | | Services | 12.91 % | Top | 3.58 % |
| | | Other | 6.06 % | | |

IV. 2. Main results and discussions

To exit the uncertainty, it has built a program of assistance to the decision by using as a tool for programming the software matlab 7.1 .

It is a probabilistic model graphics to acquire, to capitalize and exploit knowledge. The bayesian networks are the successors and natural inheritors of symbolic approaches, connexionnistes and statistics of the Artificial Intelligence and Data Mining. The exploitation of our bayesian network has allowed us to find 43 configurations (c1, c2, c3 etc.) on the 1725 possible configurations which affect remarkably of probabilities $P(c/X9 = 1)$ high, as shown in the histogram below.

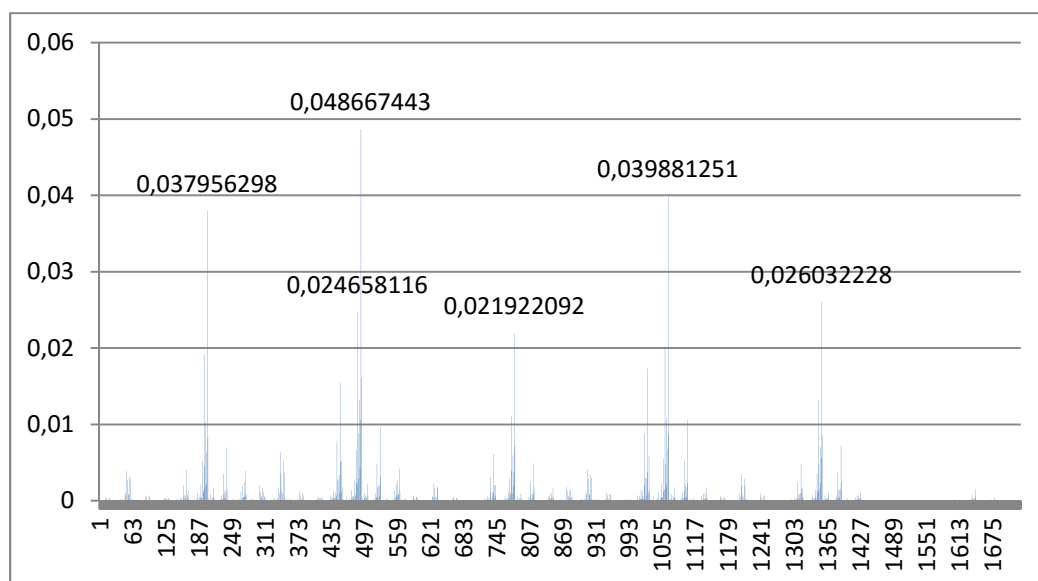


Figure 1: Histogram of probabilities

Has the reading of the table below, it is clear that the age of the borrower influence negatively and significantly the rate of failure in the reimbursements. The generally accepted view that the young borrowers are of the groups at greatest risk of repayment is confirmed by our results. This result suggests that the borrowers of middle age (30 to 45 years) have a strong propensity to repay their debts that the least elderly (less than 30 years) and the older (over 45 years) ; therefore, there is a significant relationship between age and the level of arrears. This result contradicts the results found by *Lanha (2002), Honlonkou et al. (2006) and Nawai (2012) Margaritis (2003)*.

As regards the monitoring of the credit and the proximity of the installation of the geographic area of the institution in an urban area rather than rural, it is found that the absence of visit of the borrower by the institution during the period of repayment increases the propensity of having unpaid. The duration of credit with respect to it has a negative influence and significant on the delinquency rate.

In effect, the duration of the loan, among borrowers who have unpaid debts, is greater than 12 months and those that do not have unpaid less than 12 months. Which means that more than the schedule of a loan is long, less the credit will be well repaid. This result

Confirms the hypothesis of the relaxation which stipulated that the first deadlines are very monitored and the latest deadlines least monitored leads to more of failure. These results confirm those found by *Lanha (2002) and Honlonkou and al (2006)*.

Table 3: Table of frequencies of 43 configurations Retainers

| | | | |
|--------------------------|----------------|----------------|--------|
| Duration | 0 | 1 | |
| Frequency | 58.14 % | 41.86 % | |
| Sex | 0 | 1 | |
| Frequency | 30.23 % | 69.77 % | |
| Intervention Area | 0 | 1 | |
| Frequency | 32.56 % | 67.44 % | |
| Age | 0 | 1 | 2 |
| Frequency | 13.95 % | 83.72 % | 2.33 % |

The results presented above, it is certain to say that the unpaid amounts observed in the credit portfolio of Enda inter-arab are linked not only to the characteristics of the borrower, but also to factors related to the institution.

Consider the example given above for a borrower having the following characteristics:

Table 4.4 : Example of a client of the data base of Enda-inter-arab

| Age | Intervention Area | Sex | Level of education | Sector of activities | Marital Status | Amount of credit | Duration of credit | P(c/X9 = 1) |
|----------|-------------------|-------|--------------------|----------------------|----------------|------------------|--------------------|-------------|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0.05 |
| 30 <X<50 | Urban | Woman | Bin Level | Trade | Married | 500 <X<1500 | <To12 months | *** |

The probability of solvency ($PX9 = 1/c$) of this individual is equal to 79.326 %. In other words it has calculated the probability to be solvent knowing the characteristics "c" of the individual.

In considering all other things being equal, we can conclude that a borrower (wife) married who belongs to an interval of age between 30 and 50 years, close to the

Place of installation of the office of the IMF, having a level of instruction tray or more, and who request a microcredit not exceeding the 1500 dt reimbursed in a period not exceeding 12 months , will have a greater chance to get of this borrowing and it does not present a risk of non-payment of the loan to the IMF. Its request for granting of credit will be directed to the commission or the experts of the FMF automatically to validate or not for that this client enjoys a microcredit.

Because of this, our model is both useful for the IMF (Institution of Microfinance) to help them classify its customers into 2 groups, the first group who represent the good borrowers and the other which at risk of non-repayment. Our bayesian network will aim to develop a model of Credit Scoring allowing the IMF (Institution of Microfinance) to predict the probability of default of new credit applicants.

The granting of credit by an Microfinance institution (MFI) presents risks whose main remains the default of counterparty which is very often the result of unpaid which indicate a future failure of the client.

The outstanding debt is a failure of the debtor which is located in the inability to settle its debt within the time limits or who lack to its obligations such as mentioned in the contract of loan. The unpaid amounts are also the deferral or the loss of the products of interest.

The applications of bayesian networks in the field of banking and finance are still rare, or at least are not published. But this technology has a very important potential for a number of applications falling within this area, as the financial analysis, the scoring, the evaluation of the risk or the detection of fraud.

In the first place, the Bayesian networks offer a unified formalism for the manipulation of the uncertainty, in other words the risk, including the taking into account is essential when it comes to financial decision.

Then, the possibility of pair expertise and learning is very important here, not only because the two sources of knowledge are in general available in this area, but also and especially because this capacity can help to respond to the problem of structural changes of environment.

The use of bayesian models is particularly suited for several reasons:

- The bayesian networks have enabled us to pair the knowledge of the experts and the data available.
- They allow conditioning the risks and therefore to better assess the losses incurred.
- They allow you to identify the levers for reducing risk.
- These models are transparent and easily auditable by the inspection bodies.
- The bayesian model built is a program of assistance to the decision to double role. First, it is used to differentiate between the good borrowers and bad borrowers. Second, once transformed into the decision support system, it will be a strong model for the new credit applicants to know their chance of benefit from microcredit Enda inter-arab following the instructions of the created system.

Through the establishment of a system of management of risk of credits based on the method of credit scoring which should be refined, discussed and improved by the opinion of the agents of appropriations and of the governing bodies of the institution, Enda inter-arab can only improve its performance in terms of repayment at maturity and efficiency and which constitute valuable assets to start the new phase of restructuring to which she is preparing.

In this context, Enda Inter-arab, to continue to advance and further improve its remarkable performance, and from there consolidate the position of choice that it occupies in the mechanism to fight poverty in Tunisia, is called upon to undertake its transformation into financial institution

and make sure to prepare the instruments, mechanisms and procedures adequate to succeed this mutation.

References:

- Andrea Bellucci , Alexander Borisov and Alberto Zazzaro «Does gender matter in bank-firm relationships? Evidence from small business lending » *Journal of Banking & Finance*, 2010, vol. 34, issue 12, pages 2968-2984
- Arminger, G., Enache, D. & Bonne, T. (1997). « Analyzing Credit Risk Data: A Comparison of Logistic Discrimination, Classification Tree Analysis, and Feed-forward Network », *Computational Statistics*, Vol. 12, Issue 2, pp. 293-310.
- BÉDÉCARRATS F., ANGORA R.W., LAPENU C. (2009) Is Social Performance Profitable ? The relationship between social and financial performance in microfinance, *MicroBanking Bulletin*, n° 19, 22-29.
- Bhatt, N. and S. Y. Tang (2002), « Determinants of Repayment in Microcredit: Evidence from Programs in the United States », *International Journal of Urban and Regional Research*, Vol. 26, No. 6, pp. 360-76.
- Boyle, M., Crook, J., Hamilton, R. and Thomas, L, Methods for credit scoring applied to slow payers. in L. Thomas, J. Crook and D. Edelman (eds), *Credit Scoring and Credit Control*. Oxford: Oxford University Press, 1992, 75-90.
- Brigham, E. F. (1992). *Fundamentals of financial management* (6th ed.). Forth Worth: Dryden Scan. J. Stat. 24, 1-13.
- Castillo E., Gutierrez J.M., Hadi A.S (1997). *Expert Systems and Probabilistic Network Models*. Springer, New York, 605 pp.
- Cooper and Herskovits, (1992) G.F. Cooper and E. Herskovits. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, 309-347.
- F.V. Jensen (1996). *An introduction to Bayesian Networks*. Taylor and Francis, London, United Kingdom.
- G.F. Cooper, (1990) The computational complexity of probabilistic inference using Bayesian Belief Networks, *Artificial Intelligence*, Vol. 42 (23), pp. 393- 405.
- D. Heckerman, D. Geiger, and D. Chickering (1995). Learning bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, Vol. 20, pp. 194-243.
- Daly, R., Shen Q., Aitken S (2011), *Learning Bayesian Networks: Approaches and Issues*. The Knowledge Engineering Review, Vol.26 (2), pp. 99-157
- Davis D.B, *Artificial Intelligence goes to work* . *High Technol*, (1987), Apr 16-17.
- Deschaine L., FRANCONNE F, Comparison of Discipulus TM Linear Genetic Programming Software with Support Vector Machines, Classification Trees, Neural Networks and Human Experts, White Paper. Available at: <http://www.rmltech.com/> (Accessed: 10 June 2008).

- Dinh, T.H.T. and S. Kleimeier, Dinh(2007). A credit scoring model for Vietnam's retail banking market. *International Review of Financial Analysis*, Vol.16(5), pp. 471-495.
- D. Margaritis, Learning Bayesian network model structure from data, (2003) (PhD Thesis of CMU-CS-03-153).
- D. Koller, N. Friedman, Probabilistic Graphical Models: Principles and Techniques, The MIT Press, Cambridge,MA/London, England, (2010).
- E. Castillo, J.M. Gutierrez, A.S. Hadi, Expert Systems and Probabilistic Network Models, Springer-Verlag, (1997).
- Fan Liu, Zhongsheng Hua, Andrew Lim. (2015), Identifying future defaulters: A hierarchical Bayesian method, *European Journal of Operational Research*, Vol. 241, pp. 202-211.
- Fatemeh Jamaloo, et al. "Discriminative CSP Sub-Bands Weighting Based On DSLVQ Method In Motor Imagery Based BCI" *International Journal of Online Engineering (IJOE)*, 5 (3), (2015), 156-161.
- G. Claeskens, N.L. Hjort, Model Selection and Model Averaging, Cambridge University Press, Cambridge, (2008).
- Hardy W. E., Jr., Adrian J. L. Jr. (1985), A linear programming alternative to discriminant analysis in credit scoring. *Agribusiness*, Vol. 1(4), pp. 285-292.
- Hartarska et Nadolnyak (2008) «An Impact Analysis of Microfinance in Bosnia and Herzegovina» *World Development Vol.21, No.1*
- Honlonkou, A. N., Acclassato, D. H. and Quenum, C. V. C. (2006), DÉterminants de la Performance de Remboursement Dans Les Institutions de MicroFinance au BÉNin. *Annals of Public and Cooperative Economics*, 77: 53–81.
- S. Hojsgaard, D. Edwards, S. Lauritzen, Graphical Models with R, Springer, New York, 2012.
- J. Cheng, R. Greiner, J. Kelly, D. Bell, W. Liu (2002), Learning Bayesian networks from data: an information-theory based approach, *Artificial Intelligence*, Vol.137, pp.43-90.
- J. Pearl, Probabilistic Reasoning in Intelligent Systems, Networks of Plausible Inference, Morgan Kaufmann, (1988).
- Johnson R. W., Kallberg J. G., Management of accounts receivable and payable. New York: Wiley,(1988).
- Korb K. B., Nicholson A., Bayesian artificial intelligence. Chapman, Hall/CRC, Boca Raton, FL, 2nd edition, (2010).
- Labie M., (2004) « Micro-finance : un état des lieux », *Mondes en développement, Vol. 32, 2, n°126, 2004, p.9-23.*
- Lanha Magloire, « Résolution des problèmes d'information en micro-finance Analyse à partir de la stratégie de Vital-Finance Bénin », *Mondes en développement, 3/2002 (n° 119)*, p. 47-62.

- L.E. Sucar and M. Martinez-Arroyo (1998), Interactive structural learning of bayesian networks. Expert Systems with Applications, Vol.15, pp. 325-332.
- Lobna A, A_f M, Sonia Z-G (2016), The Consumer Loans Payment Default Predictive Model : An Application in a Tunisian Commercial Bank, Asian Economic and Financial Review, Vol. 6 (1), pp.27-42.
- M. Henrion, Propagating uncertainty in Bayesian Networks by probabilistic logic sampling, (1988) Proceedings of the Fourth Conference on Uncertainty in Artificial Intelligence, pp. 149-163.
- Mouley S., Réformes et restructuration du système bancaire et financier en Tunisie : Quelle vision et quel plan stratégique prioritaire? In Cahier du Cercle des Economistes de Tunisie, (2014) Number 4.
- N. Friedman, D. Koller, Being Bayesian about network structure. A Bayesian approach to structure discovery in Bayesian Networks (2003), Machine Learning, Vol. 50, pp. 95-125.
- R.E. Neapolitan, Learning Bayesian Networks, Prentice Hall, Inc., Upper Saddle River, NJ, USA, 2003.
- Oreski, S., Oreski, G (2014), Genetic algorithm-based heuristic for feature selection in credit risk assessment. Expert Systems with Applications, Vol.41(4), 2052-2064.
- Pitt, M.M. & Khandker, S. R. (1998), The impact of group-based credit programs on the poor households in Bangladesh: does the gender of participants matter? Journal of Political Economy, Vol.106 (5), pp.958-996.
- Roslan, A. H. and A. K. Mohd Zaini (2009), « Determinants of Microcredit Repayment in Malaysia: The Case of Agrobank », Humanity and Social Sciences Journal, Vol.4(1), pp.45-52.
- Salazar, G. L. (2008), « An Analysis of Repayment among Clients of the Microfinance Institution Esperanza International, Dominican Republic », Journal of Agricultural Economics, Vol. 90 (5), pp.1366-1391.
- Schreiner, M., (2004), « Benefits and pitfalls of statistical credit scoring for microfinance ». Savings and Development, Vol.28(1), pp. 63-86.
- Tenenhaus M. (2000), La Régression Logistique PLS, Journées d'Etudes en Statistique, Modèles Statistiques pour données Qualitatives, 261-273.
- Thierry Montalieu «les institutions de micro crédit ; quelles pratiques bancaires pour quels effets»
- Thomas. L, A Survey of Credit and Behavioural Scoring; Forecasting financial- risk of lending to consumers (2000), International Journal of Forecasting, Vol.16(2), pp.149 -172.
- Vaclav K. (2015), Genetic algorithms for credit scoring : Alternative fitness function performance comparison, Expert Systems with Applications, Vol. 42, pp. 2998-3004.
- Van Gool, J., Verbeke, W., Sercu, P., Baesens, B. (2011), « Credit scoring for microfinance: is it worth it? » International Journal of Finance & Economics, Vol.17(2), pp.103-123.