



“SAS software made it possible to build the fundamentals for a powerful, usable, and comprehensible credit risk solution.”

Bart Baesens

PhD candidate in Applied Economic Sciences at the Katholieke Universiteit Leuven

The Power to Know®...

Katholieke Universiteit Leuven
Credit-risk solutions based on neural networks



“How sure are we that our customers will be able to repay their debt on time?” Financial institutions ask themselves this question hundreds of times a day. The development of credit risk analysis methods that allow a manager to decide whether or not to grant a loan is a very complex problem. Bart Baesens, PhD candidate in the data mining group headed by Professor Jan Vanthienen of the faculty of Applied Economic Sciences at the K.U.Leuven, has developed a model based on neural networks that is both powerful and comprehensible. His first choice for software to build, train, and evaluate these networks was SAS®. Baesens relies heavily on its high number of advanced training algorithms and its easy-to-use graphical user interface. The result is absolutely fascinating and has opened the door for many more applications in the future.

Difficult decisions with high impact

Credit risk analysis is not an easy job. A single wrong decision can cause a great deal of money to be lost just as a right one can build fortunes. The person making these decisions needs to be able to understand and explain the reasoning behind them. Needless to say, any tool that can correctly evaluate the risk and simplify the decision process will be in great demand around the financial world. Every financial institution is looking for reliable and easy-to-use decision support systems. Advanced, but sometimes very complex statistical models are currently used to decrease the number of wrong decisions. Unfortunately, these models are rarely user friendly and can only be understood by specialists.

“An intelligent credit scoring system has to be both accurate and comprehensible”, points out Baesens. And he has developed an easy-to-use credit risk decision table based on real-life data provided by a major Benelux financial institution.

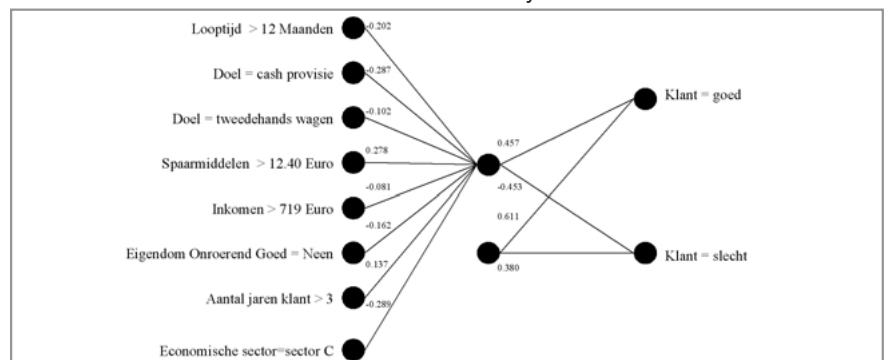
SAS software for building powerful neural networks

SAS® Enterprise Miner™ and SAS/STAT® provide advanced support for

developing high-performance credit scoring systems. SAS software includes a high number of advanced training algorithms for neural networks (see textbox for more information on neural networks). In SAS Enterprise Miner, the information is usually visualized as a flow chart, making it attractive and easy-to-use. Advanced accuracy measurement techniques are available too. “Its inherent ability to model complex algorithms, together with its friendly graphical user interface, are the main reasons SAS was chosen for credit risk analysis”, recalls Baesens. “SAS simplifies a very powerful, but complex algorithm like a neural network in such a way that very strong applications can be built with it.”

The result: useful and comprehensible rules

The final goal of this investigation is a comprehensible rule set instead of a complex mathematical model. This set is extracted from the neural network model by using rule extraction or tree extraction techniques. “The complexity of a mathematical model is far too high to allow frequent, efficient use”, Baesens declares. “On the other hand, a compact and clear decision table will facilitate both fast and correct decisions. It is also a tool for the credit manager to explain the decision to the customer easily.”



A complex mathematical model

1. Savings Account	<= 12.40 Euro												> 12.40 Euro				
2. Economical sector	Sector C	other											-				
3. Purpose	-	cash provisioning						second-hand car				other	-				
4. Term	-	<= 12 months				> 12 months				-		-	-				
5. Years client	-	<= 3		> 3		<= 3		> 3		<= 3		> 3		-	-		
6. Owns Property	-	Yes	No		-	-	Yes	No		Yes	No		-	-	-		
7. Income	-	-	<= 719 Euro		> 719 Euro		-	-	-	-	<= 719 Euro		> 719 Euro		-	-	-
1. applicant = good	-	x	x	-	x	-	x	-	x	x	-	x	x	x			
2. applicant = bad	x	-	-	x	-	x	-	x	-	-	x	-	-	-			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14			

A compact and comprehensible credit scoring decision table allows fast, correct credit-risk analysis

Credit scoring survival analysis: fundamentals for additional decisions

For the credit grantor, it's not only about the decision whether to grant the loan. It's also quite important to know at which point in time the customer can be expected to fail to reimburse. This is called survival analysis and is a crucial additional factor in the final decision. This analysis facilitates setting such parameters as the loan duration and maximum credit limit. It makes debt provisioning and repayment behavior monitoring possible.

According to Baesens, SAS/STAT® belongs to the most powerful software available for survival analysis. But because neural networks are not normally equipped for survival analysis, he is planning more research before publishing his findings in this area.

Bart Baesens' investigation has initiated a variety of potential future applications. Company bankruptcy prediction, the calculation of customer lifetime value, and customer retention applications are just a few examples. Fraud detection, a problem banks regularly struggle with, is another interesting application area. But none of this would have been possible without the curiosity of a gifted investigator untangling neural networks and equipped with the right tools from the very start.

Neural networks

A neural network is a complex mathematical model inspired by the functioning of the human brain. It's a useful and powerful tool for analyzing huge amounts of data and for learning to find patterns in data. These networks are particularly well suited for developing accurate credit scoring systems, especially when compared to other classification techniques, such as logistic regression, discriminant analysis, decision trees and rules, and other exotic mathematical methods.

Before training a neural network, the investigator must preprocess the data sets. Only clean data sets are useful: missing data, discrepancies and inconsistencies will decrease the quality of the neural network. After preprocessing, the investigator can then train the network by using so-called training algorithms. Finally, its performance is calculated, using the appropriate measures to reflect the system's accuracy. If the accuracy is poor or in some way deficient, then it follows that many wrong decisions will result. The credit scoring system will be considered incompetent and thus of no use as a useful tool in the financial arsenal.

Several neural network architectures are available for SAS® Enterprise Miner™. More information can be found at www.sas.com.



SAS Institute
Kasteel de Robiano
Hertenbergstraat 6
3080 Tervuren
Belgium

SAS Institute s.à.r.l.
204 route d'Arlon
L-8010 Strassen
Luxembourg

Tel.: +32(2) 766 07 00
Fax: +32(2) 766 07 77

Tel.: +35(2) 26 11 84
Fax: +35(2) 26 31 11 84

www.sas.com/belux