Credit Risk Early Warning System using Fuzzy Expert Systems

Igor Kaluđer
Sunoptos d.o.o.
Črešnjevec 62, 10000 Zagreb, Croatia
igor.kaludjer@sunoptos.com

Goran Klepac
University College for
Applied Computer Engineering Algebra
Ilica 242, 10000 Zagreb, Croatia
goran@goranklepac.com

Abstract. *We suggest an early warning system (EWS) for credit risk management based on fuzzy expert systems. Modelling process, including knowledge elicitation and univariate analysis is proposed as well as several approaches to handling model complexity, model auto tuning and validation.*
We found that such a hybrid model had more predictive power and is more robust than a purely statistical one. Even more important is that the methodological choice facilitated the discussion between stakeholders and increased involvement, acceptance and interdepartmental communication.

Keywords. Credit, risk, fuzzy, logic, expert, system, early, warning

1 Introduction

Recent financial crisis revealed that many financial institutions (FIs) failed to recognize increases in credit risk in their portfolios early enough to take appropriate action. Main reasons for this include: too much focus on underwriting and compliance, lack of dedicated organizational unit (often called monitoring or EWS unit) and personnel, lack of interdepartmental and intragroup communication, poor data quality and inadequate IT support.

An early warning system is essential to any financial institution exposed to credit risk. It allows the FI to recognize signs of increase in credit risk (early warning signals) early enough and differentiate between clients whose default can be prevented by taking appropriate actions from those where a more aggressive strategy is optimal (e.g. liquidating the collaterals and exiting the business relationship). In that sense, EWS helps to minimize the losses related to credit risk. EWS also provides a competitive advantage over other creditors: first creditor to identify a severe increase in credit risk and start collecting early will collect more than others. Another benefit is that an EWS allows the FI to be proactive by alerting and advising the client to take action sooner rather than later and avoid delinquencies and default.

We propose a methodology based on experience we got from implementing early warning systems in several large financial institutions. While there is a fair amount of research done and published about systemic risk early warning systems (those predicting a system wide financial crisis) [1] [2] [3] [4] , same cannot be said for EWS at an individual institution level, which is what we are concerned with [5] [6]. An obvious reason for this is that protect their intellectual property and seldom allow for publication of any details. Another reason might be that this part of the credit risk management process has been largely neglected by the ever increasing regulation keeping the FIs occupied with capital adequacy, risk parameter estimation, provisioning methodology, etc.

2 Problem definition

In its most simple form, an EWS can function on a single signal level, alerting the users whenever one or more signals go over the threshold. Such a system, while easy to develop and implement, usually does not meet the requirements in respect to predictive power and/or burdens the users with too many false positives.

In the more sophisticated version, we are dealing with a classification problem. The scope of the system are performing clients with a significant credit balance. The goal is to assign each client to a class corresponding with an optimal strategy for that client at a given point in time. A strategy is a collection of actions usually laid out on a timetable based on days past due (e.g. email the client at 10 DPD, phone call at 15 DPD, freeze credit limits at 30 DPD, etc.). An example is given in Figure 1 with 5 classes, ranging from a class for clients that show no signs of credit risk to a class for the most severe cases which should be forwarded to the workout department immediately.

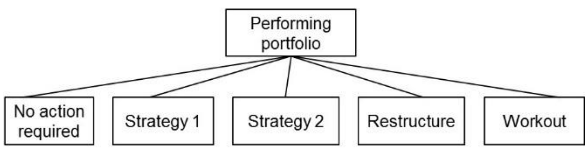

Figure 1. Classification tree

As can be seen in Table 1, there are two types of errors. False negatives are cases where EWS predicted that clients will not default and therefore no action is required, but they actually do default. This error reduces efficiency of the system. The second error, false positives, arises when the EWS predicts an increase in credit risk, but the client does not actually default. This error reduces effectiveness. In order to construct the confusion matrix all strategies expect "No action required" are grouped into one "Action required" class. This approach can also be applied in model validation as it simplifies some of the complexity involved with manual overrides (e.g. risk manager can override the model output and select a different strategy). More details are provided in the validation section.

	Outcome	
Classification	Performing	Default
No action required	True negative	False negative
Action required	False positives	True positives

Table 1 - Confusion matrix

The goal of the system is to minimize errors, while available resources are the constraint that must be met. For example, the system should not classify more clients in the "restructure" class than the restructuring unit can manage.

As the essential goal of the EWS is to estimate credit risk, a comparison to credit rating models and risk parameter estimation model should be made. The main difference between EWS and credit rating models is that false positives have much bigger impact in rating than in EWS. By overestimating credit risk in its rating model, FI will turn down viable transactions and potentially loose some of its good clients, misprice transactions, etc. False positives in EWS however, only place an unnecessary burden on the staff. Hence, EWS can focus more on early recognition rather than accuracy. Opposed to probability of default (PD) estimation models, EWS are concerned with a longer outcome window than 1 year which is a regulatory requirement for PD.

3 Fuzzy Expert Systems

3.1 Expert systems

An expert system is a computer system that emulates human decision making [7]. Every expert system is composed of two main parts: knowledge base and an inference engine. Knowledge base contains the available factual and heuristic knowledge about the domain subject. Knowledge is formalized in the system by some form of knowledge representations. The most widely used representation in expert systems are production rules, or simply, rules. Rules have two parts: conditions and actions. For example:

IF interest coverage ratio < 1 AND rating < BB THEN choose strategy 2.

There can be an infinite number of rules in the expert system, although it's usually kept under control for maintenance reasons. Any rule can have an arbitrary long set of conditions and actions. Again, it's advisable to keep the rules as simple as possible so they can be easily interpreted.

The inference engine operates on knowledge base in order to reach the goal of the problem. The inference (or reasoning) engine chooses which rules to fire, makes an assessment of the outputs and chooses which rules to fire next, and so on, until it reaches the goal. This approach is called forward chaining, the opposite being backward chaining in which the inference engine starts from the goal and work its way backwards.

3.2 Fuzzy logic

Unlike traditional (crisp) logic in which a statement is either true or false, in fuzzy logic a variable can have a truth value anywhere between 0 and 1. Consider an example with interest coverage as a financial indicator (Figure 2).

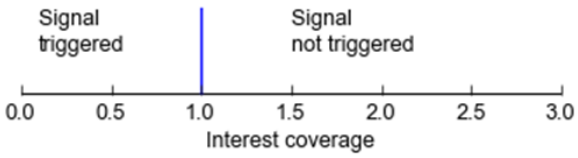

Figure 2 - Crisp logic example

In traditional, crisp logic we have to define a threshold (in this case equals 1) below which the signal will be triggered. So for clients with interest coverage just slightly above the threshold the signal will not trigger. Usually this is not in line with expert reasoning and business logic. In fuzzy logic we can define linguistic variables for interest coverage.

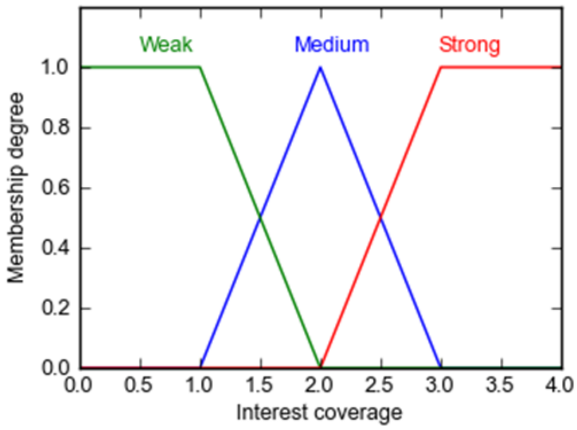

Figure 3 - Fuzzy logic example

There are three linguistic variables in this example: weak, medium and strong interest coverage. Each linguistic variable is defined with a membership function which determines the membership degree for all original, numerical values. Although somewhat similar, membership degrees are not to be confused with probability. Values of interest coverage less than 1 have 100% membership degree for the weak (i.e. they are 100% weak). As we move along from 1 to 2, the membership function declines linearly and values larger than 2 are no longer considered to be weak at all (i.e. their membership degree for weak is 0%). Values in between 1 and 2 are members of both, weak and medium, fuzzy sets (e.g. value 1.5 is equally weak and medium – membership degree to both is 50%).

3.2 Fuzzy expert systems

Fuzzy expert system is an expert system that, instead of Boolean logic, uses fuzzy logic. A typical rule in a fuzzy expert system is:

IF interest coverage is weak AND client rating is weak THEN choose strategy 2.

The main difference between a fuzzy and the ordinary expert system is the use of linguistic variables in rules definition.

4 Modelling process

After setting the goal and the scope of the system the model development process starts by interviewing the domain experts, in this case business relationship managers, risk analysts and managers, etc. As the methodology is heavily dependent on domain expert knowledge, this phase of the development process (knowledge elicitation) is of critical importance to successful model development, implementation and use. All the stakeholders and domain experts from different organizational units should be involved (business side, risk management, IT, legal, etc.).

4.1 Early warning signals

In the first step a comprehensive list of data sources for early warning signals is made. A non-exhaustive list includes:

- Internal data
- Group data (leasing, insurance, factoring, ...)
- Financial statements
- Macroeconomic and industry analyses
- Credit bureau
- Capital markets
- Government databases (land registry, subsidies, official papers, ...)
- Media
- Payment transactions

Individual signals are defined as well as their scope (e.g. some signals are applicable only to large corporate clients). Care must be taken to constrain the number of signals, otherwise the system will be too complex to interpret and maintain. Usually the number of signals is between 15 and 100.

Signals are then fuzzified that is they are transformed from a numerical to a linguistic variable. Not all signals need to be fuzzified; for some of them crisp logic works better and they can be included in rules as numerical or categorical variables.

Once signals have been defined and fuzzified their quality is checked by univariate analysis against historical data. There are three dimensions to signal quality: accuracy, time to default and workload. Clients are sampled on the observation date (only performing clients) and outcome variable is defined as default event occurrence during the outcome window. Clients for which default has not occurred are referred to as good, while those that defaulted are referred to as bad. Accuracy is then assessed using the usual measures for a binary outcome: weight of evidence (WoE) and information value (IV).

$$WoE = \ln\left(\frac{Relative\ Frequency\ of\ Goods}{Relative\ Frequency\ of\ Bads}\right) \quad (1)$$

$$IV = \sum[(Relative\ Frequency\ of\ Goods_i - Relative\ Frequency\ of\ Bads_i) * WoE] \quad (2)$$

where *i* is i-th group of a given variable. Alternatively, accuracy can also be assessed using the gini coefficient.

Time to default is average time between the first occurrence of the signal and the actual default. Signals where time to default is smaller are preferred. There are early warning systems which are almost solely based on financial indicators. While they suffice in terms of accuracy they usually fell short when it comes to time to default because they can only be raised once the financial statements are prepared and delivered to FIs which is often too late.

Workload is defined as the number of clients for which the signal has been raised. Staff that will operationally manage these clients has limited time and one of the main goals of EWS is to help them focus their attention to clients where timely action can make a difference.

The output of this phase is a long list of signals with all its linguistic variables and for each variable a measure for accuracy, time to default and workload. Based on these three measures of signal quality a selection is made by the domain experts and model developers.

4.2 Rules

A list of auxiliary variables is made (e.g. customer segment, industry, geographic region). These variables are not early warning signals in themselves but rather serve to increase the accuracy of a given rule. For example, a certain signal may only be valid or accurate enough to be used only within a certain segment of clients or within a particular industry.

Signals (linguistic variables) and auxiliary variables are combined into rules. For example:

IF rating downgrade is large AND rating is bad THEN choose strategy 2.

All rules for which at least one signal is triggered are fired simultaneously.

Rules can have weights making them more or less important in the overall model. Rule weights also serve as a basis for simple auto tuning of the model.
After rules have been defined the univariate analysis described for signals takes place only on individual rule level. The process of signal and rule definition, fuzzification and univariate analysis is iterative – it repeats until a satisfactory solution has been found and accepted by model developers, domain experts and end users.

5 Model

5.1 Model architecture

Model architecture is depicted in Figure 4. System collects necessary data and feeds it into the model scoring engine. Signal values are calculated and fuzzified. Next, support for each rule is calculated using standard fuzzy logic operators (AND, OR, NOT). Defuzzification can be done using standard methods [8], but if we want our system to react to even a single signal alert we need an alternative approach. One simple and effective way to accomplish this is that for each class we calculate membership degree as the maximum support of rules which result in this particular class. This way system will be very sensitive because only rules with maximum support for each class are taken into account.

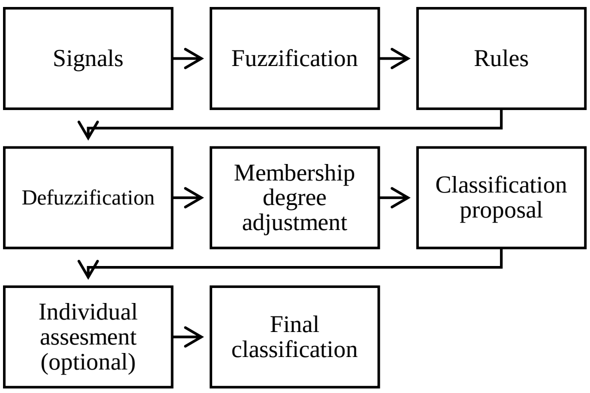

Figure 4 - Model architecture

After defuzzification every class has a corresponding membership degree. In order to reduce complexity we can directly manipulate the membership degrees. More details are given in the following sub chapter. Each class now has a membership degree assigned and the model can propose a classification. We do this by assigning a threshold (Figure 5). One of the classes should be designated as default in case a particular client does not reach any of the thresholds. If a client exceeds the threshold of two or more classes, a resolution logic must be defined. One simple logic is that most conservative class wins. For example, if the client exceeds thresholds in “no action required” and “close monitoring required” classes, the latter will be the class proposed by the model.

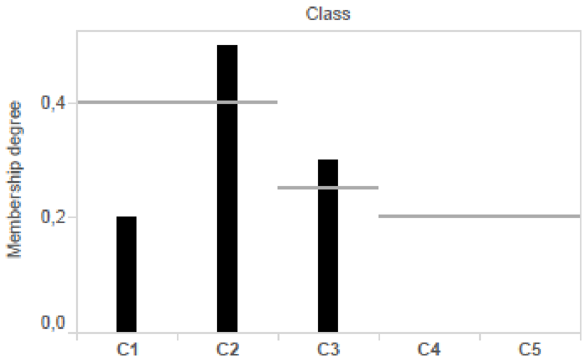

Figure 5 - Classification resolution

Thresholds must be easily changeable in order to tune the system and align it to available resources.

An optional step, usually done for large clients, typically corporates, institutions and sovereigns, is that EWS only proposes a classification where clients in classes other than “no action required” are individually and manually assessed. Final classification is then made by expert judgement.

5.2 Handling complexity

A common error in fuzzy expert systems development is having too many indicators and rules. Such a system is hard to interpret, implement and very hard to maintain. One approach to handling complexity of the model is rule blocks (Figure 6). Several signals about rating downgrade and default of parties connected to a particular client (owners, key customers, etc.) are merged into a general one describing all connected customers. One rule can then be defined using “connected customers” rather than having six different rules.

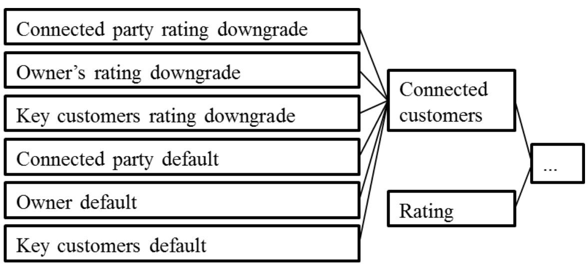

Figure 6 - Rule blocks

Another approach is to segment the entire model. For example, separate sub models can be made for different customer segments: retail, SMEs, corporate, etc. This effectively increases the total number of rules, but provides a clear delineation of separate sub models.

Finally, membership degrees can be altered directly. For example, clients with exposure fully covered with high quality collateral (e.g. cash deposit) should be treated less conservatively than others. One option is to include this differentiation into every rule, but this significantly increases the complexity of the model. Another is to segment the model into two submodels: one for clients fully covered with high quality collateral and the other for the rest. This also complicates the model. By directly adjusting the membership degrees of classes we can avoid overcomplicating the model and still get the differentiation required. For example we can lower the membership degree of fully covered clients by certain absolute or relative amount for particular classes.

5.3 Auto tuning

Because rules have weights (rules are fuzzy as well), the system can be auto tuned by adjusting the rule weights so that they correspond to a measure of predictive power. The system continuously monitors the accuracy of every rule and adjust the weight correspondingly. Weights for rules which are declining in accuracy are automatically decreased and vice versa.

The adjustment can be done with a simple function connecting the accuracy measure (e.g. information value or gini coefficient) to rule weights or by a more sophisticated approach using Bayesian statistics or neural networks.

5.4 Validation

Validation of expert based early warning systems is particularly difficult due to long data history required and several other specifics compared to statistical models. The system can be validated using the same approach applied in signal and rule univariate analysis complemented with ROC curve and related measures, but there are several problems with back testing early warning systems.

The main issue is how to define the outcome variable for the historical period. One cannot just use the default clients as “bad” because the “good” ones might be good because early signs of credit risk were identified and actions taken that prevented default. So in those cases the EWS would be unjustly punished in terms of accuracy. One way out of this problem is to define a set of actions and count clients with those actions as “bad”. But then again is the question whether these clients recovery is due to these actions or not.

Besides the usual, statistical approach to model validation, backtesting the model on particular historical cases is useful. A number of clients which already defaulted or were saved from default by taking appropriate and timely action are selected as test cases. Historical data for these clients is prepared and the system is run at several points in time prior to default (e.g. 3, 6, 12 and 24 months prior to default). System output is discussed with domain experts. If the system successfully classified these clients early enough and the domain experts have no major remarks, the model can be considered valid. An example of historical simulation is given in Table 2.

Test case	**Model outcome at N-th month prior to default**				**Domain experts' remarks**
	3m	**6m**	**12m**	**24m**	
Case 1	R	R	S1	S1	Correct & timely
Case 2	S1	NA	NA	NA	Too late
Etc.					

Table 2 - An example of historical simulation, where S1 = Strategy 1 (action required), R = Restructuring and NA = No Action Required

6 Results

We found the fuzzy expert system to have more predictive power than a pure statistical model, especially for outcome windows longer than 1 year. They are also not prone to usual problems of statistical models: historical data availability and quality, sample size, bias, etc. By exploiting expert knowledge the system is also more robust (not prone to overfitting).

The modelling process has several side benefits which proved to be beneficial to successful model implementation and use. First of all, it includes all the stakeholders and domain experts in one place and facilitates communication between experts which otherwise seldom takes place. Rules are defined in natural language which also makes the communication far easier than logistic regression models, for example.

Another major benefit of fuzzy expert systems is interpretability. Early warning systems are especially required to explain themselves so the end user can decided on appropriate action. It is absolutely essential that the EWS does not function as a black box, but rather explain the decision by showing which rules and signals are responsible for the outcome. All of this made user acceptance much easier compared to statistical and data mining models.

On the down side, there are major issues with model validation for which there is no industry standard practice yet. Also the development process is rather long and has to involve a large number of stakeholders making it difficult to reach a consensus. Finally, unless developed in specialized software (e.g. Fuzzytech) which can produce production ready software components, implementation can be challenging for IT departments.

7 Conclusion and future research

Early warning system built using fuzzy expert systems proved to be a solid alternative to pure statistical models. By leveraging expert knowledge, ability to act upon a single signal (variable) and user involvement it provides several major advantages in terms of predictive power, robustness, interpretability, transparency of assumptions, etc.

There is very little research published in this area. Most of the available publications deal with early warning systems from a macro-economic, system wide perspective [3] [1] [4], while those few that take the micro perspective are data driven, statistical or data mining models [2] [5].

Validation continues to be an issue and more research is required. Promising field of future research is the auto tuning of the model. Especially using Bayesian techniques and neural networks. A major area still largely unexplored are payment transactions in financial institutions. Finally, social network analysis seems a perfect addition to early warning system, but very little research has been published on the subject so far.

References

[1] Davis, P. E.; Karim, D. Comparing early warning systems for banking crises. *Journal of Financial Stability*, vol. 4, no. 2, pp. 89-120, 2008.

[2] Koyuncugil, A. S.; Ozgulbas, N. Financial early warning system model and data mining application for risk detection. *Expert Systems with Applications*, vol. 39, no. 6, pp. 6238-6253, 2012.

[3] Bussiere, M.; Fratzscher, M. Towards a new early warning system of financial crises. *Journal of International Money and Finance*, vol. 25, no. 6, pp. 953-973, 2006.

[4] Tung, W.L.; Quek, C; Cheng, P. GenSo-EWS: a novel neural-fuzzy based early warning system for predicting bank failures. *Neural Networks*, vol. 17, no. 4, pp. 567-587, 2004.

[5] Yang, B; Li, L. X.; Ji, H.; Xu, J. An early warning system for loan risk assessment using artificial neural networks. *Knowledge-Based Systems*, vol. 14, no. 5-6, pp. 303-306, 2001.

[6] Tang, T. C.; Chi, L. C.; Predicting multilateral trade credit risks: comparisons of Logit and Fuzzy Logic models using ROC curve analysis. *Expert Systems with Applications*, vol. 28, no. 3, pp. 547-556, 2005.

[7] Jackson, P. *Introduction to Expert Systems*, Addison Wesley, 1998.

[8] Van Leekwijck, W; Kerre, E. E. Defuzzification: criteria and classification. *Fuzzy Sets and Systems*, vol. 108, no. 2, pp. 159-178, 1999.

[9] Kandel, A. *Fuzzy Expert System*, CRC Press, 1991.

[10] Zadeh, L. The role of fuzzy logic in the management of uncertainty in expert systems. *Fuzzy Sets and Systems*, vol. 11, no. 1-3, pp. 197-198, 1983.

[11] Zadeh, L. Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A. Zadeh (*Advances in Fuzzy Systems: Application and Theory*), World Scientific Pub Co Inc., 1996.

[12] Grosan, C.; Abraham, A. *Fuzzy Expert Systems in Intelligent Systems*, Springer Berlin, 2011, pp. 219-260.

[13] Chen, L. H.; Chiou, T. W. A fuzzy credit-rating approach for commercial loans: a Taiwan case. *Omega*, vol. 27, no. 4, pp. 407-419, 1999.

[14] Siler, W.; Buckley, J. J. *Fuzzy Expert Systems and Fuzzy Reasoning*. Wiley-Interscience, 2004.

[15] Zadeh, L. Fuzzy Sets. *Information and Control* vol. 8, pp. 338-353, 1965.