

A Financial Investment Tutoring System

M. Delgado Calvo-Flores¹, E. Gibaja Galindo^{*2}, C. Molina Fernández³ and J. Núñez Negrillo¹

¹ Departamento de Ciencias de la Computación e Inteligencia Artificial, E.T.S. de Ingeniería Informática, Universidad de Granada, 18071 Granada, Spain.

² Departamento de Informática y Análisis Numérico, Campus de Rabanales, Edificio Albert Einstein, Universidad de Córdoba, 14071 Córdoba, Spain.

³ Departamento de Informática, Campus las Lagunillas, Edificio A3 (Ingeniería y Tecnología), Universidad de Jaén, 23071 Jaén, Spain.

Stock market trading is one of the most popular forms of investment both on a corporate and individual level, yet despite the number of studies carried out, the “stock market world” is so complex that there is still no universally accepted method. In the real world, there are a number of restrictions (e.g. capital available, maximum share volume, etc.) which complicate the problem, and we can therefore consider decision-making to be expert and heuristic. This has led us to consider the development of a web-system prototype which uses association rules, fuzzy logic, and expert systems to facilitate the learning of a certain investment strategy. This system has been designed for users from the academic community who are learning about the world of finance, although it is also suitable for other types of users, such as risk managers, shareholders, and bank customers.

Keywords: Educational system, association rules, OLAP, knowledge discovery, Fuzzy, CLIPS, stock market

1. Introduction

1.1. Application Domain: the Strategy Selection Problem

There are a large number of ways of investing in the stock exchange: stock purchases, future trading, warrant purchases, government bonds, etc., and thanks to the different techniques, it is possible to obtain the estimated profit and risk of a certain action. Much has been written about this subject over the last hundred years, and many articles and applications have been published ([1, 2, 3, 4]).

The basis for any financial study are stock market records, i.e. the different market values which a share (Telefónica, EON, etc.), an asset (gold, fuel, corn, etc.), an index (Ibex 35, Dax 30, Nasdaq 100, etc.), or any other stock exchange element (warrants, etc) has had over time. In our case, we have evaluated hundreds of strategies on hundreds of values, and we therefore have a database of thousands of strategic market units (SMU) for making selection studies. Each SMU has a series of easily calculable characteristics: profit, potential, Sharpe ratio, consistency, series of losses, volatility, monthly risk, number of tradings per year, reliability, etc., but the main variables are profit and volatility. Volatility is a measure of the variance-based risk, and the greater the volatility, the greater the risk. By way of extremes, on one hand, we might have a system which is extremely profitable but which has a high degree of risk, and on the other, a strategy with a fixed interest rate and which is much less profitable, but with zero risk. This also reveals the fuzzy nature of the terms with which the expert usually works.

* Corresponding author: e-mail: egibaja@uco.es, Phone: +34 957 212 172

1.2. Motivation and aims of this work

Despite the number of studies carried out, the “*stock market world*” is so complex that there is still no universally accepted method when it comes to selecting a strategy. The ideal solution is to find a strategy which consistently produces maximum profit with minimum estimated risk.

Although there are a series of basic rules (i.e. highly volatile strategies are not advisable, very consistent strategies tend to produce hardly any profits, etc.), knowledge about the selection of the best investment portfolio is expert knowledge which has been acquired over many years of study and by solving real situations. Since this knowledge is difficult to transmit in an organized way, we have been led to consider the development of a tutoring system prototype to facilitate the learning and spreading of this heuristic knowledge. This system has been designed for users from the academic community who are learning about the world of finance, and it is presented as an attractive method for practicals. Nevertheless, it is also suitable for other types of users, such as risk managers, shareholders, and bank customers.

The informative nature of this system is another aspect which conditions the choice of solution. Since the information must be made available to the largest possible number of users, the system must be set up online with a client/server architecture. Given the characteristics of the application domain presented, the tutor uses recorded data, association rules, fuzzy logic, and an expert system to accomplish its mission.

2. Description of the Multidimensional Model

In order for the tutoring system to obtain the necessary knowledge (expressed in the form of fuzzy association rules), the data are organized according to a multidimensional model.

2.1. Multidimensional Model

The use of OLAP systems has recently been extended to new fields of knowledge due to their analysis capabilities and intuitive data access. Some of these fields have complex domains which must be modeled and require the multidimensional model to be capable of managing imprecision. Under these circumstances, the use of fuzzy logic in a multidimensional model helps in two ways:

1. The model is able to work with concepts which are not well defined.
2. Data access is more intuitive since the model can represent concepts which are more natural for users (e.g. close to average, not very low values, etc.).

Various different proposals have already been published and we propose the use of the one presented in [5] with the following features:

1. Multiple hierarchies can be defined on each dimension.
2. The model can manage imprecision in the concept definition (dimensions and facts) and in the relationships between them. As Han ([6]) mentioned, the use of OLAP systems for data mining processes (referred to as OLAM or online analytical mining) is particularly interesting. Various methods for association rule extraction have also been proposed ([7, 8]).

2.2. Description of Fuzzy Terms

With expert help, the membership functions of the different discrete values to the fuzzy sets are defined for each variable in our domain. We will use three linguistic labels: bad, normal, and good. Let us assume that a strategy has a 75% reliability. According to our fuzzy set definition, it will be bad with a membership degree of 0, normal with a membership degree of 0.6, and good with a membership degree of 0.45.

3. Description of the Tutoring System

3.1 System Structure

The system comprises the following modules:

- **A battery of cases:** We have compiled a complete battery of real cases. Each case is characterized by a fuzzy quality value given by an expert and a series of values obtained by statistical methods (e.g. estimated profit, potential, Sharpe ratio, consistency, series of losses, volatility, risk, etc). By knowing these factors, an expert is able to say whether an investment strategy is feasible, extremely advisable, unadvisable, or totally impractical. The examples analyzed correspond to European indexes (Ibex-35, Dax-30, Cac-40, Eurex-50, Nasdaq-100, Russell-1000), although the process may be easily used on any stock market.
- **A knowledge base:** This gathers the system's knowledge about the application domain (i.e. the strategy evaluation problem) and comprises a set of fuzzy association rules. In order to obtain this, an OLAP data mining technique has been applied (see Section 2). These rules represent the expert's knowledge about the domain and therefore serve as a reference for evaluating user response.
- **An inference engine:** This enables an expert's reasoning process to be emulated and to provide the system's responses. We should not forget the fuzzy nature of the problem which requires an inference engine to be used with these capabilities, and this is why Fuzzy CLIPS [9] was selected.
- **The justifier module** shows the system's conclusions. In this particular case, the conclusions are shown in the form of association rules which enable the user to understand the reasoning followed during the inference process.

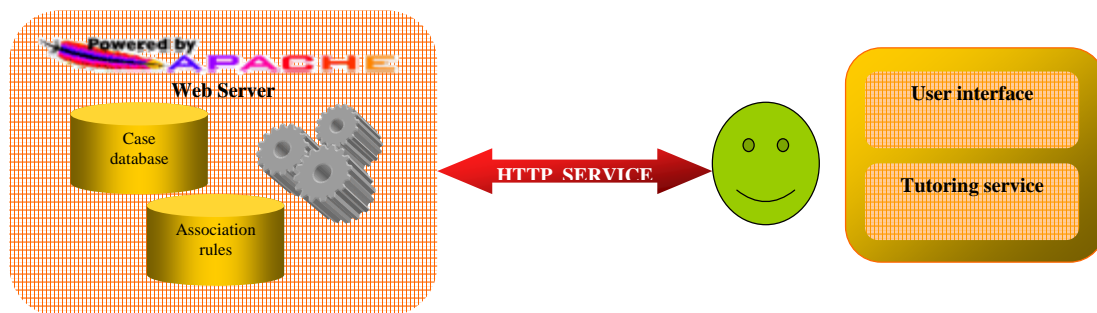


Fig. 1 System architecture

3.2 System Operation

The stages of system operation are described as follows. Initially, the user is presented with a real case from the case base so that it may be allocated a fuzzy valuation. The data input interface is a form which the user completes to send the input data to the server when the inference request is made. This enables the user to express the inputs in a fuzzy way, and in order to facilitate input of the fuzzy evaluation, a

menu with qualitative terms will be used. The system then evaluates the user's response from the knowledge base. The FCLIPS inference engine will use all the activated association rules in order to obtain a single measurement for the quality of the expert, which will be contrasted with that of the user. Finally, the system shows the user the result of this comparison and explains the expert's answer (the decision-making rules) so that they may learn the reasoning method followed by the expert and what the most important characteristics to analyze are.

3.3 System Implementation

The system has a client/server architecture which enables resources to be centralized on the application server, software to be updated without changing the clients, the system to be accessed in a unified way, and which facilitates the diffusion of knowledge.

The following tools have been used for implementation:

- **Multidimensional model:** to obtain association rules.
- **Fuzzy Clips:** to implement the inference engine.
- **HTML:** to program the web client.
- **PHP:** to program server functionality.

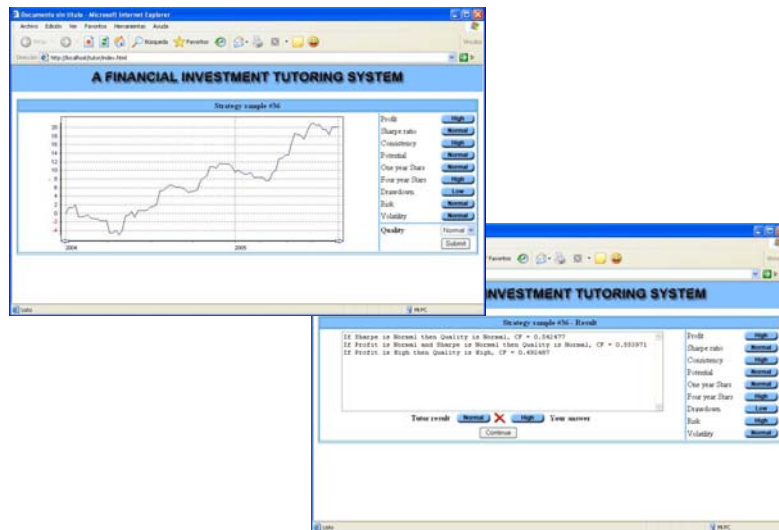


Fig. 2 Final appearance of the system

4. Operation Example

The designed software is intuitive and easy to use. Once the application has been opened, a series of fuzzy values are presented for each variable and also a graph with the strategy's accumulated performance over time (see Figure 2). Having entered the answer, the system queries the inference engine. After receiving the inference engine's response and the rules involved in the decision-making process, they are displayed in such a way that the user can progressively assimilate the concepts. For example, if the inputs are good profit, normal risk, normal potential, and normal Sharpe ratio, the system will infer a quality value of 8.66 and 7.64 according to the defuzzification method (mean-of-maximum - MOM or the centre of gravity - COG, respectively). The quality is measured on a 1-10 scale, and the value obtained corresponds to a value between normal and good. In fact, rules 167 and 6 have been fired and concluded that the quality was normal and as has rule 1, which concluded that the quality was good.

5. Conclusions and Future Work

We will finish this paper with a series of conclusions and lines of future research.

- This type of system is not common on Internet and therefore represents a ground-breaking innovation.
- The web architecture provides important advantages:
 - Great accessibility: it can be consulted any time, anywhere.
 - Wide diffusion of expert knowledge: the system's ability to reproduce expert knowledge enables widespread distribution of this knowledge at a reasonable cost.
- It solves problems involving fuzzy data: this is an important feature since it adapts to the nature of the strategy selection problem.
- It is capable of spreading expert knowledge by explaining the solution so that the user can learn the reasoning followed by the system.
- The user learns the characteristics to be observed by interacting with the query interface.

We are currently considering the possibility of tackling more complex problems (e.g. the choice of an investment portfolio) which would bring us even closer to the reality of the world of finance.

While the results obtained are satisfactory, they do lead us to consider the inclusion of this tool within a more general intelligent tutoring system as a module to generate exercises which enable the knowledge learnt by a student to be evaluated. The idea is to facilitate adaptation to the user through a register of questions/answers and the selection of the level of difficulty. In this way, basic cases would initially be presented to the user and the difficulty would be increased as the percentage of correct answers rises.

References

- [1] Baba, N., Inoue, N. et al "Utilization of Soft Computing Techniques for Constructing Reliable Decision Support Systems for Dealing Stocks". IJCNN'02: Proceedings of the 2002 International Joint Conference on Neural Networks, Honolulu, Hawaii.
- [2] Chapman, A.J. "Stock Market Trading Systems Through Neural Networks: Developing a Model", International Journal of Applied Expert Systems, Vol. 2, no. 2, 1994, pp 88-100
- [3] Liu, N. K. and Lee, K. K. "An Intelligent Business Advisor System for Stock Investment." Expert Systems 14(3): 129-139. 1997
- [4] Skabar, A., Cloete, I. "Discovery of Financial Trading Rules". Proceedings of the IASTED International Conference on Artificial Intelligence and Applications. 2001
- [5] Molina, C; Rodríguez-Ariza, L; Sánchez, D.; Vila, M.A. "A New Fuzzy Multidimensional Model". Accepted paper for IEEE Transaction on Fuzzy Systems.
- [6] Han, J., OLAP Mining: Integration of OLAP with Data Mining, in 'IFIP Conf. on Data Semantics', pp. 1-11, 1997.
- [7] Zhu, H. (1998), 'On-Line Analytical Mining of Association Rules', PhD thesis, Simon Fraser University.
- [8] Alhadj, R. & Kaya, M. (2003), Integrating Fuzziness into OLAP for Multidimensional Fuzzy Association Rules Mining, in 'ICDM '03: Proceedings of the Third IEEE International Conference on Data Mining', IEEE Computer Society, Washington, DC, USA, pp. 469.
- [9] Orchard, R.A. "Fuzzy CLIPS Version 6.04A. User Guide". Integrated Reasoning. Institute for Information Technology. National Research Council Canada. October 1998.