

Johann-Adolf Mueller and Frank Lemke

Self-Organising Data Mining

An Intelligent Approach To Extract Knowledge From Data

Extracted from Chapter 7 with permission of the authors.

7.5.2 Special module: Finance

KnowledgeMiner provides a small module for supporting financial trading decisions. Its intention is integrating self-organising data mining into financial decision making by applying the idea of predictive control based on the example described in section 2.3.2. In result, a self-adjusting, predictive, and rather fuzzy trading indicator is obtained.

An analysis of financial markets (section 2.3.2) points out that

- a. financial markets are integrated. An intermarket financial analysis needs adequate tools for modelling of complex systems.
- b. financial markets are fuzzy processes that are not susceptible to a purely quantitative analysis. According to Zadeh's principle of incompatibility, as the complexity of a system increases, our ability to make significant statements about the system's behaviour diminishes so that, finally, precision and significance are becoming almost mutually exclusive characteristics. The closer one look at a real-world problem the fuzzier becomes its solution (see also section 1.3).
- c. financial systems are nonlinear and instationary.

Starting from these facts, two sorts of models are closely integrated.

Predictive Model

Applying the ideas of section 6.2 (fig. 6.2), there is to find an algorithm that satisfies the above requirements and that also fits the task of prediction. When using daily price data, the first choice are nonparametric modelling methods, and here Analog Complexing especially (section 5.2), because

- it covers market behavior by searching for similar market situations on a set of representative market player's data. Analog Complexing uses exclusively the observed data to predict actual market developments capturing in this way any market influences stored in a pattern's price data respectively a pattern's continuation.

- financial markets, on a daily price basis, are the most fuzzy and noisy processes and require adequate models. Analog Complexing forms models from the data itself ensuring equal fuzziness of object and model.
- nonlinearity and dynamics of the process is described by the method inherently due to its self-referencing modelling approach. Instationarity can be explained by repeating modelling daily on updated data accomplishing moving modelling. A synthesis of several models indicating different market situations and different predictions is integrated in the algorithm. A result is that a measure of prediction uncertainty is available.
- finally, Analog Complexing is a very fast working modelling method and can be applied at many assets likewise.

The asset's recent volatility is used by KnowledgeMiner to automatically determine the forecast horizon of the predictive model following the rule "the more volatile the asset, the shorter the forecast horizon". Common values are between 1 and 5 days.

Decision Model

Driven by Zadeh's principle of incompatibility and based on the Moving Average Convergence Divergence (MACD) principle (2.3.2), we have modified this indicator for KnowledgeMiner in two directions:

1. One important disadvantage of the MACD is that only historical data are used to describe trend developments of financial processes by some smoothing. Thus, the resulting trading signal will probably lag advantageous trading points in time. Financial trading has a forward looking nature. An exclusive retrospective reflection of the market is not sufficient, obviously, for decision making. Any time lag of the decision model forced by averaging or smoothing has to be compensated (ideally) or minimized (practically) by a predictive approach. A predictive model generates prediction uncertainty. A forward looking decision model should reflect this uncertainty as well.
2. The difference of the MACD and the Signal line is a line (let's call it the indicator line) that oscillates around zero. Concluding from the MACD principle described in section 2.3.2, the common decision rule of the MACD indicator is that each time the indicator line crosses the zero line, a corresponding trading signal is suggested (positive indicator: buy; negative indicator: sell). When applying the MACD to a series of random numbers as a reference model, many buy/sell trades are generated by the indicator while a correct model should suggest no trades at all. Increasing the averages for MACD calculation will not really help, it will decrease the absolute of the indicator's amplitude only. We concluded that, for a given set of averages, there is a fuzzy, time variant interval of uncertainty $(- \delta, \delta)$, $|\delta| > 0$, between which only random changes of the examined time series will be reflected. The uncertainty δ can be estimated by GMDH, for example, using the reference model and the asset's volatility, a common characteristics of financial time series. Implementing an uncertainty estimation model into the MACD indicator, a kind of fuzzy decision model is obtained. A signal is generated when the asset's actual price changes is expected not to fit into the asset's recent trend pattern, that is, when the indicator line is close or above (below) the uncertainty interval's border. This interval is computed daily reflecting the rule "the more volatile the asset, the larger decision uncertainty" and must be considered fuzzy (fig. 7.34/ 7.35).

Indicator Calculation

The modified MACD indicator can be applied in two ways:

- a. using the data stored in the spreadsheet and
- b. predictive model based indicator calculation.

a. Spreadsheet based calculation

Here, data stored in the spreadsheet are used to calculate the indicator. These data can be an asset's historical price data, but also future expectations or assumptions of its development. Selecting the "Spreadsheet Based Calculation" mode and choosing "Generate Self-Adjusting Trading Indicator" from the Finance menu (fig.7.33), the indicator will be calculated and graphically shown beginning from the selected row of the column (fig. 7.34a). The advantage of this option is that it allows what-if type scenarios by considering assumed future values or known future events and their expected impact on the equity's price while getting different indicator results and signals.

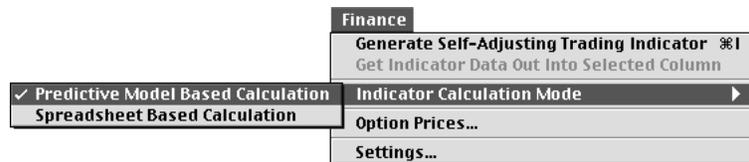
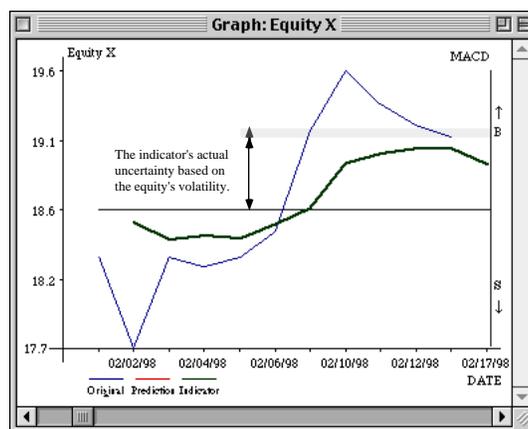


Fig. 7.33: The Finance menu of KnowledgeMiner

926	01/28/98	19.19	13493
927	01/29/98	18.50	18918
928	01/30/98	18.31	14504
929	02/02/98	17.69	56852
930	02/03/98	18.31	35948
931	02/04/98	18.25	15196
932	02/05/98	18.31	21341
933	02/06/98	18.50	18066
934	02/09/98	19.19	44167
935	02/10/98	19.60	
936	02/11/98	19.38	
937	02/12/98	19.23	
938	02/13/98	19.15	

a. Spreadsheet based calculation: The graph will be displayed from the selected row (02/10/98 to 02/13/98)



b. Corresponding indicator line (green) displayed for the chosen past period (02/02/98 - 02/10/98) and the forecasted period (02/11/98 - 02/17/98) (the one day shift is due to exponential smoothing of the indicator)

Fig. 7.34: MACD indicator calculated on historical price data and expected data stored in the spreadsheet

b. Predictive model based indicator calculation

When this option is selected (fig. 7.33, default setting), an AC model of the examined asset is used internally to generate a predictive indicator. If the document contains no AC model of that asset, this model will be generated by KnowledgeMiner automatically. In contrast to *a.*, here the uncertainty of the asset's prediction (fig. 7.17) is transformed into indicator prediction uncertainty correspondingly. Figure 7.35 shows the equity curve with its predictions, the predicted indicator along with its most optimistic, most likely and most pessimistic future reactions, and the decision model's actual uncertainty. This graph can be obtained easily by choosing "Generate Self-Adjusting Trading Indicator" from the Finance menu on a daily basis (when using daily close prices).

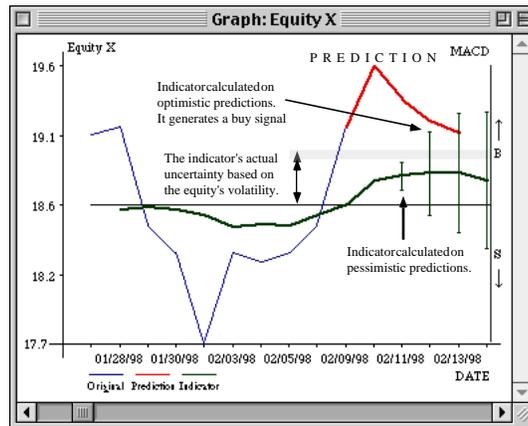


Fig. 7.35: Predictive model based indicator: Uncertainty of signals based on vagueness of prediction and the equity's recent volatility

Since this module will evolve fast, the following Internet address can be referenced for updated information:

<http://www.knowledgeminer.net>