| Petroleum-Gas University of Ploiesti BULLETIN | Vol. LXIII No. 3/2011 | 105-112 | Economic Sciences Series |
|---|---|---|---|

# The Analysis of Currency Exchange Rate Evolution using a Data Mining Technique

Mădălina Cărbureanu

Petroleum-Gas University of Ploieşti, Bd. Bucureşti 39, Ploieşti, Romania
e-mail: mcarbureanu@upg-ploiesti.ro

## Abstract

*The evolution of leu-euro rate of exchange has a definite influence on Romanian real economy, permanently being searched that balanced exchange rate, to assure the financial market stability. Hereby, it is necessary the permanent analysis of the currency exchange rate, including the main factors that influence its fluctuations. With a view to achieving this aim, it is necessary a set of informatics tools dedicated to exchange rate evolution analysis, tools that use various techniques. In the present paper, it is used a data mining technique, namely classification for analyzing the evolution of leu-euro rate of exchange for a certain period of time. For implementing the results obtained through this technique, it will also be developed a system for analyzing the evolution of this currency exchange rate. Such a system can be a useful tool for economists, financial analysts, statisticians etc., in studying the exchange rate, also in predicting its evolution.*

**Key words:** *currency exchange rate, data mining, classification, analysis*

**JEL Classification:** *C38, C81*

## Introduction

The political, economic and social events from a certain period of time have a positive or negative influence on currency demand and supply, with instant effects on exchange rate evolution. Moreover, there is a set of factors with direct impact on leu-euro exchange rate evolution, such as: gross domestic product, the buying capacity on inner market, the inflation evolution (consumer price index), the interest evolution, a set of psychological factors (such as the population and economic agents' confidence in a certain currency), monetary board, the credits accorded by credit institutions etc[1].

In achieving the classification and developing the proposed system, we have considered the following factors: consumer price index (CPI), the monetary board, the credits accorded by credits institutions and gross domestic product (GDP). These factors were chosen, because through them it is carried out the monetary authority's influence on the currency exchange rate evolution.

The *structure* of the paper is organized as follows:

[1] Enache, C., Mecu, C., *Economie politică*, Editura Fundației România de Mâine, Bucureşti, 2007, pp. 298.

- Data mining techniques as an analysis and prediction instrument; this section introduces the main applications and advantages of using data mining techniques in the financial domain;
- The application development and the results interpretation, a section that presents the C5.0 data mining algorithm, the decision rules obtained through the application of this algorithm on the training data supplied by Romanian National Bank (BNR) and the statistical results are interpreted;
- The classification results implementation, section in which the classification results are used for building a system for leu-euro exchange rate evolution analysis, using CBuilder programming environment. There are also presented the main advantages of the proposed system and it is achieved the prediction of leu-euro rate of exchange for May 2011;
- Conclusions, where there is an emphasis upon the original elements brought by the author through this paper.

## Data Mining Techniques as Instrument for Analysis and Prediction

Banks and other financial institutions are maintaining huge electronic data repositories, data that contains valuable information for these institutions. Because these repositories are huge, the human analyst cannot cope with so much data, being very difficult to extract the useful information without the help of a specific tool[2].

The data mining techniques represent such a tool that solves different types of problems from banking and finance domains, by finding patterns, correlations, rules sets, causalities etc., and helps the human analyst in the process of analysis and prediction of some financial tasks evolution, such as: currency exchange rate, stock market, bank bankruptcies, financial risk, credit rating, loan management, bank customer profiling, money laundering etc. Furthermore, some data mining packages offer statistical methods, such as principal components, logistic regression, correspondence analysis[3] etc., for financial predictions[4]. From the data mining techniques category, the ones that are mostly applied in the banking and financial domain, respectively in the currency exchange rate analysis and prediction are those presented in Figure 1[5].

The problem of analysis and prediction of the currency exchange rate evolution can be solved using the classification data mining technique. The classification problem involves predicting whether the exchange rate has increased or decreased in a certain period of time.

The classification is the process through which it is identified a set of common features and models that describe data classes or concepts.

For predicting the foreign exchange rate evolution a large number of classification models were developed, models that provide insight into attributes that are important for this specific application. In finance, many applications combine data mining techniques with various finance and accounting models, such as capital asset pricing model and the Kareken–Wallace model[6].

---

[2] *** Data Mining in Banking and Finance, at http://www.scribd.com/doc/28085348/Data-Mining-in Banking-and-Finance.

[3] Ioniță, I., Cărbureanu, M., *Correspondence Analysis in the Banking System*, Petroleum-Gas University of Ploiești, Economic Sciences Series, Vol. LXII, No. 2/2010, pp. 102-109.

[4] Kovalerchuk, B., Vityaev, E., *Data mining in finance. Advances in Relational and Hybrid Methods*, Kluwer Academic Publisher, Boston/ Dordrecht/London, 2000, pp. 16.

[5] ***Data Mining for Financial Applications, at http://www.math.nsc.ru/AP/ScientificDiscovery/PDF/ data_mining_for_financial_applications.pdf.

[6] Zhang D., Zhou L., *Discovering Golden Nuggets: Data Mining in Financial Application*, IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, vol. 34, no. 4, 2004, pp. 513-514.

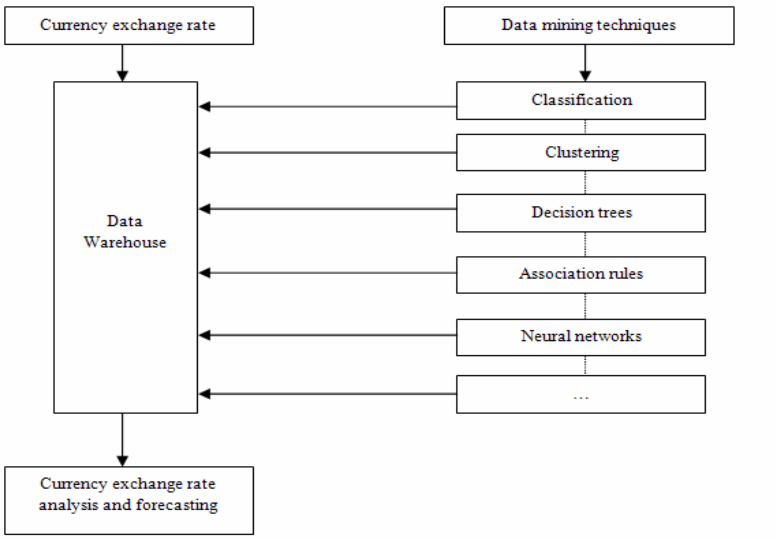


**Fig. 1.** Data mining techniques

The analysis and forecasting using data mining techniques are essential in all business organizations. The ability to accurately predict the future is essential to many decision activities in banking, finance, marketing, production, and other business areas in order to improve the organization's performance in decision making[7].

## Application Development and Interpretation of Results

For the analysis of leu-euro rate of exchange, classification[8] as a data mining technique will be employed. The results obtained through the application of this technique will be implemented into a system developed in CBuilder programming environment, system whose goal is to achieve the analysis and prediction of leu-euro exchange rate.

Aiming at developing this system, first of all, it is necessary to obtain a set of rules through classification, respectively through the usage of C5.0 algorithm implemented in See5 system[9]. C5.0 algorithm generates the so-called classifiers under decision trees form or under if-then rules form[10] .

Among the advantages of using this algorithm we mention[11] :

- improved algorithm for classification building (raised accuracy, minimal memory consuming, raised processing speed);
- the generated decision tree, respectively the generated set of rules it is, from the dimension point of view, considerably much smaller than the decision trees or set rules generated through other data mining algorithms, such as ID3 or C4.5.

---

[7] Maimon, O., Rokach, L., *Data mining and knowledge discovery handbook*, second edition, Springer Science+Business Media, New-York, 2010, pp. 507-508.
[8] Gorunescu, F., *Data mining. Concepte, modele şi tehnici*, Editura Albastră, Cluj-Napoca, 2006, pp. 23-24.
[9] See5: An Informal Tutorial, at http://www.rulequest.com/see5-win.html.
[10] Khosrow-Pour, M., *Emerging Trends and Challenges in Information Technology Management*, Editura Idea Group, 2006, pp. 282-283.
[11] ibidem

The data base is built using values supplied by BNR (monthly bulletins and interactive data base), for the main factors (medium leu-euro exchange rate, consumer price index, monetary board, credits and gross domestic product) that influence the leu-euro exchange rate evolution, for January 2007-April 2011 period of time[12].

The data base that contains the necessary data for achieving the analysis has the structure presented in Table 1.

**Table 1.** The data base

| Year | Month | Medium leu-euro exchange rate | Consumer price index (IPC) (%) | Monetary board (thousands lei) | Credits accorded by credits institutions (millions lei) | Gross domestic product (GDP) (millions lei) |
|---|---|---|---|---|---|---|
| 2007 | January | 3.3937 | 4.01 | 106626394 | 108365 | 68841.5 |
| | February | 3.3824 | 3.81 | 109614957.4 | 111187 | 68841.5 |
| | March | 3.3694 | 3.66 | 112696490.6 | 116233 | 68841.5 |
| | April | 3.3349 | 3.77 | 113134878.8 | 119415 | 87063.7 |
| | May | 3.285 | 3.81 | 112826481.8 | 123784 | 87063.7 |
| | June | 3.2264 | 3.8 | 116276131.8 | 129191 | 87063.7 |
| | July | 3.1337 | 3.99 | 120040473.9 | 135592 | 111035.1 |
| | August | 3.2237 | 4.96 | 124457523.1 | 141924 | 111035.1 |
| | September | 3.3466 | 6.03 | 126678916.4 | 147687 | 111035.1 |
| | October | 3.3525 | 6.84 | 128873129.1 | 154191 | 137768.5 |
| | November | 3.4707 | 6.67 | 136170985.2 | 161214 | 137768.5 |
| | December | 3.5289 | 6.57 | 148115532.6 | 167956 | 137768.5 |
| 2008 | January | 4.1409 | 7.26 | 147531171.4 | 171020 | 91130.3 |
| | February | 4.1179 | 7.97 | 149762195.6 | 184721 | 91130.3 |
| | March | 4.0879 | 8.63 | 151859227.6 | 191649 | 91130.3 |
| | April | 4.1285 | 8.62 | 157088216.3 | 197304 | 115074.3 |
| | May | 4.1743 | 8.46 | 157605324.9 | 203389 | 115074.3 |
| | June | 4.2396 | 8.61 | 161495383.8 | 209356 | 115074.3 |
| | July | 4.2611 | 9.04 | 161297857 | 215555 | 138323.7 |
| | August | 4.2389 | 8.02 | 162351393 | 221609 | 138323.7 |
| | September | 4.2642 | 7.3 | 166092031.4 | 226730 | 138323.7 |
| | October | 4.2798 | 7.39 | 162522556.7 | 229118 | 159430.4 |
| | November | 4.2931 | 6.74 | 164727400.5 | 229321 | 159430.4 |
| | December | 4.2925 | 6.3 | 174027815.6 | 228482 | 159430.4 |
| 2009 | January | 4.2327 | 6.71 | 176103960.9 | 229131 | 96616.7 |
| | February | 4.2839 | 6.89 | 176205108.1 | 227622 | 96616.7 |
| | March | 4.2821 | 6.71 | 175288253 | 227264 | 96616.7 |
| | April | 4.1954 | 6.45 | 176365560.8 | 227012 | 112073 |
| | May | 4.1689 | 5.95 | 177304666.6 | 225516 | 112073 |
| | June | 4.2126 | 5.86 | 180325700.7 | 224888 | 112073 |
| | July | 4.2168 | 5.06 | 181383974.9 | 226671 | 130288.7 |
| | August | 4.2185 | 4.96 | 183963270.7 | 225569 | 130288.7 |
| | September | 4.2389 | 4.94 | 183819079.4 | 229325 | 130288.7 |

[12] Romanian National Bank (BNR), Monthly bulletins, 2007-2011, at http://www.Bnro.ro/ PublicationDocuments.aspx?icid=1182.

*Table 1 (cont.)*

| | | | | | | |
|---|---|---|---|---|---|---|
| | October | 4.2848 | 4.3 | 183992054.2 | 230025 | 152295.3 |
| | November | 4.2881 | 4.65 | 185553385.3 | 229168 | 152295.3 |
| | December | 4.2248 | 4.74 | 189630290.7 | 231561 | 152295.3 |
| 2010 | January | 4.1409 | 5.2 | 185996884.4 | 232750 | 97263.3 |
| | February | 4.1179 | 4.49 | 187539413.4 | 232832 | 97263.3 |
| | March | 4.0879 | 4.2 | 189976874.3 | 233683 | 97263.3 |
| | April | 4.1285 | 4.28 | 190920697.3 | 234658 | 117126.5 |
| | May | 4.1743 | 4.42 | 192858601.5 | 236309 | 117126.5 |
| | June | 4.2396 | 4.38 | 195087511.4 | 238464 | 117126.5 |
| | July | 4.2611 | 7.14 | 192349853.5 | 238444 | 139408.3 |
| | August | 4.2389 | 7.58 | 195719515.7 | 239980 | 139408.3 |
| | September | 4.2642 | 7.77 | 195738785.4 | 242188 | 139408.3 |
| | October | 4.2798 | 7.88 | 194803995 | 241727 | 159842.7 |
| | November | 4.2931 | 7.73 | 197287251.2 | 243604 | 159842.7 |
| | December | 4.2925 | 7.96 | 202763379.2 | 246441 | 159842.7 |
| 2011 | January | 4.2622 | 6.99 | 199216247.8 | 246259 | 162240.34 |
| | February | 4.2472 | 7.6 | 198130351.8 | 247166 | 162240.34 |
| | March | 4.1646 | 8.01 | 196330942.3 | 247314 | 162240.34 |
| | April | 4.0992 | 8.34 | 196484004.9 | 247803 | 162727.06 |

Regarding the target attribute that is leu-euro rate of exchange, for framing the values associated to it, it was necessary the building of three intervals taking into consideration the exchange rate fluctuations through time, as it can be observed in Table 2.

**Table 2.** Leu-euro exchange rate intervals

| **Leu-euro exchange rate** | | |
|---|---|---|
| **low** | **medium** | **high** |
| <3.5 | [3.5; 4.2] | >4.2 |

Appling C5.0 algorithm on the data from table 1, it was obtained a set of decision rules as it follows:

- if monetary<=136170000 then leu-euro exchange rate=low;
- if monetary>136170000 and credit<=203389 then leu-euro exchange rate=medium;
- if credit>247166 then leu-euro exchange rate=medium;
- if IPC<=6.57 and monetary>136170000 and GDP<=117126.5 then leu-euro exchange rate=medium;
- if monetary>136170000 then leu-euro exchange rate=high.

C5.0 algorithm is used to classify all four available attributes, of which the monetary board attribute has the greatest contribution to the classification.

From those fifty-two cases used for classification under decision rules form, there were misclassified only two cases, from which an error of 3.8% resulted, smaller than the error generated through classification under decision tree form, as it can be observed in table 3.

**Table 3.** Statistical results

| **Statistical data** | **Decision rules** | **Decision tree** |
|---|---|---|
| Error | 3.8% | 5.8% |
| Misclassified cases | 2 | 3 |
| Correct classified cases | 50 | 49 |

According to the results from table 3, it can be observed that the classification under decision rules form is more efficiently from statistical point of view than the classification under decision tree form. The application of C5.0 algorithm it also generated a so-called confusion matrix, as it can be observed in table 4.

**Table 4.** Confusion matrix

| a | b | c | |
|---|---|---|---|
| 11 | 0 | 0 | a-class “low” |
| 0 | 15 | 0 | b-class “medium” |
| 0 | 2 | 24 | c-class “high” |

The interpretation of the generated confusion matrix is the following:

- two of the cases *medium* were misclassified as *high*;
- the others fifty cases (the principal diagonal) are correctly classified.

## Classification Results Implementation

For implementing the results obtained through classification, it was developed an analysis system in CBuilder programming environment, using the set of rules obtained through C5.0 application and the statistical results previously presented.

The system interface is presented in figure 2.

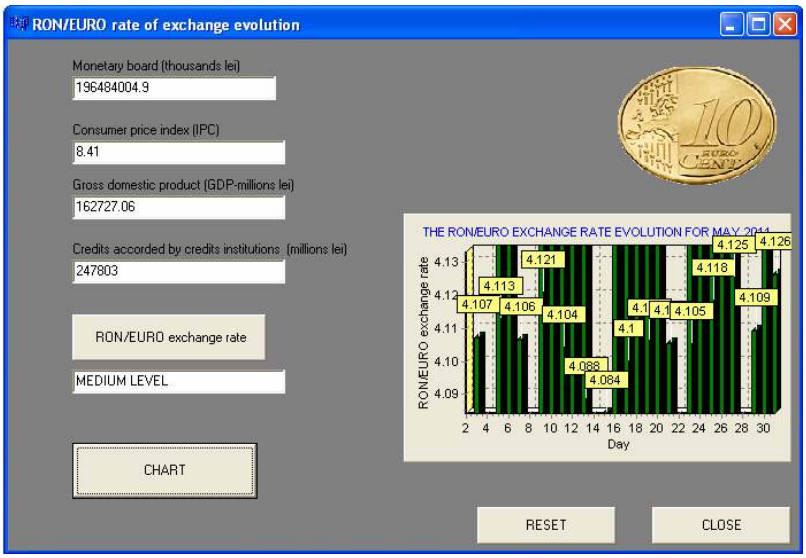


**Fig. 2.** Analysis system interface

The developed system for leu-euro exchange rate analysis offers to the user the following facilities:

- based on monetary board, consumer price index (IPC), gross domestic product (GDP) and credits values, the users (economists, financiers, statisticians etc.) can know the leu-euro exchange rate evolution (low, medium and respectively high level of exchange rate);
- the possibility for graphical representation of exchange rate evolution for a certain month;
- the possibility to predict the evolution of leu-euro exchange rate.

For the relevance of the developed system, we tested it with other data, different from the training data, as it is shown in table 5.

**Tabel 5.** System validation

| Year | Month | Medium leu-euro exchange rate | Monetary board (thousands lei) | Consumer price index (IPC) (%) | Gross domestic product (GDP) (millions lei) | Credits accorded by credits institutions (millions lei) |
|---|---|---|---|---|---|---|
| 2006 | April | 3.4911 Low (<3.5) | 94157000 | 6.92 | 75967 | 75903 |
| | November | 3.4954 Low(<3.5) | 101016000 | 4.67 | 112476 | 102762 |

In table 6, using the values for monetary board, IPC, GDP and credits, values supplied by BNR, it is achieved the prediction of leu-euro exchange rate evolution for May 2011.

**Tabel 6.** System results

| Year | Month | Medium leu-euro exchange rate | Monetary board (thousands lei) | Consumer price index (IPC) (%) | Gross domestic product (GDP) (millions lei) | Credits accorded by credits institutions (millions lei) |
|---|---|---|---|---|---|---|
| 2011 | May | Medium [3.5; 4.2] | 196484004.9 | 8.41 | 162727.06 | 247803 |

As we can observe in table 6, in May 2011 it is estimated a medium leu-euro exchange rate, within [3.5; 4.2] lei interval.

## Conclusions

The goal of this paper was to highlight the applicability and utility of a data mining technique, namely classification in the analysis and prediction of leu-euro exchange rate evolution. In this sense, using the set of rules obtained through classification, it was developed a prototype system that can be a useful tool for economists, financial analysts, statisticians, etc., in currency exchange rate analysis and prediction.

What differentiates the prototype system from other systems in the same category (systems for currency exchange rate evolution analysis and prediction) is the fact that in developing it, it was employed a data mining technique as an analysis and forecasting instrument. From the available data mining techniques, we chose classification due to its large applicability in banking and financial domains and because there were developed a large number of classification models for currency exchange rate analysis and prediction.

The analysis and forecasting of currency exchange rate evolution are essential for understating the financial trends and fluctuations on the market. For achieving a relevant analysis and prediction of currency exchange rate evolution, huge data bases are processed. For processing this type of data bases, it is necessary a tool for helping the human analyst in his work. The data mining techniques offer such tools that prove to be the essential element in achieving a relevant analysis and prediction.

The developed system can be extended with other options, such as the possibility of analyzing and forecasting other pairs of currency (USD/RON, EUR/USD etc.) in a graphical form, for a longer period of time. Moreover, the system can be improved for maintaining a history of the past forecasts for different currency pairs and it can be added with an exchange rate calculator.

Therefore, all above-mentioned issues recommend the usage of data mining techniques in the analysis and prediction of the financial market indicators for a company's or an institution's successful activity.

## References


1. *** See5: An Informal Tutorial, at http://www.rulequest.com/see5-win.html.
2. *** Romanian National Bank (BNR), Monthly bulletins, 2007-2011, at http://www.Bnro.ro/ PublicationDocuments.aspx?icid=1182.
3. *** Data Mining in Banking and Finance, at http://www.scribd.com/doc/28085348/Data-Mining-in-Banking-and-Finance.
4. *** Data Mining for Financial Applications, at http://www.math.nsc.ru/AP/ScientificDiscovery/PDF/ data_mining_for_financial_applications.pdf.
5. Enache, C., Mecu, C., *Economie politică*, Editura Fundației România de Mâine, Bucureşti, 2007.
6. Gorunescu, F., *Data mining. Concepte, modele şi tehnici*, Editura Albastră, Cluj-Napoca, 2006.
7. Ioniță, I., Cărbureanu, M., *Correspondence Analysis in the Banking System*, Petroleum-Gas University of Ploieşti, Economic Sciences Series, Vol. LXII, No. 2/2010, pp. 102-109.
8. Khosrow-Pour, M., *Emerging Trends and Challenges in Information Technology Management*, Editura Idea Group, 2006.
9. Kovalerchuk, B., Vityaev, E., *Data mining in finance. Advances in Relational and Hybrid Methods*, Kluwer Academic Publisher, Boston/ Dordrecht/London, 2000.
10. Maimon, O., Rokach, L., *Data mining and knowledge discovery handbook*, second edition, Springer Science+Business Media, New-York, 2010.
11. Zhang D., Zhou L., *Discovering Golden Nuggets: Data Mining in Financial Application*, IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, vol. 34, no. 4, 2004, pp. 513-514.


# Analiza evoluției cursului de schimb valutar prin utilizarea unei tehnici de data mining

## Rezumat


*Evoluția cursului de schimb valutar leu-euro influențează în mod categoric economia reală a României, căutându-se în permanență acel curs valutar de echilibru, care să asigure stabilitatea pieței financiare. Astfel, este necesară analiza permanentă a cursului valutar, precum şi a principalilor factori care îi influențează fluctuațiile. Pentru aceasta este nevoie de o serie de instrumente informatice dedicate analizei evoluției cursului de schimb valutar, instrumente care au la bază o multitudine de tehnici. În acest articol este utilizată o tehnică de data mining, şi anume clasificarea pentru analiza evoluției cursului leu-euro pentru o anumită perioadă de timp. Pentru implementarea rezultatelor obținute prin aplicarea acestei tehnici, se va dezvolta şi un sistem pentru analiza evoluției acestui curs de schimb. Un astfel de sistem poate fi un instrument util economiştilor, analiştilor financiari, statisticienilor etc., în studiul cursului de schimb, precum şi în previzionarea evoluției acestuia.*