Dorina Marghescu| Shuhua Liu| Peter Sarlin

Evaluation of a Fuzzy C-Means Model in Currency Crisis Prediction

TURKU CENTRE *for* COMPUTER SCIENCE

TUCS Technical Report
No 970, April 2010

Evaluation of a Fuzzy C-Means Model in Currency Crisis Prediction

Dorina Marghescu

Åbo Akademi University, Department of Information Technologies
Data Mining and Knowledge Management Laboratory
dmarghes@abo.fi

Shuhua Liu

Academy of Finland
Åbo Akademi University, Department of Information Technologies
Institute for Advanced Management Systems Research
sliu@abo.fi

Peter Sarlin

Åbo Akademi University, Department of Economics
psarlin@abo.fi

TUCS Technical Report
No 970, April 2010

Abstract

In this paper, we analyze financial crisis data using the Fuzzy C-means (FCM) clustering technique. The aim is to investigate the extent to which this method is useful in predicting currency crises in emerging economies. The focus is on the currency crises that took place in Asia during 1997. We analyze a dataset that concern 23 developing countries all over the world and five variables that have been experimented in other similar studies. Thus, we build, using the FCM technique, an early warning system for forecasting the Asian crises based on historical data. The approach consists of clustering the data using the FCM technique and then classifying the clusters into early warning and tranquil clusters. We compare the prediction results of the model with those obtained by applying probit analysis. This study shows that the training accuracy of a FCM model is better than that of the probit model, while the prediction accuracy calculated for the test datasets is slightly worse. However, the FCM model has a good explanatory feature of different currency crises, that is, diverse crises over different periods of time and in different countries are assigned to various clusters whose centers can be regarded as representative conditions for financial instability. Therefore, we conclude that the high classification accuracy in training, coupled with the intuitive explanations of different crisis situations are indicators that a model based on the FCM technique could be useful for analyzing and predicting financial instability. However, future work is intended to improve the current model with respect to its predictive performance.

Keywords: Fuzzy C-means, probit analysis, currency crisis, financial instability, early warning signals, prediction, evaluation

1. Introduction

The literature concerning the analysis and prediction of financial crises is growing at a fast rate now, because of an increased interest among economists in this subject due to the recent sub-prime crisis in USA and debt crises in Dubai and Greece among other countries. Theoretical and empirical studies try to explain the phenomenon and build models of how financial crises emerge. Financial crises are broadly defined as the occurrence of a sudden fall in the value of assets and/or financial institutions that may have dramatic effects on the economies. Currency crises are an instance of financial crises and they are characterized by a devaluated or floated fixed exchange rate due to massive capital outflows. A currency crisis is often caused by speculative attacks, for example, the Asian crises in 1997.

Particularly interesting for research are those empirical models that attempt to predict the occurrence of a financial crisis a long time before this appears, so that decision makers can take actions in order to prevent the crisis or mitigate its effects. These empirical models are known under the name of early warning systems. In this paper, we study the use of the Fuzzy C-means (FCM) clustering technique as a basis for an early warning system intended to identify signs of an impending currency crisis. The time frame in which such signs can appear (i.e., signaling horizon) usually consists of 24 months before the occurrence of the unwanted event (Kaminsky et al. 1998).

The most cited early warning system for currency crises has been developed by Kaminsky, Lizondo and Reinhart (1998); the so-called “signals” approach and denoted by KLR hereafter. This system uses 15 macroeconomic variables regarding 20 economies (15 developing and 5 industrialized countries) during 1970–95. The variables have been chosen based on theoretical studies and data availability. In total, 76 currency crises were analyzed. The KLR approach identifies the indicators that are able to signal successfully the occurrence of a crisis in the near future (i.e., within 24 months from the signal). Once the indicators are identified, the approach can be used to monitor each of these variables and observe when they exceed a determined threshold.

The KLR approach has been evaluated in many studies of predicting the Asian crises (e.g., Furman and Stiglitz 1998). One notable evaluation of KLR and also a proposed alternative of that approach is the study by Berg and Pattillo (1999a). They reproduced the KLR study on the same data and found similar results, though small differences were observed in the importance of some indicators, differences that could be due to the errors in the raw data that the International Monetary Fund has revised in time. However, more importantly is that Berg and Pattillo have proposed an alternative to the KLR approach, namely a model based on probit analysis that outperforms the in-sample accuracy of the former model. As a consequence, since the publication of the Berg and Pattillo (1999a) study, many empirical models have been built and tested (Dumitrescu et al. 2010).

In this paper, we build and evaluate a model based on the FCM clustering technique for analyzing and predicting currency crises in Asia in 1997. Our approach consists of clustering the data into similar groups, and then classifying the clusters into early warning and tranquil clusters. The research question is to what extent a FCM-based model using historical data from emerging economies would have been successful in predicting the Asian crises in 1997. The importance of the question in the current global economic context is that if its answer is encouraging, then the approach stimulates further research on analyzing indicators of recent crises using a similar model. The motivation of using FCM in our model is given by a recent study of the currency crisis in Finland in 1992, where the fuzzy clustering approach has been used to determine significant fluctuations in economic conditions that corresponded to different phases of a crisis cycle (Liu and Lindholm 2006). Liu and Lindholm analyzed single-country quarterly data covering the period 1984–1994. The FCM algorithm was used to identify crisis early warning signals and to show the corresponding cyclical fluctuations of economic indicators of currency crises; in their model, cluster centers indicate representative economic conditions, and cluster memberships, the fluctuations. Moreover, the choice of a clustering technique was driven by the encouraging results we obtained by applying the Self-Organizing Map technique to the same problem (Sarlin and Marghescu 2010). Self-Organzing Map is a neural-network based clustering and visualization technique developed by Teuvo Kohonen in 1981 (Kohonen 2001). The clustering and mapping of data onto a 2-dimensional grid takes into account the similarities between the data points.

2. The FCM model of predicting currency crises

2.1 The FCM technique

The FCM algorithm was developed by Jim Bezdek (1981) and is the mostly used fuzzy clustering method. It recognizes hyper-spherical clouds of points in a multidimensional space. A cluster center is determined as the weighted average over the data points; the weights being the cluster membership values to that cluster. The cluster centers are representative points of all the data assigned to that cluster. The fuzzy characteristic of the method determines whether a data point belongs strongly to one cluster and weakly to the other clusters, or belongs to all clusters with similar degrees of memberships, depending on the inherent structure of the dataset. The degree of overlapping between clusters is controlled by a parameter called fuzziness indicator, henceforth denoted by m. A small value of this parameter (i.e., close to 1) results in a crisp clustering model, namely the popular C-means technique (MacQueen 1967).

The algorithm starts with a random initialization of the membership values to a specified number of clusters, c. The algorithm then tries to update in an iterative manner the membership values so that the data points will have stronger membership values to the clusters whose centers are more similar with those data points. The similarity is measured using a distance function, usually the Euclidean distance.

2.2 Data

The data are a replica of the dataset analyzed by Berg and Pattillo (1999b), henceforth BP. The dataset consists of monthly data regarding 23 developing countries[1] observed during 1970:1–1997:12. The independent variables are the foreign exchange reserve growth (RESG), the export growth (EXPG), the real exchange rate overvaluation relative to trend (RDEV), the current account deficit relative to GDP (CANE), and the short-term debt in relation to reserves (STDR2).

[1] The countries included in the analysis are Argentina, Bolivia, Brazil, Chile, Colombia, India, Indonesia, Israel, Jordan, Korea, Malaysia, Mexico, Pakistan, Peru, Philippines, South Africa, Sri Lanka, Taiwan, Thailand, Turkey, Uruguay, Venezuela, and Zimbabwe.

The predicted variable is a binary variable that indicates whether or not a currency crisis occurs within 24 months. This variable is denoted in the dataset by C24 and is henceforth referred to by this name. Its value is 1 if the data point is a pre-crisis month, and 0, otherwise. The pre-crisis months represent in our model the early warning signals that we aim to predict. A currency crisis is defined in a similar way as in BP and KLR; a crisis is defined to occur when the sum of a weighted average of monthly percentage depreciation in the exchange rate and monthly percentage declines in reserves exceeds its mean by more than three standard deviations. Moreover, to avoid that crises are associated with high inflation, the sample is split into periods with low and high inflation, whereafter separate indices are constructed for each of the samples. The pre-crisis period is defined to start 24 months before the crisis episode measured by the earlier definition.

The foreign exchange reserve growth and export growth are on a given month defined as the percentage change in the level of the variable with respect to its level a year earlier. These two variables have been multiplied by -1, and thus they actually represent export loss and reserve loss. The real exchange rate is defined with respect to the US dollar on a bilateral basis and the overvaluation measured as the percentual deviation from a deterministic time trend. The current account in relation to GDP is the ratio of a moving average of the current account over the previous twelve-months in relation to a moving average of the GDP over the same period. It has originally been multiplied also by -1, and thus it represents a current account deficit (Kaminski et al. 1998). The short-term debt is measured in relation to foreign exchange reserves and has been collected from the Bank for International Settlements.

Four of the variables (i.e., reserve growth, export growth, real exchange rate overvaluation and current account deficit) are macroeconomic fundamentals that according to early theoretical models (e.g., Krugman 1979) help predict the occurrence of a currency crisis. The fifth variable (i.e., the short-term debt in relation to reserves) is a measure of countries' vulnerability to an economic panic (Radelet and Sachs 1998), which, in the light of newer theoretical models, can be used to predict governments' reluctance to combat speculative attacks with high interest rates.

In order to create a model that is comparable with the BP model, we separate the data into training and test sets similarly as in their paper; the training set consists of observations from 1986:1–1995:4, the first test set consists of observations from 1995:5–1996:12 (denoted henceforth by test set 1), and a second test set contains data from 1997:1–12 (test set 2). In addition, this separation enables an analysis of whether the model is able to predict the Asian crises in 1997. Moreover, we transform the variables into percentile form, using the same procedure as BP, namely the transformation is based on the country's distribution for a variable over the whole in-sample period from 1970:1–1995:4. By applying this transformation, the variables become normalized to the range [1,100] and data from different countries, comparable.

2.3 Training the models

To create a predictive model of currency crisis, we first cluster the training data using the FCM technique to obtain groups of similar economic conditions. Then, we classify (label) the resulted clusters into early warning (i.e., pre-crisis) and tranquil clusters. We evaluate the goodness-of-fit of the model by calculating a set of performance measures based on the contingency matrix on the training data. We create several such models for different values of the parameter c (the number of clusters). Based on the in-sample (training) accuracy, we select the best fitted model for prediction. Figure 1 presents the main steps in creating the FCM model.

The first step in the training process consists of applying the FCM algorithm to the training data, given a specified number of clusters *c*. These clusters are then labeled based on the type of data points assigned to them. One or more clusters are labeled as early warning (EW) clusters if they contain a relatively high proportion of pre-crisis months (i.e., C24=1). Otherwise they are labeled as tranquil (TL) clusters. After the labeling is completed, we evaluate the in-sample accuracy of the model. This training process is repeated for different values of the parameter *c*. For prediction, we choose those models that have better in-sample accuracy than a benchmark model obtained by applying probit analysis (i.e., the BP model).

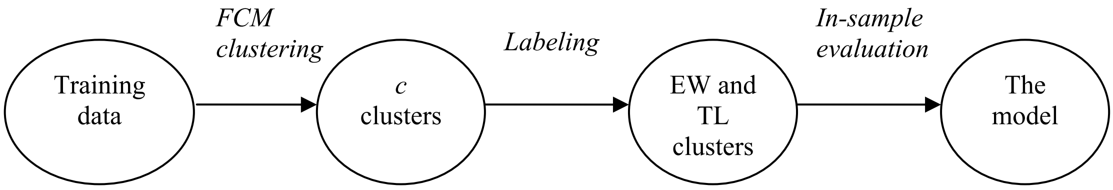

Figure 1. Obtaining the predictive FCM model from historical data

For obtaining the models, we experimented with different values of the FCM algorithm's parameters (i.e., the fuzziness indicator, *m*, and the number of clusters, *c*). Based on earlier experience on a dataset consisting of indicators for currency crises (Liu and Lindholm 2006), we set *m* to 1.5. The number of clusters *c* is not known *a priori*, even if, intuitively, we would want to have a two-cluster model: one cluster with data points representing early warning signals, and the other representing tranquil periods. However, for the reason that the dataset is quite heterogeneous, consisting of countries from different regions and observed during a long time period, we apply the FCM algorithm with different values of *c*, ranging from 2 to 2400.

Labeling the clusters into early warning and tranquil clusters

In the training process, the FCM algorithm assigns to each cluster C_i (i=1, 2,…, c) a data point x_k (k=1,2,…, n) with different degrees of membership (μ_{ik}). After the memberships values are calculated for all data points, we determine a partition of the dataset so that the data points are assigned to the clusters with the highest membership values. Thus, the partition of the dataset will consists of clusters C_i (i=1,2,…, c), where each cluster contains the points x_k that have the maximum membership values to that cluster (1).

$$C_i = \left\{ x_k \,\middle|\, \mu_{ik} = \max_{j=1,\dots,c} \{\mu_{jk}\},\ \forall k = 1,\dots,n \right\}. \quad (1)$$

Next, we label the clusters as early warning (EW) and tranquil (TL) clusters (see Figure 1). The labeling procedure is performed as follows. We define the EW clusters as being those in which the pre-crisis months represent a large proportion. Based on the proportion of pre-crisis points in the cluster, we define a variable associated to each cluster and, therefore, to the data points assigned to it, that represents the probability that the cluster is an EW cluster (2).

$$P(C_i = EW) = \frac{n_i^{PC}}{n_i}, \quad (2)$$

where $P(C_i = EW)$ is the probability that the cluster i is an early warning cluster (EW), n_i is the number of data points assigned to cluster i, and n_i^{PC}, the number of pre-crisis months assigned to cluster i.

After we calculate the probability of each cluster of being an EW cluster, we sort the clusters in the descending order of this probability; the clusters with a high probability of being EW clusters are the ones indicating most accurately the occurrence of an imminent currency crisis.

To decide which clusters to be included in the model as EW, we calculate a number of performance measures for different models that are built by including one by one the clusters in the order of their EW probabilities. The computed performance measures are based on the contingency matrix, namely the overall accuracy, recall and precision of early warning signals, and recall and precision of tranquil periods, and the false positives rate (see Witten and Frank 2005). In the Appendix we provide a brief summary of the measures used. In order to select the EW clusters with an acceptable rate of false alarms, we apply a threshold value to the false positive rate. For this operation, we use the same value of the threshold as BP, namely 10 percent. This means that, when the overall false positive rate exceeds 0.1, we do not include more EW clusters.

In-sample evaluation and model selection

The in-sample evaluation of different models obtained for different numbers of clusters is then used to choose the model that will be used for prediction. By running several experiments, we observed that the model may suffer from overfitting; when *c* approaches the number of data points in the training set, the in-sample accuracy converges to 100%, while the out-of-sample accuracy decreases. Therefore, given that the test data do not follow exactly the same patterns as the training data, the number of clusters in the model should not be too large.

To avoid model overfitting, we use two criteria to select the predictive model. The first criterion is based on the overall accuracy; we select the first model, in the ascending order of *c*, which reaches overall in-sample accuracy as good as the BP model, namely 81.6%. However, given the facts that (1) the overall accuracy is influenced by the accuracies of classifying both the pre-crisis and the tranquil months, (2) the proportion of tranquil months exceeds the one of pre-crisis (i.e., the dataset is imbalanced), and (3) our interest is in identifying as many as possible of the pre-crisis months or early warning signals, the second criterion for choosing the model for prediction is to select the first one that reaches the recall rate of early warning signals (i.e., pre-crisis months) in the probit model, namely 34.1%.

Because the FCM models are sensitive to the initial values of the memberships, and consequently, two models with the same parameter values may differ in performance quite significantly, we run 30 experiments for each value of *c* and calculate averages of all performance measures mentioned above. Thus, if the average accuracy or average recall rate over the 30 trials reaches 81.6% or 34.1%, respectively, the model is selected.

2.4 Evaluation of the models: Out-of-sample accuracy

We measure the models' performance by examining whether the test data points belong to an EW or a TL cluster. This gives the prediction (out-of-sample) accuracies of the models, which we then compare with the results of the probit model. The assignment of test points to one of the clusters in the FCM models is done by calculating the degrees of membership to all clusters and selecting the one with the highest value.

3. Experiments and results

The reproduction of the probit model

We obtained identical results of applying probit analysis to the currency crisis prediction as in BP, thus we ensured that the results of the FCM models are fully comparable with our benchmark. The performance[2] of the probit model on both training and test data sets is shown in Table 1.

Table 1. The performance measures of the probit model

Dataset	***TP***	***FP***	***TN***	***FN***	***Precision EW***	***Recall EW***	***Precision TL***	***Recall TL***	***Accuracy***	***FP rate***
Train set	126	211	1893	244	37.4 %	34.1 %	88.6 %	90.0 %	81.6 %	10.0 %
Test set 1	87	54	267	33	61.7 %	72.5 %	89.0 %	83.2 %	80.3 %	16.8 %
Test set 2	55	22	168	17	71.4 %	76.4 %	90.8 %	88.4 %	85.1 %	11.6 %

The FCM models

The FCM models are created in accordance with the procedure presented in Figure 1. Out of several models, we select for prediction two models: one with 8 clusters and another with 20 clusters. The model with 8 clusters was selected because it was the first to match the first criterion (average accuracy is higher than in the probit model, see Section 2). This model has one EW cluster and the rest are labeled as TL clusters. The model with 20 clusters was the first one that matched the second criterion (the recall rate of early warning signals exceeds the one obtained in the probit model). This model has three EW clusters. Both models have lower false positive rates than the chosen threshold of 10%.

The classification (in-sample) and prediction (out-of-sample) performance measures of the 8-cluster model are presented in Table 2. The results indicate that the 8-cluster FCM model's out-of-sample overall accuracy is slightly better for the first test set, while measured by the second test set, the probit model is better. However, the recall rates in both test sets are slightly worse in the FCM model.

Table 2. The performance measures of the FCM model with m=1.5 and c=8

Dataset	***TP***	***FP***	***TN***	***FN***	***Precision EW***	***Recall EW***	***Precision TL***	***Recall TL***	***Accuracy***	***FP rate***
Train set	114	197	1907	256	36.7%	30.8%	88.2%	90.6%	81.7%	9.4%
Test set 1	78	44	277	42	63.9%	65.0%	86.8%	86.3%	80.5%	13.7%
Test set 2	50	22	168	22	69.4%	69.4%	88.4%	88.4%	83.2%	11.6%

Table 3 presents the performance measures of the 20-cluster model. This model has a higher overall performance on the training data, but slightly worse performance on the test datasets.

[2] The performance measures and the notations used in the tables are defined in the Appendix.

Table 3. The performance measures of the FCM model with m=1.5 and c=20

Dataset	*TP*	*FP*	*TN*	*FN*	*Precision EW*	*Recall EW*	*Precision TL*	*Recall TL*	*Accuracy*	*FP rate*
Train set	157	174	1930	213	47.4%	42.4%	90.1%	91.7%	84.4%	8.3%
Test set 1	70	46	275	50	60.3%	58.3%	84.6%	85.7%	78.2%	14.3%
Test set 2	51	25	165	21	67.1%	70.8%	88.7%	86.8%	82.4%	13.2%

4. Comparison of the models

In this section, we compare the three models: probit (BP), 8-cluster FCM model, and 20-cluster FCM model. The performance measures of the models for test set 1 and test set 2 are summarized in Tables 4 and 5.

Out-of-sample performance

On the test set 1, the two FCM clustering models have similar performance with the BP probit model, but slightly worse in terms of recall of pre-crisis periods. However, the false positive rates are lower in the clustering models.

Table 4. Out-of-sample performance on test set 1

Model	*TP*	*FP*	*TN*	*FN*	*Precision EW*	*Recall EW*	*Precision TL*	*Recall TL*	*Accuracy*	*FP rate*
Probit	87	54	267	33	61.7 %	72.5 %	89.0 %	83.2 %	80.3 %	16.8 %
FCM *c*=8	78	44	277	42	63.9 %	65.0 %	86.8 %	86.3 %	80.5 %	13.7 %
FCM *c*=20	70	46	275	50	60.3 %	58.3 %	84.6 %	85.7 %	78.2 %	14.3 %

Table 5 shows similar statistics for test set 2, however the difference in the recall of the EW signals is only a few pre-crisis periods. The table shows that, in comparison to the clustering models, the probit model both predicts more precisely and recalls a higher share of the pre-crisis periods.

Table 5. Out-of- sample performance on test set 2

Model	*TP*	*FP*	*TN*	*FN*	*Precision EW*	*Recall EW*	*Precision TL*	*Recall TL*	*Accuracy*	*FP rate*
Probit	55	22	168	17	71.4 %	76.4 %	90.8 %	88.4 %	85.1 %	11.6 %
FCM *c*=8	50	22	168	22	69.4%	69.4%	88.4%	88.4%	83.2%	11.6%
FCM *c*=20	51	25	165	21	67.1%	70.8%	88.7%	86.8%	82.4%	13.2%

The ROC curves of the models

Figures 2–4 depict the ROC curves (see the Appendix) for comparing the overall performance of the probit model with the selected clustering models. They present the performance of the models regardless the threshold chosen (in our case, false positive rate of 10%). The true positive and false positive rates that are used for plotting the ROC curves of the FCM models are calculated based on different threshold values set on the probability that a cluster is labeled as an EW cluster. Figure 2 shows that generally, the 20-cluster FCM model is better than or very close to the BP model, for any chosen threshold. However, on the test datasets, generally the BP model performs better, with some variations from case to case.

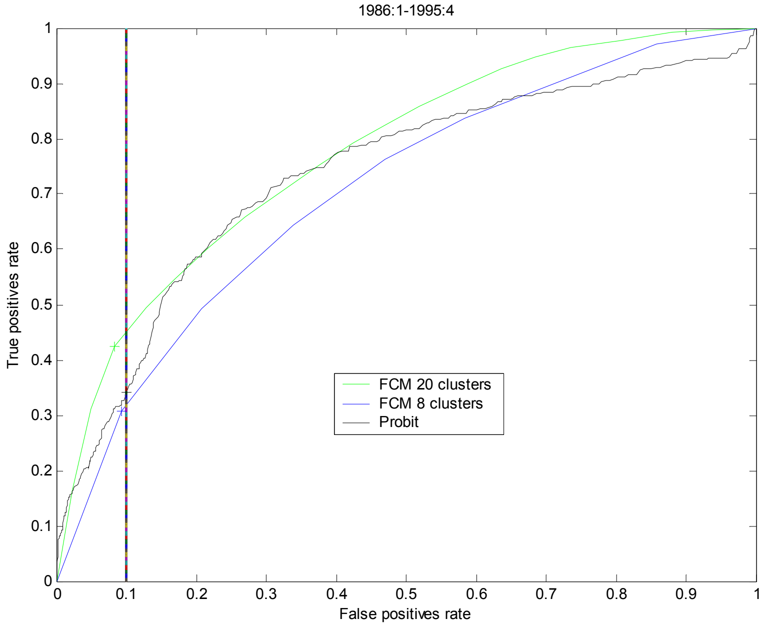

Figure 2. ROC curve on the training set

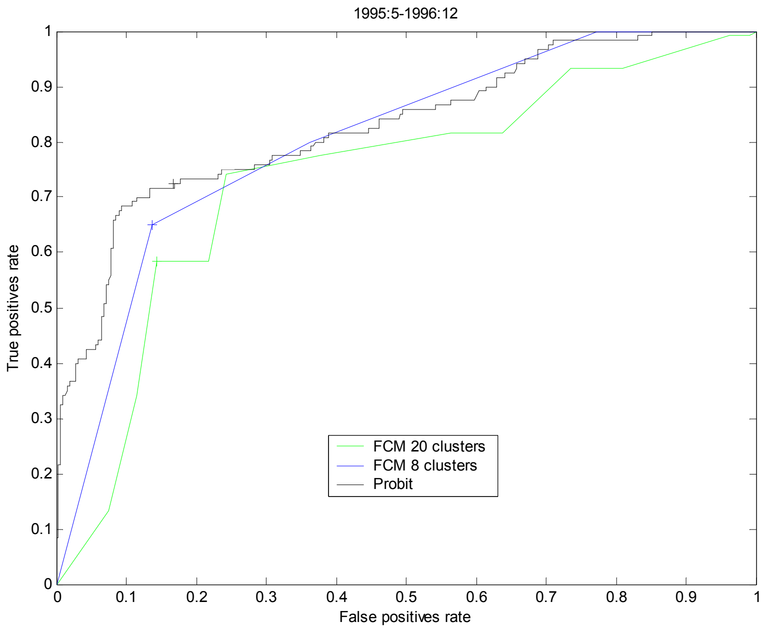

Figure 3. ROC curve on the test set 1

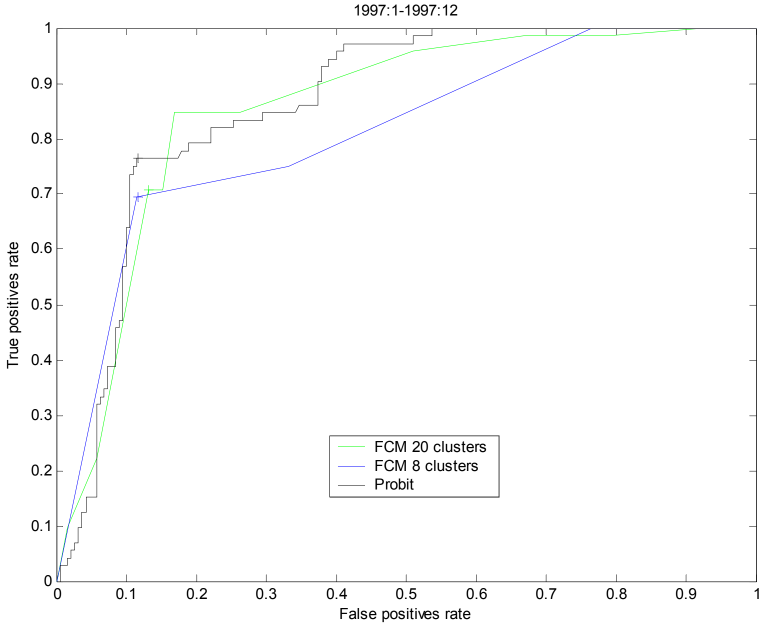

Figure 4. ROC curve on the test set 2

Description of early warning signals

In addition to the performance measures, we compare the FCM models with the probit model in terms of explanatory power of the currency crises, in particular, the economic conditions that characterize these phenomena. Table 6 presents the output of the probit analysis by indicating the extent to which each variable in the model is significantly contributing to the prediction of currency crises.

Table 6. Probit model

Parameter	***Estimate***	***Std. Error***	***Z***	***Sig.***
Intercept	-2.4749	.132	-18.715	.000
RESG	.0069	.001	5.608	.000
EXPG	.0019	.001	1.485	.138
RDEV	.0050	.001	3.964	.000
CANE	.0109	.001	8.903	.000
STDR2	.0043	.001	3.541	.000

The model shows that only export growth is not influencing significantly the probability of an early warning. All other variables are significant, and the estimates indicate the importance or strength of each variable in the model. The current account deficit, the reserve growth, the exchange rate devaluation, and the short term debt are in the descending order influencing the dependent variable, the probability of a currency crisis within 24 months. However, the model does not indicate critical levels of each variable, but determines the strength of the relationships between the explanatory variables and the predicted variable. Moreover, the positive signs of the estimates point out that higher values of the independent variables are associated with higher probability of a currency crisis.

Despite the lower out-of-sample accuracy, the clustering models are useful for characterizing the EW signals, i.e. pre-crisis periods, in terms of the macroeconomic conditions. For this purpose, the cluster representatives are used. Table 7 presents the cluster centers of the 8-cluster FCM model. In this model, only one cluster is identified as an EW, namely cluster 3. This cluster is characterized by a small exchange rate overvaluation, and high values for all other variables.

Table 7. The cluster centers of the FCM model with m=1.5 and c= 8 clusters

Cluster ID	***RESG***	***EXPG***	***RDEV***	***CANE***	***STDR2***
Cluster 1	45.62	41.72	71.75	69.77	19.37
Cluster 2	65.72	34.53	28.65	42.95	69.73
Cluster 3 (EW)	**75.06**	**63.24**	**36.53**	**78.08**	**65.33**
Cluster 4	69.63	68.08	35.94	33.09	52.56
Cluster 5	60.38	65.96	71.83	39.36	39.73
Cluster 6	26.73	64.10	46.48	40.07	36.37
Cluster 7	24.36	42.84	22.30	22.97	77.59
Cluster 8	22.24	34.40	31.69	22.96	20.67

In Table 8, the cluster centers of the 20-cluster FCM model are shown. In this model, three clusters are identified as EW clusters, namely clusters 14, 6 and 4. Cluster 14 is characterized by a medium value of exchange rate overvaluation, a significant reserve loss coupled with a high current account deficit and a medium short term debt relative to reserves. Cluster 6 is characterized by significant losses in reserves and exports, as well high exchange rate overvaluation and short term debt relative to reserves. Finally, cluster 4 is described by very high values of all variables, except the exchange rate overvaluation.

Table 8. The cluster centers for the FCM model with m=1.5 and c= 20 clusters

Cluster ID	***RESG***	***EXPG***	***RDEV***	***CANE***	***STDR2***
cluster 1	37.049	34.26	24.901	67.033	79.885
cluster 2	18.968	64.352	33.973	50.8	43.369
cluster 3	48.893	41.121	81.993	78.494	13.583
cluster 4 (EW)	**83.956**	**69.958**	**30.625**	**83.099**	**81.412**
cluster 5	74.54	31.978	20.151	35.434	73.163
cluster 6 (EW)	**72.27**	**64.467**	**80.781**	**42.859**	**62.968**
cluster 7	30.633	29.827	36.887	71.812	22.621
cluster 8	69.245	74.586	35.746	71.321	39.549
cluster 9	12.676	24.934	25.172	16.523	17.659
cluster 10	21.627	29.091	20.946	14.931	79.233
cluster 11	70.211	52.41	27.05	20.387	28.479
cluster 12	75.745	74.155	38.967	29.963	70.863
cluster 13	71.933	74.702	67.707	41.707	27.406
cluster 14 (EW)	**80.717**	**37.293**	**50.647**	**78.737**	**48.117**
cluster 15	28.012	66.546	23.536	18.723	21.997
cluster 16	51.287	39.357	45.665	43.81	65.791
cluster 17	31.242	62.507	75.477	56.256	27.166
cluster 18	35.492	67.284	63.796	27.333	37.554
cluster 19	24.86	73.886	19.761	26.383	81.027
cluster 20	34.481	33.816	64.57	27.685	21.71

Analysis of the Asian currency crises using the FCM models

The Asian countries and the corresponding currency crises are presented in Table 9. Tables 10–12 show the performance of the three models in classifying and predicting the economic conditions in these countries in terms of early warning signals and tranquil periods. None of the models correctly classified the early warning signals present in the training dataset. However, the models differ in performance when predicting the economic conditions in different countries (e.g., Indonesia and Thailand). On the other hand, the crises in Malaysia and Taiwan are predicted with 100 percent accuracy by two of the models in both test sets. The prediction of the crises in Korea and Philippines has the lowest accuracy rate.

Table 9. Crises episodes in Asian countries during 1986 - 1997

Country	***Crisis episode in training set***	***Crisis episode in test set 1***	***Crisis episode in test set 2***
Indonesia	1986	-	1997
Korea	-	-	1997
Malaysia	-	-	1997
Philippines	1986	-	1997
Taiwan	1987	-	1997
Thailand	-	-	1997
Total number of crisis episodes	3	0	6

Table 10. Correctly classified early warning signals in Asia with the probit model

Country	***Recall EW in training set***	***Recall EW in test set 1***	***Recall EW in test set 2***
Indonesia	0%	100.00%	54.55%
Korea	0%	35.71%	100.00%
Malaysia	0%	100.00%	100.00%
Philippines	0%	0.00%	0.00%
Taiwan	0%	100.00%	100.00%
Thailand	0%	100.00%	90.91%
Over all 6 countries	0% (0/29)	75.82% (69/91)	73.44% (47/64)

Table 11. Correctly classified early warning signals in Asia with the 8-cluster FCM model

Country	***Recall EW in training set***	***Recall EW in test set 1***	***Recall EW in test set 2***
Indonesia	0%	100.00%	100.00%
Korea	0%	0.00%	100.00%
Malaysia	0%	94.44%	72.73%
Philippines	0%	0.00%	0.00%
Taiwan	0%	100.00%	100.00%
Thailand	0%	100.00%	100.00%
Over all 6 countries	0% (0/29)	69.23% (63/91)	78.13% (50/64)

Table 12. Correctly classified early warning signals in Asia with the 20-cluster FCM model

Country	***Recall EW in training set***	***Recall EW in test set 1***	***Recall EW in test set 2***
Indonesia	0%	15.38%	9.09%
Korea	0%	50.00%	100.00%
Malaysia	0%	100.00%	100.00%
Philippines	0%	0.00%	45.45%
Taiwan	0%	100.00%	100.00%
Thailand	0%	55.56%	100.00%
Over all 6 countries	0% (0/29)	57.14% (52/91)	75.00% (48/64)

Using the FCM clustering models, one can analyze the economic conditions prior to a currency crisis based on the assignments of the pre-crisis months to different clusters. The graphs in Figures 5–10 depict the evolution of a country's economy in terms of the cluster to which it belongs in a given month, and the probability of being an early warning (based on the similarity with past conditions).

For each country, two line graphs are drawn that depict the cyclical fluctuations as described by the data points' assignments to different clusters. In the first line graph, the cluster identifiers are sorted in ascending order based on the clusters' probability of being EW clusters. Thus, the clusters placed higher on the vertical axis correspond to a higher probability of being EW. Consequently, months that are assigned to those clusters have a higher probability of being part of a pre-crisis period. The second line graph shows the EW probabilities assigned to the data points based on their cluster assignment. The red doted areas on both graphs represent the pre-crisis periods.

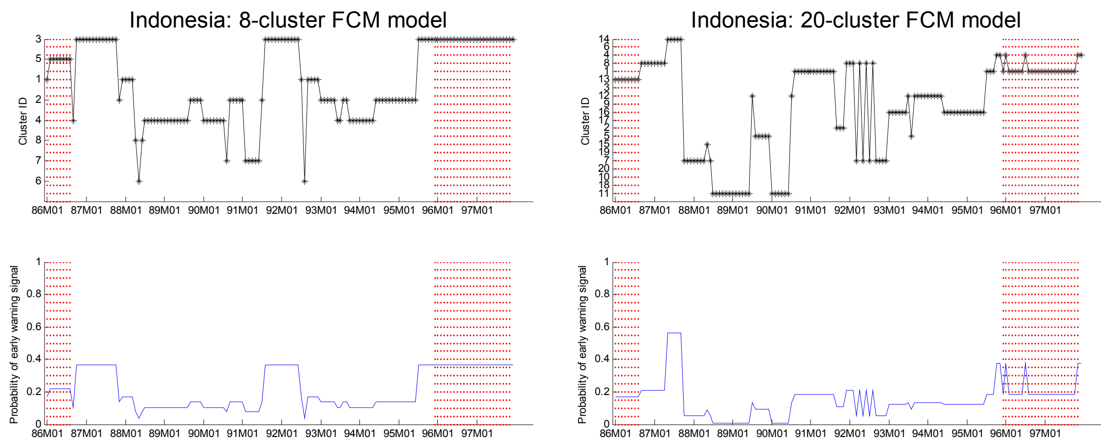

Figure 5. Prediction of Indonesia early warning signals using the 8- and 20-cluster models

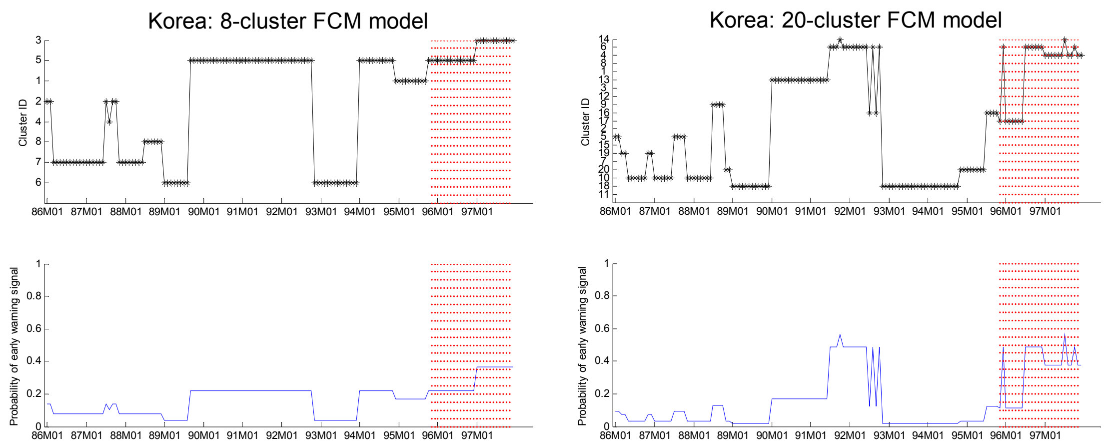

Figure 6. Prediction of Korea early warning signals using the 8- and 20-cluster models

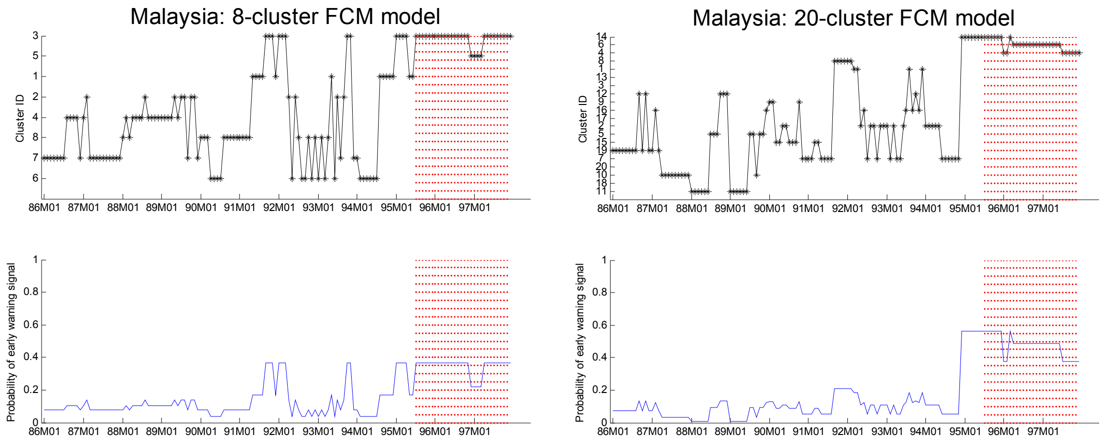

Figure 7. Prediction of Malaysia early warning signals using the 8- and 20-cluster models

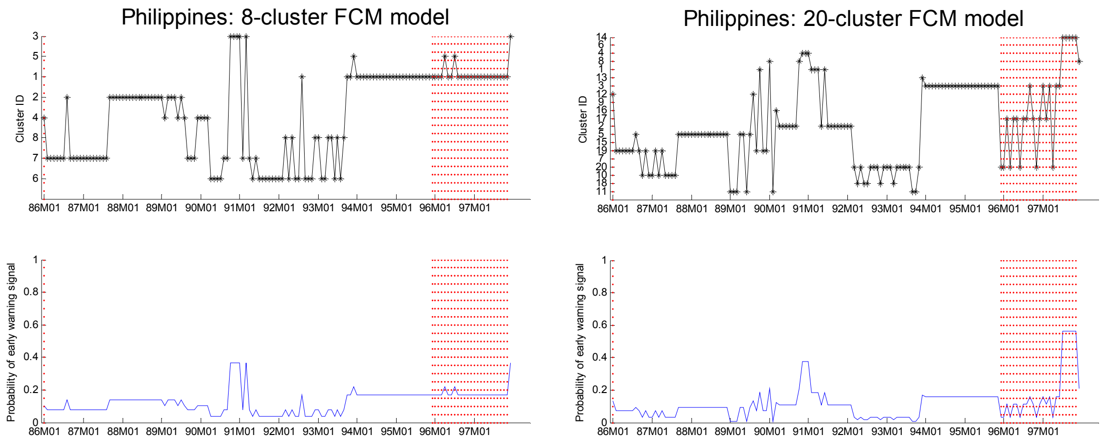

Figure 8. Prediction of Philippines early warning signals using the 8- and 20-cluster models

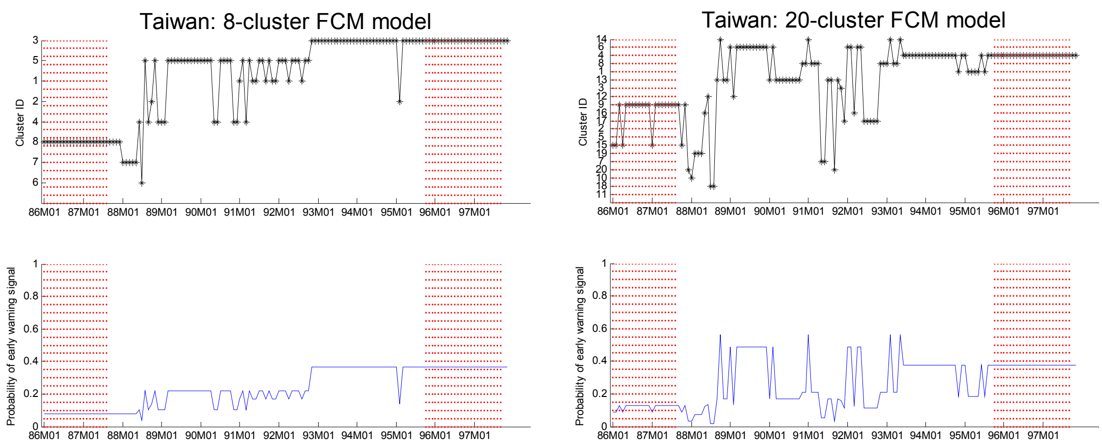

Figure 9. Prediction of Taiwan early warning signals using the 8- and 20-cluster models

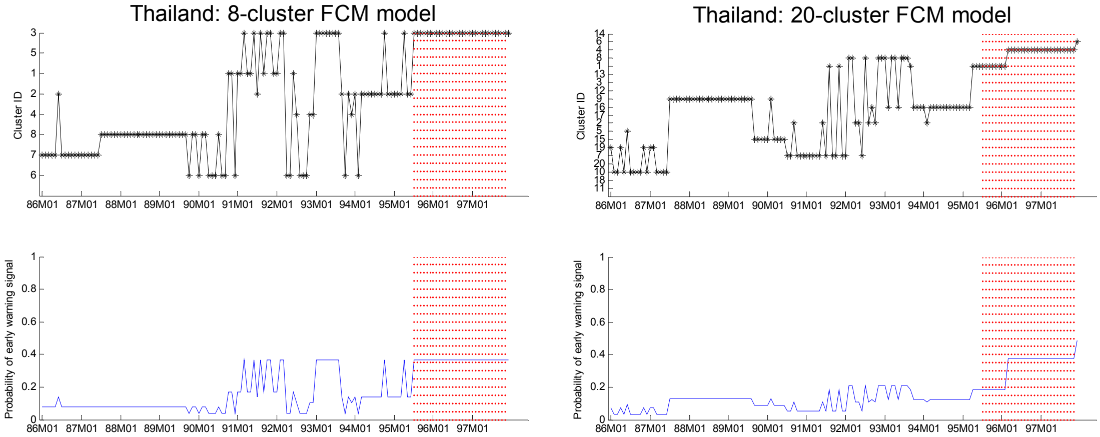

Figure 10. Prediction of Thailand early warning signals using the 8- and 20-cluster models

Based on the FCM models of currency analysis and prediction, one can learn the characteristics of the crises in different countries and time periods. For example, Figure 10 shows that in Thailand the economy prior to the currency crisis in 1997 was characterized by very high values of all variables, but exchange rate overvaluation (according to the 8-cluster model). Though the 20-cluster model has a lower prediction accuracy than the 8-cluster model, the former one shows that in the beginning of the pre-crisis period, the Thai economy was characterized by low values of the export loss and reserve loss, but quite high values of the current account deficit and short term debt relative to reserves (Cluster 1 in the 20-cluster model). Looking at the percentile values of Thailand during the pre-crisis months we observe that indeed in Thailand, in the beginning of the pre-crisis period (i.e., 1995:07–11) the export and reserve losses were relatively small, and afterwards the reserve loss started to increase considerably. Therefore, the 20-cluster model appears to be more accurate for describing the crisis in Thailand, even if the prediction accuracy is lower than the BP or the 8-cluster models.

This kind of analysis can be used to interpret the conditions prior to currency crises in all countries of interest. Therefore, despite that the prediction accuracy may vary between models, the explanatory feature of the FCM models is useful for economists and decision makers in understanding and monitoring the economic conditions that lead to financial instability.

5. Conclusions

In this paper, we built and evaluated an early warning system for predicting currency crises in Asia during 1997. The model is based on the Fuzzy C-Means technique applied on historical monthly data covering 23 developing countries during 1970–1997. First we clustered the data, and then classified the clusters into early warning and tranquil clusters. We evaluated the model by comparing it with a probit model in terms of accuracy measures on both training and test datasets. A 20-cluster model shows better in-sample accuracy than the probit model, while the out-of-sample accuracy is slightly worse. However, the explanatory feature of the model derived from assigning different pre-crisis periods to clusters with different characteristics can be regarded as an advantage of the FCM-based models. As a next step in the evaluation and development of the early warning system, we intend to analyze in more detail the type of crises that were not correctly predicted by the model. Moreover, we plan to improve the model as to its predictive performance by taking into account to a larger degree the information given by the membership values to different clusters.

Acknowledgements

We thank Andrew Berg and Catherine Pattillo for providing us with the data. Financial support from Academy of Finland (Grant Number 111692 and 127592) and Jubileumsfonden 1968 vid Åbo Akademi is gratefully acknowledged.

References

Berg, A., and Pattillo, C. (1999a) Predicting currency crises - the indicators approach and an alternative. *Journal of International Money and Finance* 18(4): 561-586.

Berg, A., and Pattillo, C. (1999b) What caused the Asian crises: An early warning system approach. *Economic Notes* 28(3); 285-334.

Bezdek, J. C. (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms.* Plenum Press: New York.

Dumitrescu, E.-I., Hurlin, C., and Candelon, B. (2010) How to evaluate an Early Warning System?, HAL Working Papers, halshs-00450050_v1.

Furman, J., and Stiglitz, J.E. (1998) Economic crises: evidence and insights from East Asia. *Brookings Papers on Economic Activity* 2: 1-135.

Kaminsky, G., Lizondo, S., and Reinhart, C. (1998) Leading indicators of currency crises. *International Monetary Fund Staff Papers* 45: 1–48.

Kohonen, T. (2001) Self-Organizing Maps. Verlag: Berlin.

Krugman, P. (1979) A model of balance-of-payments crises. *Journal of Money, Credit and Banking* 11: 311–325.

Liu, S., and Lindholm, C. (2006) Assessing the early warning signals of financial crises: A fuzzy clustering approach. *International Journal of Intelligent Systems in Accounting, Finance & Management* 14 (4): 179-202.

MacQueen, J. (1967) Some methods for classification and analysis of multivariate observations. In *Proc. of Fifth Berkeley Symposium on Math. Stat. and Prob.* 1: 281-297.

Radelet, S., and Sachs, J. (1998) The East Asian Financial Crisis: Diagnosis, Remedies, Prospects. *Brookings Papers on Economic Activity* 1: 1-90.

Sarlin, P., and Marghescu, D. (2010) Visual predictions of currency crises: A comparison of Self-Organizing Maps with probit models, submitted to *Computational Statistics & Data Analysis*.

Witten, I. H., and Frank, E. (2005) *Data Mining: Practical Machine Learning Tools and Techniques*, second edition, Morgan Kaufmann Publishers.

Appendix: Performance measures based on the contingency matrix

In a classification problem with two classes (denoted by +1 and -1), the results of the classification can be summarized in a 2x2 contingency table as follows:

		Predicted class	
		+1	-1
Actual class	+1	TP	FN
	-1	FP	TN

TP is the number of *true positives*, i.e., the number of correctly classified instances in class +1.
TN is the number of *true negatives*, i.e., the number of correctly classified instances in class -1.
FP is the number of *false positives*, i.e., the number of instances in class -1 classified incorrectly in class +1.
FN is the number of *false negatives*, i.e., the number of instances in class +1 classified incorrectly in class -1.

Based on these measures, different ratios can be calculated, for example, *recall*, *precision*, *accuracy*, and *false positive rate*. Recall and precision are calculated for each class.

Recall positives = TP/(TP+FN)
Recall negatives = TN/(TN+FP)

Precision positives = TP/(TP+FP)
Precision negatives = TN/(TN+FN)

Accuracy = (TP+TN)/(TP+TN+FP+FN)

False positive rate = FP/(FP+TN)

In addition to these measures, one can plot the *ROC curve* (Receiver Operating Characteristic curve) and based on it calculate the *area under ROC curve* (*AUC*). The ROC curve plots the recall or the true positive rate against the false positive rate at various values of a threshold based on which the classification is done. The ROC curve shows the trade-off between the benefits and costs of choosing a certain threshold. When two models are compared, the best one has a higher benefit, expressed in terms of true positive rate on the vertical axis, at the same cost, expressed in terms of false positive rate on the horizontal axis.

TURKU
CENTRE *for*
COMPUTER
SCIENCE

Joukahaisenkatu 3-5 B, 20520 Turku, Finland | www.tucs.fi

University of Turku
- Department of Information Technology
- Department of Mathematics

Åbo Akademi University
- Department of Information Technologies

Turku School of Economics
- Institute of Information Systems Sciences

ISBN 978-952-12-2423-2
ISSN 1239-1891