World Academy of Science, Engineering and Technology
International Journal of Economics and Management Engineering
Vol:8, No:2, 2014

Towards Assessment of Indicators Influence on Innovativeness of Countries' Economies: Selected Soft Computing Approaches

Marta Czyżewska, Krzysztof Pancerz, Jarosław Szkoła

***Abstract*—**The aim of this paper is to assess the influence of several indicators determining innovativeness of countries' economies by applying selected soft computing methods. Such methods enable us to identify correlations between indicators for period 2006-2010. The main attention in the paper is focused on selecting proper computer tools for solving this problem. As a tool supporting identification, the X-means clustering algorithm, the Apriori rules generation algorithm as well as Self-Organizing Feature Maps (SOMs) have been selected. The paper has rather a rudimentary character. We briefly describe usefulness of the selected approaches and indicate some challenges for further research.

***Keywords*—**Assessment of indicators, innovativeness, soft computing.

I. Background

According to last research, there are many challenges for policy makers in appropriate business innovativeness stimulation. The contemporary business propensity for innovations is influenced by many factors. There are several research activities referring the identification and assessment of the impact of individual factors on the innovativeness level of countries' economies.

Our research is based on the European Innovation Scoreboard (EIS) [1] which presents the overall innovativeness performance of the European Union 27 countries (EU27). The average performance for the latest EIS 2011 is measured using the composite indicator "Summary Innovation Index" built on data for 24 indicators, see Fig. 1.

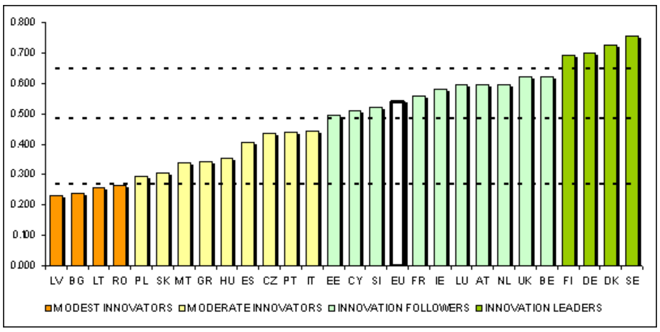

Fig. 1 Summary Innovation Index 2011 [1]

M. Czyżewska, K. Pancerz, and J. Szkoła are with University of Information Technology and Management in Rzeszow, Poland (e-mail: mczyzewska@wsiz.rzeszow.pl, kpancerz@wsiz.rzeszow.pl, jszkola@wsiz.rzeszow.pl).

According to the results of the research published in EIS 2011, the countries are divided into four groups:

1. Innovation leaders (their performance is 20% or more above the average of the EU27).
2. Innovation followers (it is less than 20% above but more than 10% that of the EU27).
3. Moderate innovators (it is less than 10% below but more than 50% below the average of the EU27).
4. Modest innovators (it is below 50% of the average of the EU27).

The indicators describing the Summary Innovation Index for European Union 27 countries we have chosen to present the analysis for this paper purposes refer to business activities area.

The indicators describe the firm activities within the innovation field, e.g.: investments in R&D, cooperation in the process of innovation introduction, intellectual property rights protection. They also describe the output of those efforts measuring the innovativeness of SME sector and economic effects of innovative activities resulting in the employment in knowledge – intensive activities, exports of high-tech products and services, new to markets and new to firm innovations, revenues from licensing and patenting. We assume these indicators are conditioning business influence on the overall level of national economies innovativeness.

The indicators described in the paper are listed below in their original meaning (numbered as they are published in The European Innovation Scoreboard 2011):

2.1.1. Business R&D expenditure as % of GDP (the indicator captures the formal creation of new knowledge within firms. It is particularly important in the science-based sectors as: pharmaceuticals, chemicals and some areas of electronics, where new knowledge is created in or in close cooperation with R&D laboratories).

2.1.2. Non-R&D innovation expenditure as % of total turnover (the indicator includes the investment in equipment and machinery and the acquisition of patents and licenses as well as measures the diffusion of new production technology and ideas. It does not include R&D expenditures).

2.2.1. SMEs innovating in-house as % of SMEs (the indicator measures the degree to which SMEs introduce new or significantly improved products or production processes that have innovated in-house).

2.2.2. Innovative SMEs co-operating with others (% of all SMEs). The indicator measures the degree to which SMEs are involved in innovation co-operation. It shows the flow of knowledge between public research institutions and private companies and also between companies.

2.3.1. PCT patent applications per billion GDP (in PPP€) – the indicator measures the number of Patent Cooperation Treaty (PCT) patent applications.

2.3.2. PCT patent applications in societal challenges per billion GDP (in PPP€) - the indicator measures PCT applications in health technology and climate change mitigation.

2.3.3. Community trademarks per billion GDP (in PPP€) - the indicator measures trademarks valid across the European Union registered with Office for Harmonization in the Internal Market in Alicante.

2.3.4. Community designs per billion GDP (in PPP€) - the indicator measures designs valid across the European Union registered with Office for Harmonization in the Internal Market.

3.1.1. SMEs introducing product or process innovations as % of SMEs (the indicator reflects the introduction of new products or services and processes in manufacturing SMEs).

3.1.2. SMEs introducing marketing/organizational innovations as % of SMEs (the indicator captures the non-technological innovation among SMEs - introduced in marketing and within their organizations).

3.2.1. Employment in knowledge-intensive activities as % of total employment (Knowledge-intensive activities are defined as those industries where at least 33% of employment has a university degree - ISCED5 or ISCED6).

3.2.2. Medium and high-tech product exports as % of total products exports (The indicator measures the technological competitiveness of the EU, i.e., the ability to commercialize the results of research and development (R&D) and innovation in the international markets. Medium and high-tech products are the source of high value added and well-paid employment).

3.2.3. Knowledge-intensive services exports as % of total services exports (The indicator measures the competitiveness of the knowledge-intensive services sector. Exports of knowledge-intensive services are measured by the sum of credits in Extended Balance of Payments Services Classification: 207, 208, 211, 212, 218, 228, 229, 245, 253, 254, 260, 263, 272, 274, 278, 279, 280 and 284).

3.2.4. Sales of new to market and new to firm innovations as % of turnover (This indicator measures the share of new or significantly improved products in total turnover and includes both products new to the firm and products which are also new to the market. The indicator thus captures both the creation of new technologies represented by the sales of new to market products and the diffusion of these technologies (new to firm products).

3.2.5. Licence and patent revenues from abroad as % of GDP (This indicator captures disembodied technology and also other types of innovations acquisition from abroad).

II. Computer Tools

Soft computing became a computer science area of study in 1990s [2]. It includes a variety of methods (e.g., neural networks, fuzzy logic, evolutionary computation, etc.) to effectively employ modes of reasoning that are approximate rather than exact. Contradictory methods, belonging to the hard computing area, are characterized by precision and certainty which bring a high computational cost. Therefore, computation, reasoning, and decision making should exploit, wherever possible, the tolerance for imprecision and uncertainty.

In our research, the main attention is focused on selecting proper computer tools implementing the soft computing paradigm for solving the problem of assessment of indicators influence on innovativeness of countries' economies taking into consideration period 2006-2010. For each indicator described in Section I, each country is described by five element time series (vectors) consisting of normalized scores determined for five consecutive years (from 2006 to 2010). A fragment of exemplary data (indicator 2.1.1) subjected to our analysis is shown in Table I.

TABLE I
A Fragment of Exemplary Data Subjected to Analysis

Country/Year	2006	2007	2008	2009	2010
BE	0.543	0.556	0.565	0.565	0.556
BG	0.039	0.047	0.052	0.056	0.116
CZ	0.405	0.384	0.362	0.371	0.405
DK	0.763	0.763	0.845	0.884	0.884
DE	0.754	0.750	0.789	0.810	0.806
...	...	...	...	...	...

Due to the vector description of items (countries), we cannot use simple methods which allow only finding correlations among individual values, i.e., for a given year. The analysis within one year leads to linear ordering of countries for a given indicator as, for example, it is presented in the European Innovation Scoreboard [1], see Fig. 2.

For data vectors, there is a need to use more sophisticated methods for finding correlations. For experiments described in this paper, we have selected the X-means clustering algorithm [3], the Apriori rule generation algorithm [4] as well as Self-Organizing Feature Maps (SOMs) [5]. The basic step is to use a clustering process. Clustering algorithms examine data to find groups (clusters) of items (vectors, objects, cases) that are similar to each other and dissimilar to the items belonging to other groups. The similarity between items is often based on a measure of the distance between them [6]. Different clustering algorithms address various facets and properties of clusters. A variety of clustering algorithms has been proposed in the literature (cf. [7]). In our investigations, we are interested in algorithms characterized by a lack of a priori knowledge about a number of clusters created during the clustering process. Among algorithms satisfying this property, we can distinguish the following ones:

- hierarchical clustering,
- X-means clustering,
- ant based clustering,
- Self-Organizing Feature Maps (SOMs).

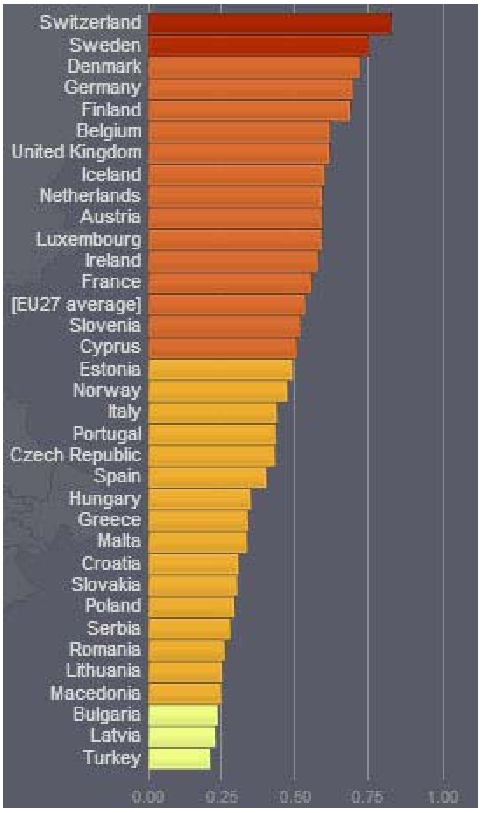

Fig. 2 Exemplary linear ordering of countries according to a selected indicator [1]

III. Experiments

In our experiments, we carried out two types of analyses of indicators described in Section I. In the first analysis, countries were clustered using the X-means algorithm, individually for each indicator. After this procedure, we have obtained a map of countries belongingness to clusters for each indicator. A fragment of clustering results is shown in Table II.

TABLE II
A Fragment of Clustering Results

Country/Indicator	$i_{2.1.1}$	$i_{2.1.2}$	...	$i_{3.2.5}$
BE	cluster4	cluster1	...	cluster2
BG	cluster2	cluster1	...	cluster1
CZ	cluster1	cluster1	...	cluster1
DK	cluster3	cluster2	...	cluster2
DE	cluster3	cluster1	...	cluster1
...	...	...	...	...

After a clustering process, the Apriori algorithm has been used two find some associations between belongingness to clusters. The Apriori algorithm generates the so-called association rules. Below, we have listed some of them (each rule is supplemented with the so-called confidence factor, the greater the confidence factor is, the more certain the rule is):

1. **If** $i_{2.3.1}$ = cluster1 and $i_{3.2.5}$ = cluster1, **then** $i_{3.2.3}$ = cluster1 (the confidence factor *conf* = 0.93). The rule can be interpreted as follows. If countries belong to clusters representing the low value of index $i_{2.3.1}$ and the low value of index $i_{3.2.5}$, then, in 93% of cases, they also belong to a cluster representing the low value of index $i_{3.2.3}$. Referring to the indicators we can state that if the country is low in both the rank of *PCT patent applications* and *Licenses and patent revenues from abroad* in 93% cases we see also a weak position of the country in *Knowledge-intensive services exports*. On this basis, we can state that PCT patenting and selling licenses abroad are crucial in gaining competitive advantage in knowledge-advanced fields on the international market.
2. **If** $i_{2.3.3}$ = cluster2, **then** $i_{3.2.5}$ = cluster1 (the confidence factor *conf* = 0.93). The rule can be interpreted in the following way. If countries belong to a cluster representing the low value of index $i_{2.3.3}$, then, in 93% of cases, they also belong to a cluster representing the low value of index $i_{3.2.5}$. Referring to the interpretation of the indicators we can assume that if countries belong to a cluster representing the low value of *Community trademarks* than they also belong to a cluster of a low value of *Licence and patent revenues from abroad*. On this basis, we can derive what countries are in poor position regarding the overall intellectual property rights protection.

In the first analysis, we have used mentioned algorithms (X-means, Apriori) implemented in a computer tool called WEKA [8]. WEKA is a collection of machine learning algorithms for data mining tasks.

In the second analysis, we have used a special kind of neural networks called Self-Organizing Feature Maps (SOMs). This approach is described more formally in [9]. We have proposed some modification of the clustering process using SOMs to improve classification results and efficiency of the learning process. As the result of a clustering process of the set of time series (vectors) corresponding to a given indicator, we obtain the so-called minimal spanning tree with respect to distances between feature vectors and centroids of clusters. Such trees enable us to made non-linearly ordered comparison of countries according to indicators considered in period 2006-2010. Below, we have listed exemplary trees.

The tree for indicator $i_{3.2.1}$ (*Employment in knowledge-intensive activities*) is shown in Fig. 3. According to the European Innovation Scoreboard 2011, the average value of the indicator is 13.5%. Countries with high shares of knowledge-intensive activities include Iceland, Ireland, Luxembourg and Switzerland. In Romania and Turkey, the share of knowledge-intensive activities is around 5%.

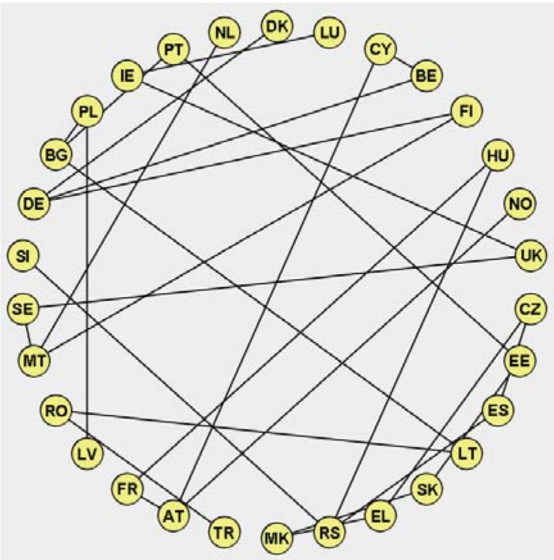

Fig. 3 A minimal spanning tree for *Employment in knowledge-intensive activities*

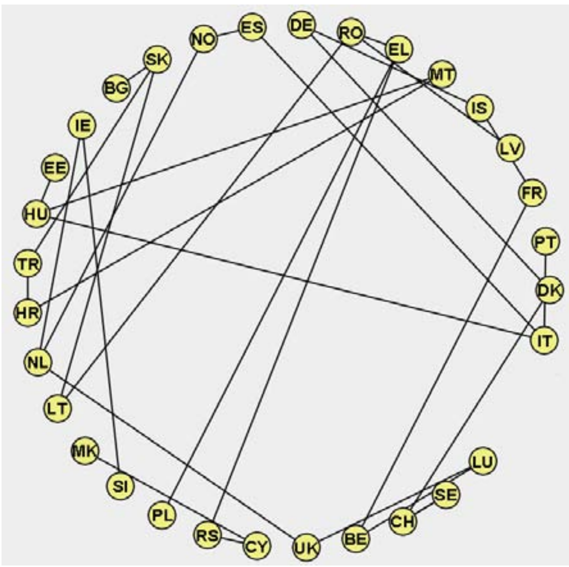

Fig. 4 A minimal spanning tree for *Business R&D expenditure as % of GDP*

The tree for indicator $i_{2.1.1}$ (*Business R&D expenditure as % of GDP*) is shown in Fig. 4. According to the EIS 2011, the highest intensity of expenditures on R&D in business sector is above 2% GDP in Denmark, Finland, Sweden and Switzerland whereas the average intensity for the EU27 is 1.25%. For 13 countries the intensity is below 0.50% GDP.

The tree for indicator $i_{3.2.2}$ (*Medium and high-tech product exports*) is shown in Fig. 5. The leaders of medium and high-tech products export are: Hungary, Malta, then Switzerland, Germany, Slovakia and Czech Republic. The low export shares are in Iceland and Norway.

The tree for indicator $i_{3.2.4}$ (*Sales of new to market and new to firm innovations*) is shown in Fig. 6. The average score of the indicator for the EU27 is 13%, but the highest values close to 25% are in Greece and Switzerland. In Norway the sales share of new or significantly improved products is below 5%.

In order to identify correlations between indicators, we need to apply some methods for comparison of topological structures of minimal spanning trees. In simple case, we can make one-to-one comparison, i.e., we compare a minimal spanning tree of one of the indicators with the one of another indicator. In the second analysis, we have used our own computer tool.

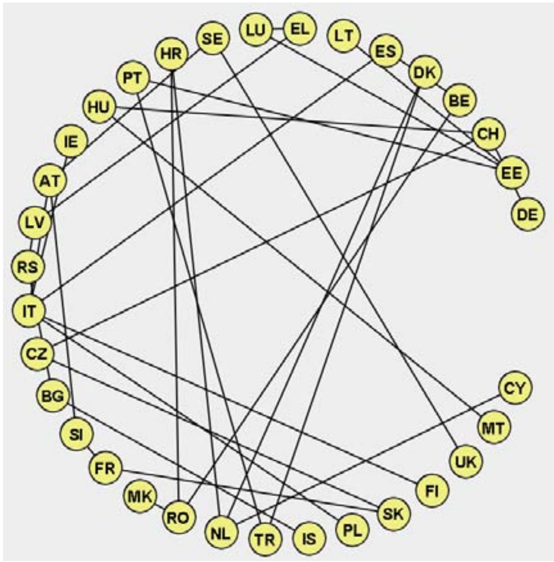

Fig. 5 A minimal spanning tree for *Medium and high-tech product exports*

World Academy of Science, Engineering and Technology
International Journal of Economics and Management Engineering
Vol:8, No:2, 2014

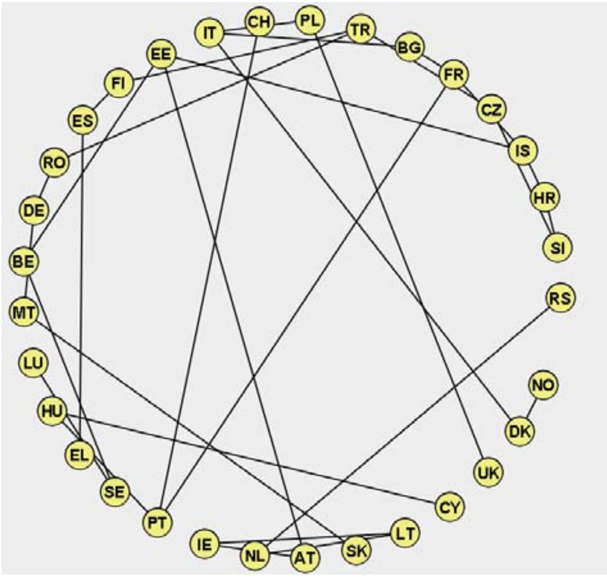

Fig. 6 A minimal spanning tree for *Sales of new to market and new to firm innovations*

IV. Conclusion

In the paper, we have shown some selected approaches based on the soft computing paradigm for assessment the influence of several indicators determining innovativeness of countries' economies. In the future, we plan to test other clustering methods, among others, that proposed in [10] based on the ant principle. It is worth noting that we need to use clustering methods without predetermined number of clusters. A fixed number of clusters can disturb the process of searching for correlations.

References

[1] European Innovation Scoreboard 2011: http://www.proinnoeurope.eu/inno-metrics/page/ius-2011.

[2] L. A. Zadeh, “Fuzzy Logic, Neural Networks, and Soft Computing,” *Communication of the ACM*, vol. 37, pp. 77–84, 1994.

[3] D. Pelleg and A. W. Moore, “X-means: Extending K-Means with Efficient Estimation of the Number of Clusters,” in *Proc. of the Seventeenth International Conference on Machine Learning*, P. Langley, Ed., Stanford, CA, USA, 2000, pp. 727–734.

[4] R. Agrawal and R. Srikant, “Fast Algorithms for Mining Association Rules in Large Databases,” in *Proc. of the 20th International Conference on Very Large Data Bases, VLDB*, Santiago, Chile, 1994, pp. 487–499.

[5] T. Kohonen, “Self-Organized Formation of Topologically Correct Feature Maps,” *Biological Cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.

[6] K. Cios, W. Pedrycz, R. Swiniarski, and L. Kurgan, *Data Mining. A Knowledge Discovery Approach*. New York: Springer, 2007.

[7] G. Gan, C. Ma, and J. Wu, *Data Clustering. Theory, Algorithms, and Applications*, SIAM, Philadelphia, ASA Alexandria, VA, 2007.

[8] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 2005.

[9] M. Czyżewska, J. Szkoła, and K. Pancerz, “Self-Organizing Feature Maps in Correlating Groups of Time Series: Experiments with Indicators Describing Entrepreneurship,” in *Proc. of the Workshop on Concurrency, Specification and Programming (CS&P 2012)*, L. Popova-Zeugmann, Ed., Berlin, Germany, 2012, vol. 1, pp. 73-78.

[10] K. Pancerz, A. Lewicki, and R. Tadeusiewicz, “Ant Based Clustering of Time Series Discrete Data - A Rough Set Approach,” in *Swarm, Evolutionary, and Memetic Computing*, ser. Lecture Notes in Computer Science, B. K. Panigrahi et al., Eds. Berlin Heidelberg: Springer-Verlag, 2011, vol. 7076, pp. 645–653.