Expert Systems with Applications 42 (2015) 562–572

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Classification of countries' progress toward a knowledge economy based on machine learning classification techniques

Mónica de la Paz-Marín [a,*], Pedro Antonio Gutiérrez [b], César Hervás-Martínez [b]

[a] Department of Management and Quantitative Methods, Loyola Andalucía University, Business Administration Faculty (ETEA), Escritor Castilla Aguayo 4, 14004 Córdoba, Spain
[b] Department of Computing and Numerical Analysis, University of Córdoba, Albert Einstein Building, 14071 Córdoba, Spain

ARTICLE INFO

Article history:
Available online 20 August 2014

Keywords:
Decision support systems
Machine learning
Knowledge economy
Hierarchical clustering
Ordinal classification

ABSTRACT

Knowledge is a key factor of competitive advantages in the current economic crisis and uncertain environment. There are a number of indicators to measure knowledge advances, however, the benefits for stakeholders and policy makers are limited because of a lack of classification models. This paper introduces an approach to classify 54 countries (in 2007–2009) according to their progress toward a knowledge economy (KE). To achieve this, the aims of this paper are twofold: first, to find clusters of countries at a similar stage of development toward KE to test if they are meaningful; hence, it will be possible to order the clusters from early KEs (last cluster) to advanced KEs (first cluster). Second, having obtained these clusters, it is possible to build various models to detect the advancement of countries toward KE from one year to another due to its classification. Then, three ordinal classifiers from the machine-learning field were compared in order to select the classifier that performs the best and to confirm the ordinal description of the clusters. Finally, an ordinal model based on the Support Vector Ordinal Regression with Implicit Constraints was selected because of its ability to classify the patterns into the clusters, confirming the appropriateness of the clusters and their ordinal nature. The proposed ordinal classifier could be used for monitoring the progress or stage of transition to KE and for analysing whether a country changes clusters, entering one that performs better or worse.

1. Introduction

Academics, policy-makers, stakeholders, consultants and the media have shown a growing interest in the relevance of knowledge creation as a key factor enabling an increase in the competitive advantages of firms and, consequently, of national economies (Von Krogh, Ichijo, & Nonaka, 2000). That is especially crucial in the uncertain, changing, ambiguous and complex environment that characterises nations today (Johannessen & Olsen, 2010). Thus, it can be said that we are merging into the so-called knowledge economy or knowledge-based economy (KE), seen as the stage that follows the industrial era, which has become almost an imperative for nations, stressing even more the role of innovation in efforts to achieve competitiveness and a sustainable economic development. The fact that governing bodies place knowledge at the core of their strategies also reveals the relevance of achieving this type of growth model.

In the last two decades, literature and research related to KE have proliferated (Aghion & Howitt, 1992; David & Foray, 2002; Drucker, 1993; Grossman & Helpman, 1991; Leydesdorff, 2006; OECD, 1996; Thurow, 1999), focusing mainly on the important role of knowledge or human capital as a source of long-term economic growth. The relevance of knowledge is clearly linked to a new growth theory, which considers knowledge (or human capital) to be an endogenous variable of economic growth. Knowledge is regarded as the basic form of capital and economic growth is driven by the accumulation of knowledge (Lucas, 1988; Romer, 1990). Other economic theories appear to examine this phenomenon: the evolutionary theory of economic change (Nelson & Winter, 1982), the national innovation systems theory (Freeman, 1987; Nelson, 1993), the knowledge gap theory (Abramovitz, 1986), the triple helix theory (Etzkowitz & Leydesdorff, 2000; Leydesdorff, 2005) or even the *N*-Duple of helices theory (Leydesdorff, 2006).

There are several real examples of the influence of knowledge on current economic growth: (1) progress in information and communications technology (ICT) that enables cheap and rapid access to knowledge and information; (2) the ever-increasing speed of scientific and technological advances; (3) global competition; and

* Corresponding author. Tel.: +34 957 2182 12; fax: +34 957 218030.
E-mail addresses: mpaz@uco.es (M. de la Paz-Marín), pagutierrez@uco.es (P.A. Gutiérrez), chervas@uco.es (C. Hervás-Martínez).

http://dx.doi.org/10.1016/j.eswa.2014.08.008
0957-4174/

(4) the new demands, tastes and customs of citizens. Based on all these, the World Bank Institute emphasises that most countries that have made rapid progress staged nationwide KE-inspired change programs (International Bank for Reconstruction, 2007).

The concept of knowledge economy (KE), although it may have its roots in Adam Smith's work, was possibly first used by Machlup (1962) and coined by Drucker (1969) and is an object of special attention in the KE report of the Organisation for Economic Co-operation and Development (OECD). According to this report, knowledge economies are those 'which are directly based on the production, distribution and use of knowledge and information' (OECD, 1996, pp. 7). The definition of the World Bank is also a widely used setting that is 'essentially an economy where knowledge is the main engine of economic growth' (Chen Derek & Dahlman, 2006, pp. 1). In these economies, emphasis is placed on intellectual capabilities rather than physical factors (Powell & Snellman, 2004) or, similarly, the share of intangible capital is greater than that of tangible capital in the overall stock of real capital (Foray, 2004).

In this context, governments must plan investments and develop strong education systems to train highly skilled workers for highly skilled jobs if they seek to achieve a knowledgeable society (Hsu, Lin, & Wei, 2008). Measurement tools, frameworks, models and methodologies help stakeholders to analyse and benchmark the capabilities of countries as knowledge-based economies. Such assessments facilitate the adoption of policies as well as the creation of national knowledge systems for holistic development.

Numerous composite indicators have been created by many organisations including the World Economic Forum (WEF), the United Nation (UN), the World Bank (WB) or the International Institute for Management Development (IMD), to name a few. These indicators have been utilised by organisations including government agencies, aid agencies and research institutions to assess the competitiveness of a nation or nations in the context of KBE. However, these indicators suffer from many shortcomings, as they can be inconsistent as they generate different ranking and scores depending on the nature and type of assessments. Moreover, in all of these studies indicators have inherited two problems. The first is the definition of the weighting scheme and the second is that they are examined at a specific time (Mimis & Georgiadis, 2013).

These indicators yield different scores and rankings depending on the nature and type of assessments, report on past performance (Al Shami, Lotfi, & Coleman, 2012), which involves many questions that must be answered subjectively (Booysen, 2002) and do not anticipate the classification of the countries where a certain KE is heading (or could head) if all variables employed in the model were known.

Classification, in general, is one of the most frequent decision-making tasks in human activity. A (supervised) classification problem occurs when an object needs to be assigned into a class based on a number of observed attributes related to that object. The most common approach to classification considers that a class variable is composed of non-ordered labels, i.e., a variable exhibits a nominal nature and the categories cannot be ordered. However, many multi-criteria classification problems involve classifying data into classes that have a natural order (ordinal problem) (Zopounidis & Doumpos, 2002). Ordinal classification techniques have broad applications in which it is natural to rank instances such as information retrieval (Chu & Keerthi, 2007; Herbrich, Graepel, & Obermayer, 1999), econometric modelling (Mathieson, 1995), credit risk (Doumpos, Kosmidou, Baourakis, & Zopounidis, 2002; Xu, Zhou, & Wang, 2009) or gen analysis (Pyon & Li, 2009), to name a few.

Consequently, in this study, the problem has been addressed by using ordinal classifiers in order to test which ordinal model performs best. A priori, the dependent variable (the cluster or class previously obtained) has an ordinal consideration as can be seen in the myriad of examples that imply a ranking of countries in socioeconomic issues: the current Rating Agencies (i.e., Moody's, Standard and Poor's or Fitch), the Global Competitiveness Index of the International Monetary Fund, the Knowledge Economy Index of the World Bank, the Innovation Union Scoreboard of the European Commission, the Human Development Index of the United Nations, the ranking of universities, and so on.

In accordance with the above, the first aim of this work focuses on obtaining homogeneous groups of countries in relation to their progress toward KE. Thus, a hierarchical clustering (an unsupervised algorithm) was applied to detect behavioural patterns. As a result, a number of clusters were set and the characteristics of each one were defined. The second aim of this paper is to build a model for the classification of 162 country–year observations (54 countries in 2007, 2008 and 2009) thanks to their assignment into clusters previously obtained. In relation to the description of the clusters, it is clear they present an ordinal nature. For that reason, three ordinal classifiers were built to assign each country–year observation to its corresponding cluster and the results were compared to evaluate which performed the best due to the ordinal nature of this socioeconomic problem (see Fig. 1).

Ordinal classification algorithms yielded very good performance, and a Support Vector ordinal model was selected for the classification of countries into one of the clusters obtained. This model could help to monitor national strategies and some key features related to knowledge creation and innovation in general terms, analysing the evolution (or lack thereof) of the country to a better or worse cluster in terms of KE progress, because the cluster has an order similar to rankings.

The remainder of this paper is organised as follows: we briefly review the relevant literature on the assessment of knowledge economy in countries and classification methodologies in Section 2. Then, the methodology applied in this study is detailed in Section 3.

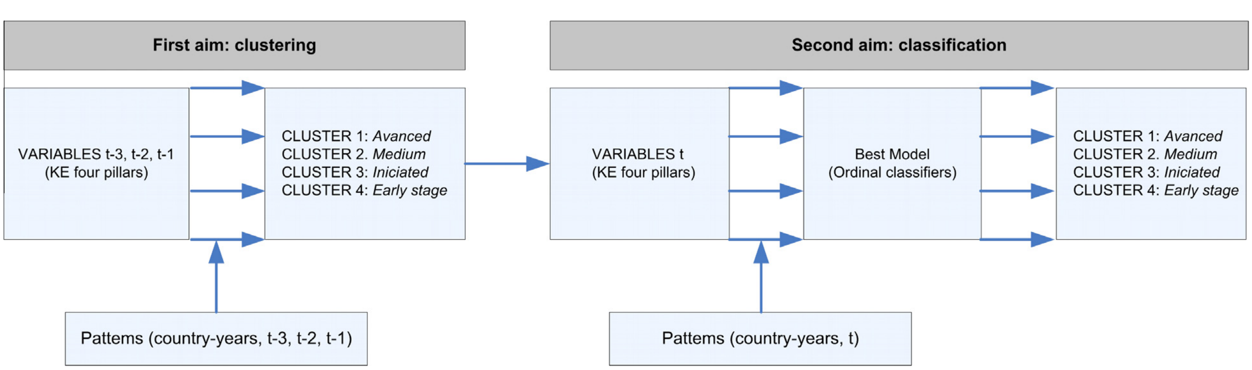

Fig. 1. Main stages and aims of the research.

In Section 4, the experimental study is carried out, including a description of the dataset and variables employed as well as a discussion of the main results. Finally, the main conclusions are drawn in Section 5.

2. Review of knowledge economy assessment methodologies

The measurement of knowledge, in general, is a core problem because its very nature implies both implicit and explicit components (Cowan, David, & Foray, 2000). In several early studies, authors have discussed the problems with the measurement of the knowledge economy (Carter, 1996; Howitt, 1996).

In relation to current KE assessment methodologies, empirical studies range from statistical techniques or econometric models to composite indicators or indices. Composite indicators are recognised as a useful tool for policy-making in setting policy priorities, in benchmarking or monitoring performance. International bodies have proposed a considerable number of them (Nardo et al., 2005), but they tend to be 'data-driven', meaning that they use data available across countries (Shapira, Youtie, Yogeesvaran, & Jaafar, 2006) and most of the data employed comes from the national accounts of the countries involved. As noted by Carter (1996), it is problematic to measure knowledge at the national level, partly because it is also difficult to measure knowledge at the firm level.

The problem of KE measurement is highly related to knowledge production, knowledge diffusion, innovation systems and the institutional environment. The theory of 'national systems of innovation' has been dominant (Lundvall, 1992; Nelson, 1993). But, while national systems of innovation can be measured in terms of sectors and institutions, the relevant contribution of the Triple Helix model (Etzkowitz & Leydesdorff, 2000; Leydesdorff & Meyer, 2006) appears as an improving alternative, measuring the extent to which innovation has become systemic instead of assuming the existence of national (or regional) systems of innovations a priori (Leydesdorff, 2012). Moreover, this model can be seen as a heuristic for researchers to take into account three spheres or layers when studying (sub)dynamics of knowledge production and innovation: industry, university and government. According to Etzkowitz and Leydesdorff (2000), in most countries these three institutional spheres overlap in their knowledge infrastructure.

On the other hand, on a micro level the rise of economies mainly driven by information and knowledge is attributed to increased intellectual capital (IC) or intangible assets. In the determination of corporate value and national economic performance, it could be asserted that IC is instrumental (Petty & Guthrie, 2000). The value created by intellectual assets is often not reflected in the financial statements of these companies; however, it must be highlighted that IC has become a prominent resource in knowledge-based economies (Kavida & Sivakoumar, 2008).

Corporate value was mostly measured by the tangible assets reflected in the book value of the companies, using a combination of both accounting and non-financial indicators. In knowledge-based economies, numerous corporate organisations have utilised these intellectual assets to their competitive advantage to create corporate value; hence, their measurement becomes crucial. Manzoni and Islam (2009) provide a leading approach in examining the field of corporate governance and performance (from suppliers through production operations and management to end users with links to other stakeholders) by means of multifaceted and multidimensional performance using Data Envelopment Analysis to integrate several non-traditional measures of performance.

The modelling and measurement of KE on a macro level follows prior efforts to measure related components, including regional human development and the building of intellectual-capital frameworks (Hanley & Malafsky, 2002). A variety of knowledge assets measurement models have also been proposed by reputable organisations, including the United Nations Economic Commission for Europe (UNECE) and the eEurope National Knowledge Assets Measurement models, among others.

Of these, the World Bank Knowledge Assessment Methodology (KAM) is one of the main knowledge economy initiatives in the world. It provides a country with a Knowledge Economy Index (KEI). Briefly, the methodology consists of a set of 148 structural and qualitative variables; notably, it can be used for benchmarking (Chen & Dahlman, 2006; World Bank, 2012). A reduced number of variables (see Table 1) serve as proxies for four relevant pillars in the development of a knowledge economy and provide a basic scoreboard:

Pillar 1: economic and institutional regime. An appropriate economic and institutional regime provides incentives for the efficient use of existing and new knowledge and for entrepreneurship.
Pillar 2: education and skills. One of the most important pillars of the knowledge-based economy is human capital. Thus, education appears as a critical element of socioeconomic life, enabling people to create and share knowledge.
Pillar 4: information and communication infrastructure. Knowledge economy is heavily based on information and communications technology (ICT). Many global changes that have taken place since the early 1970s are driven by technology. The global use of ICTs has radically modified the landscape of business, government, and social life.
Pillar 4: innovation system. This is an efficient innovation system of firms, research centres, universities, and other organisations, where knowledge diffusion and knowledge creation are typically measured by a country's ability to patent and publish scientific articles.

Regarding the methodology proposed in this study, multi-class pattern recognition, in general, has become increasingly popular. In nominal classification learning problems, no order is found among the classes. However, in ordinal classification problems,[1] target categories present an order. Because of the various synergies between data mining and operational research domains (Corne, Dhaenens, & Jourdan, 2012), a great deal of effort has been devoted to the problem of ordinal classification in operations research and computational intelligence fields, where the learning of ordinal classification models has been seen as a generalisation of some multi-criteria nominal techniques and, more recently, has led to new theoretical developments.[2] This problem appears to be different from standard regression because a distance between the labels cannot be established.

The present study employs a set of algorithms known as *threshold models* that group the majority of the proposals for ordinal regression. They focus on two main issues (Fouad & Tino, 2012): (i) how to find the optimal projection line, representing the assumed linear order of classes, onto which the input data will be projected; and (ii) how to optimally position thresholds defining the label intervals so that the margin of separation between neighbouring classes is maximised.

Arguably, the most well known group of ordered response models is based on the estimation of cumulative probabilities (Fullerton & Xu, 2012). The Proportional Odds Model (POM) is one of the first models specifically designed for ordinal regression. The method is based on extending binary logistic regression to model ordered responses. Various authors, especially McCullagh

[1] This kind of problem is also called *ordinal regression*.
[2] For a review of multicriteria models for learning ordinal data, see Sousa, Yevseyeva, Pinto da Costa, and Cardoso (2013).

Table 1
Description of variables.

Code	Variable	Description
Pillar 1: economic and institutional regime		
TNTBA	Tariff and non-tariff barriers	Score assigned to each country based on the analysis of its tariff and non-tariff barriers to trade, such as import bans and quotas as well as strict labelling and licensing requirements
REGQU	Regulatory quality	Incidence of market-unfriendly policies such as price controls or inadequate bank supervision, as well as perceptions of the burdens imposed by excessive regulation in areas such as foreign trade and business development
RULEL	Rule of law	This indicator includes several indicators, which measure the extent to which agents have confidence in and abide by the rules of society
Pillar 2: education and skills		
SCHOO	Primary enrolment (% gross)	Gross enrolment ratio is the ratio of total enrolment, regardless of age, to the population of the age group that officially corresponds to the primary level of education.
SECON	Secondary enrolment (% gross)	The ratio of total enrolment, regardless of age, to the population of the age group that officially corresponds to the secondary level of education
TERTI	Tertiary enrolment (% gross)	The ratio of total enrolment, regardless of age, to the population of the age group that officially corresponds to the tertiary level of education
Pillar 3: information and communication infrastructure		
TELEP	Telephones	Telephone mainlines, per 1000 people. Integrated services digital network channels and fixed wireless subscribers are included
FIXBI	Fixed broadband internet subscribers	The number of broadband subscribers with a digital subscriber line, cable modem or other high-speed technology, per 1000 people
INTERN	Internet users	Internet users are people with access to the worldwide network, per 1000 people
Pillar 4: innovation system		
PATEN	Patent applications	Patent grants by country of origin and patent office, per 1000 people
STJOU	Scientific journal articles	Scientific and engineering articles published by country, per 1000 people

(1980), are often credited with the idea of using the logit link for ordered responses (i.e., the POM).

In the context of Support Vector Machines (SVMs), a class of models has been developed referred to as Support Vector Ordinal Regression (SVOR). Shashua and Levin (2003) proposed a generalised formulation of SVMs applied to ordinal data based on two large-margin principles: (i) the fixed-margin principle, in which the margin of the closest pair of classes is maximised, leading to equal margins between two neighbouring classes; and (ii) the sum of margins principle, which allows for different margins and only the sum of all $Q-1$ margins is maximised (assuming there are Q ordered categories). However, the order on the $Q-1$ class thresholds was not imposed, this work being extended in the SVOR with EXplicit ordering constraints (SVOREX) formulation where the order of class thresholds is considered explicitly (Chu & Keerthi, 2007). Furthermore, the authors also presented an alternative SVOR model: SVOR with IMplicit ordering constraints (SVORIM). Both SVOREX and SVORIM are the two additional methods selected in this study for comparison.

3. Methodology

3.1. Hierarchical clustering and class description

In many real problems, decision makers must group objects into homogeneous classes. Cluster analysis allows a set of data to be divided into groups, so that the data in the same cluster are more similar to each other than data in other clusters and each data element belongs to one cluster. In order to identify an appropriate number of clusters, a variety of clustering algorithms have been proposed in recent years, but the most relevant can be grouped in *partitional* and *hierarchical* clustering (Lin, 2013; Meyer & Olteanu, 2013).

Our initial task is to employ a cluster technique to detect patterns in the selected countries with respect to their stage of transition to a KE. In the absence of class labels or knowledge about the clusters present, it is very difficult to select a criterion to judge whether one subset of features is better than another. The technique adopted in this study is hierarchical clustering (an unsupervised machine learning technique), which is used in many fields and is often preferred to non-hierarchical clustering when the number of clusters is not known or when the interest lies in the relationships between the objects, such as in taxonomical studies (Questier, Walczaka, Massarta, Bouconb, & de Jong, 2002).

Hierarchical clustering provides insight into the data by assembling all the objects into a dendrogram, so that each sub-cluster is one of its nodes, and the combinations of sub-clusters create a hierarchy—a structure that is more informative than the unstructured set of clusters in partitional clustering (Wu, Xiong, & Chen, 2009). As with other clustering techniques, one set of attributes is considered for both partitioning the data space and measuring the similarity between objects (Chen, Hsu, & Lee, 2006).

The hierarchical clustering technique was used along with the COBWEB algorithm, because of its low computational cost, with WEKA software (Hall et al., 2009). The COBWEB algorithm was developed to cluster objects in an object-attribute data set yielding a clustering dendrogram, called a classification tree, that characterises each cluster using a probabilistic description. This algorithm operates based on the so-called category utility function (CU) that measures clustering quality. The resulting classification tree of hierarchical clustering, or dendrogram, can be seen in Fig. 2. For each node, the identifier of the cluster corresponds to the number without brackets and the number in brackets is the number of country–year patterns included in the corresponding cluster. Although many clusters were derived from the COBWEB algorithm, we decided to group them in four final classes, with the aid of the hierarchical structure found by the algorithm.

The groups are referred to as classes in the figure, and they were decided by observing the country–year observations, attempting to reflect degree of transition to a KE, and also seeking to obtain representative groups with a reasonable number of elements in each class or group. A further analysis of these clusters is made based on their centroids (Table 2).

The flexibility of the resulting hierarchical clustering can be used to obtain a set of classes with an ordinal structure, which is done by grouping the different leaves according to the degree of transition to a knowledge-based economy of patterns included in each leaf. Using an economic analysis carried out by experts as a complement to the present hierarchical clustering, four clusters were selected and the profile of each group was drawn and labelled. The number of clusters was selected because four groups

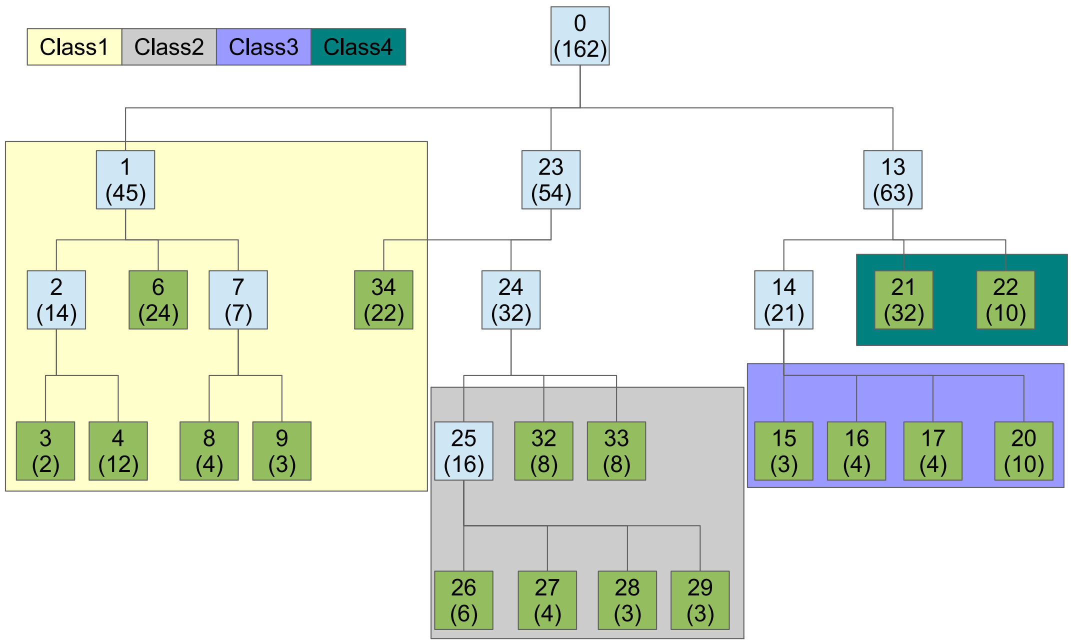

Fig. 2. Dendogram from hierarchical clustering.

achieve greater differentiation; this avoids the need for the massive information that a very fine-grained measure would require, if a greater number of clusters were chosen.

The main findings of the clustering analysis are that the countries selected fall into four groups or clusters according to the stage of their transition to a knowledge-based economy (for a detailed description of variables see Table 1 in Section 2)[3]:

Cluster 1 (*advanced KEs*). This cluster contains the country–year patterns that perform the best in the four pillars. They are economies with high-quality legal frameworks, which provide effective protection for property rights and strong support for legal regulations, intolerance of corruption, regulatory efficiency and openness to strong global commerce with an overall macroeconomic stability that minimises inherent uncertainty. They show excellent levels of patent applications and scientific publications (i.e., Switzerland, Japan and Korea) and a strong educational system, especially in the tertiary education level, a key element for a highly skilled future labour force. They are very strong in information and communication technologies (ICT) with values far above the cluster centroid in relation to the number of telephone lines, Internet users, and fixed broadband Internet subscriber variables.

Cluster 2 (*followers*). These country–year patterns follow those in cluster 1, according to the mean of each variable in this cluster with respect to the corresponding cluster centroid. They are significantly behind the first cluster, particularly at the level of patents and scientific publications. They have a legal, institutional and commercial environment with values very close to the overall average and with gross enrolment rates in primary, secondary and tertiary levels that are close to the values in the first cluster, except in primary education, where cluster 2 has a slightly higher average. Although they follow in second place, these countries are far from the patenting, scientific publications and ICT levels of the first cluster.

Cluster 3 (*moderated KEs*). This cluster groups the country–year patterns that show low performance in innovation (low number of patent applications and half the number of scientific publications), with a lower gross ratio of enrolment at each level of education (primary, secondary and tertiary). In relation to the law-rule variable that measures the extent to which agents have confidence in and abide by the rules of society, it shows very small positive values on average. The countries grouped in this cluster also show low values of ICT. In spite of this, it includes the countries with the greatest increases in broadband connection rates (i.e., Greece, Romania and Bulgaria had high average annual growth in 2007–2009).

Cluster 4 (*early KEs*). This cluster groups the country–year patterns that show a very low performance in each knowledge economy pillar. Along with Chile and Mexico, this cluster mainly represents transition economies in Europe and China. Thus, in relation to their economic incentive regimes, it must be said that they have a far from favourable framework because of the transition from a soviet-type institutional environment to a capitalist one. Hence, the special socioeconomic situation should explain the wide range of differences compared with the other clusters. With respect to the innovation pillar, the level of patenting and publishing is very low (almost non-existent in some cases) and, finally, the number of telephone lines and mobiles, as well as the number of Internet subscribers, are far from the cluster centroid.

3.2. Evaluated ordinal classifiers

For comparison purposes, different ordinal classifiers in the Machine Learning literature (see Section 2 for a brief literature review) have been included in the experimentation. The general formal framework is the following: the problem consists of predicting the label y of an input vector $\mathbf{x}$, where $x \in X \subseteq \mathbb{R}^K$ and $y \in Y = \{C_1, C_2, \ldots, C_Q\}$, i.e., $\mathbf{x}$ is in a K-dimensional input space and y is in a label space of Q different categories. The objective is

[3] In a similar way, and based on their average innovation performance, the Innovation Union Scoreboard (IUS) classifies all Member States and other European countries into four groups: innovation leaders, innovation followers, moderate innovators and modest innovators. We must say that they are commonly used terms that can be found in a number of studies related to the competitiveness of countries. An additional example can be found in the Human Development Report, 2013 (Malik, 2013): low, medium, high and very high development countries.

Table 2
Description of clusters (mean, standard deviation and country–year pattern of reference).

		Mean ± SD (country–year pattern of reference for each cluster and variable)				
CODE	VARIABLE	Cluster 1[a]	Cluster 2[b]	Cluster 3[c]	Cluster 4[d]	TOTAL
Pillar 1: economic and institutional regime						
TNTBA	Tariff and non-tariff Barriers	85.149 ± 4.575 (Japan-09)	85.400 ± 1.975 (Italy-07)	83.324 ± 5.271 (Cyprus-07)	75.638 ± 9.605 (Chile-08)	82.038 ± 7.296
REGQU	Regulatory quality	1.523 ± 0.291 (Korea Rep.-08)	1.109 ± 0.262 (Croatia-09)	0.721 ± 0.550 (Maced. FYR-08)	0.009 ± 0.539 (Tunisia-09)	0.943 ± 0.734
RULEL	Rule of law	1.640 ± 0.298 (Latvia-09)	0.912 ± 0.383 (Hungary-09)	0.300 ± 0.568 (Greece-07)	−0.394 ± 0.502 (Chile-07)	0.796 ± 0.933
Pillar 2: education and skills						
SCHOO	Primary enrollment (% gross)	102.379 ± 3.976 (Germany-09)	103.585 ± 6.220 (Hungary-09)	99.155 ± 4.520 (Bulgaria-07)	103.323 ± 7.803 (Romania-07)	101.855 ± 5.818
SECON	Secondary enrollment (% gross)	105.982 ± 11.667 (United Kingdom-07)	100.271 ± 8.094 (Lithuania-09)	92.691 ± 5.255 (Cyprus-07)	87.525 ± 5.333 (Belarus-07)	97.818 ± 11.782
TERTI	Tertiary enrollment (% gross)	66.122 ± 17.479 (Japan-07)	64.153 ± 12.894 (Czech Rep.-08)	56.251 ± 13.201 (Chile-09)	42.699 ± 19.073 (Romania-07)	58.142 ± 19.179
Pillar 3: information and communication infrastructure						
TELEP	Telephones	1619.142 ± 239.961 (Korea Rep.-07)	1587.687 ± 160.324 (Hungary-07)	1387.418 ± 207.976 (Belarus-09)	1035.833 ± 288.396 (Ukraine-07)	1425.006 ± 340.150
FIXBI	Fixed broadband internet subscribers	280.205 ± 49.153 (Estonia-07)	177.438 ± 29.787 (Latvia-08)	107.771 ± 16.798 (Cyprus-08)	41.149 ± 28.340 (Russian Fed.-09)	175.152 ± 106.165
INTERN	Internet users	768.438 ± 82.969 (Malta-09)	534.872 ± 73.289 (Spain-07)	457.583 ± 114.678 (Poland-08)	235.764 ± 94.148 (Albania-09)	542.012 ± 233.420
Pillar 4: innovation system						
PATEN	Patent applications	1.630 ± 1.224 (United Kingdom-09)	0.372 ± 0.413 (Ireland-07)	0.099 ± 0.093 (Cyprus-08)	0.075 ± 0.079 (Belarus-08)	0.782 ± 1.083
STJOU	Scientific and technical journal articles	1.864 ± 0.706 (Estonia-08)	1.104 ± 0.509 (Czech Rep.-08)	0.591 ± 0.308 (Greece-07)	0.169 ± 0.100 (Turkey-09)	1.108 ± 0.878

[a] Cluster 1 groups Australia, Austria, Belgium, Canada, Denmark, Estonia, Finland, France, Germany, Hong Kong SAR China, Iceland, Ireland-09, Japan, Korea Rep., Latvia-09, Luxembourg, Malta-09, Netherlands, New Zealand, Norway, Slovenia-09, Sweden, the United Kingdom and the United States.
[b] Cluster 2 groups Croatia-09, Cyprus-09, Czech Republic, Greece-09, Hungary, Ireland-07–08, Israel, Latvia-07–08, Lithuania, Malta-07–08, Portugal, Slovenia-07–08 and Spain.
[c] Cluster 3 groups Belarus-09, Bulgaria, Chile-09, Croatia-07–08, Cyprus-07–08, Greece-07–08, Macedonia FYR-08–09, Poland, Romania-08–09 and the Slovak Republic.
[d] Cluster 4 groups Albania, Armenia, Azerbaijan, Belarus-07–08, Bosnia and Herzegovina, Chile-07–08, China, Georgia, Macedonia FYR-07, Mexico, Moldova, Romania-07, Russian Federation, Tunisia, Turkey and Ukraine.

to find a classification rule or function $f: X \rightarrow Y$ to predict the labels of new patterns, given a training set of N points, $D = \{(\mathbf{x}_i, y_i), 1 \leqslant i \leqslant N\}$ (clusters, in our case). If the labels are ordered, so $C_1 \prec C_2 \prec \cdots \prec C_Q$, an additional constraint is found. The symbol $\prec$ denotes the ordering relationship between the different labels. Many ordinal regression measures and algorithms consider the rank of the label, that is, the position of the label in the ordinal scale, which can be expressed by the function $\mathcal{O}(\cdot)$, in such a way that $\mathcal{O}(C_q) = q$, $1 \leqslant q \leqslant Q$.

From existing ordinal regression algorithms, *threshold methods* are the most common approach. Their main goal is to model ordinal regression problems from a regression perspective in the sense that some underlying real-valued outcomes are assumed to exist, although they are unobservable. In this way, two different things are usually estimated:

– A function $f(\mathbf{x})$, which is able to predict these real-valued outcomes, uncovering the nature of the assumed underlying outcome.
– A vector $\mathbf{b} \in \mathbb{R}^{Q-1}$ (Q is the number of classes) of free parameters (thresholds) where $\mathbf{b} = (b_1, \ldots, b_{Q-1})$, which represent intervals in the range of outcomes satisfying the constraints $b_1 \leqslant b_2 \leqslant \cdots \leqslant b_{Q-1}$. Each interval is associated with a class, allowing possible different scales around the different classes.

All the ordinal classifiers selected for the experimentation follow this structure and they are briefly summarised now.

3.2.1. The Proportional Odds Model (POM)

This model is a generalisation of the logistic regression model, but considers ordinal responses. A multinomial logistic regression model is composed of $Q-1$ different linear functions, $f_q(\mathbf{x}) = \mathbf{w}_q^{\mathrm{T}}\mathbf{x}$, $1 \leqslant q \leqslant Q$, where $\mathbf{w}_q^{\mathrm{T}}$ is the vector of coefficients to be estimated from the training set. Each function estimates the probability a pattern has of belonging to each class, and the last class probability is obtained by taking into account that all probabilities must sum to one. The POM is formulated in a similar way; however, instead of modelling $Q-1$ linear responses, one linear response is estimated and the probability distribution is obtained from the value of this response. In this way, the form of the model is the following: $g^{-1}(P(y \leqslant C_q|\mathbf{x})) = b_q - \mathbf{w}_q^{\mathrm{T}}\mathbf{x}$, $1 \leqslant q \leqslant Q$, where g^{-1}: $[0, 1] \rightarrow (-\infty, +\infty)$ is a monotonic function (the inverse link function), and b_q is the threshold for class C_q. A latent variable is the main motivation of the POM, with only one linear model of the form $f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x}$. Then, a probability density function over the class labels for a given feature vector $\mathbf{x}$ is assumed. Thus, the label C_q of the training set is observed if and only if $f(\mathbf{x}) \in [b_{q-1}, b_q]$, where the function f (latent utility) and $\{b_0, b_1, \ldots, b_{Q-1}, b_Q\}$ are the parameters to be learned from the data, with $b_0 = -\infty$ and $b_Q = +\infty$. Once the model of the latent variable is defined, $f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + \varepsilon$, where ε is the random component with zero expectation, $E[\varepsilon] = 0$, which is distributed according to $F\varepsilon$, the key is to select a proper distribution assumption, $F\varepsilon$, for ε. The cumulative model is obtained by choosing the inverse distribution $F\varepsilon^{-1}$ as the inverse link function g^{-1}. The most common choice for $F\varepsilon$ is the standard logistic function, where the logit is modelled in the following manner:

$$logit(y \preceq C_q|\mathbf{x}) = \ln(odds(y \preceq C_q|\mathbf{x})) = \ln\left(\frac{P(y \preceq C_q|\mathbf{x})}{1 - P(y \succ C_q|\mathbf{x})}\right) = \mathbf{w}_q^{\mathrm{T}} \cdot \mathbf{x} + b_q \quad (1)$$

This model form allows direct estimation of the probabilities

$$P(y = C_q|\mathbf{x}) = P(y \preceq C_q|\mathbf{x}) - P(y \preceq C_{q-1}|\mathbf{x})$$

The coefficients of the model are estimated from the sample by using the design matrix of a multivariate Generalised Linear Model (GLM). There exist different methods for calculating the maximum likelihood estimate $\widetilde{\mathbf{w}}_q^{\mathrm{T}}$ (the main difficulty is introduced by the nonlinear link function; McCullagh, 1980).

3.2.2. Support Vector Ordinal Regression (SVOR).

The Support Vector Machine (SVM) (Cortes & Vapnik, 1995) may be the most common kernel learning method for classification. It can be thought of as a basis function model with a kernel computing the inner product on transformed input vectors $\phi(\mathbf{x})$. Here, $\phi(\mathbf{x})$ denotes the input pattern $\mathbf{x}$ in a very high dimensional space, which is related to $\mathbf{x}$ by a specific transformation (Cortes & Vapnik, 1995). All computations are done using only the reproducing kernel function, which is defined as:

$$k(\mathbf{x}, \mathbf{x}') = \langle \phi(\mathbf{x}) \cdot \phi(\mathbf{x}') \rangle \tag{2}$$

where $\cdot$ denotes the inner product in the high dimensional space. The idea is that linear models in this space will be nonlinear in the original one.

SVMs are based on the idea of separating the two different classes—they are first defined for two classes and then extended to the multiclass case—using a hyperplane specified by its normal vector $\mathbf{w}$ and the bias b, $\mathbf{w} \cdot \phi(\mathbf{x}) + b = 0$.

The optimal separating hyperplane is the one that maximises the distance between the hyperplane and the nearest points of both classes (called margin) for a bi-class problem, resulting in the best prediction for unseen data. Additionally, hard margins are replaced by soft margins, which allow the handling of noise and pre-labelling errors. *Slack-variables*, ξ_i, are used to relax the hard-margin constraint. The optimal separating hyperplane with maximal margin is a Quadratic Programming (QP) optimisation problem. To deal with the multiclass case, a "1-versus-1" approach is usually considered (Hsu & Lin, 2002).

SVM formulation has been adapted to the ordinal regression setting (Support Vector Machines for Ordinal Regression or SVOR), by defining a different threshold b_j for each class, and adapting the QP problem. Instead of deciding the class by the sign of the projection $\mathbf{w}^{\mathrm{T}} \cdot \mathbf{x}$, the corresponding real line is divided into consecutive intervals with a threshold vector $\mathbf{b}$, resulting in parallel hyperplanes with the same $\mathbf{w}$ and different thresholds b_j. In this paper, two different implementations for this idea are considered (Chu & Keerthi, 2007):

- SVOR with EXplicit constraints (SVOREX). The QP problem includes explicitly a set of constraints assuring the order between the thresholds, while the slacks for the jth parallel hyperplane consider patterns of class j and $j + 1$.
- SVOR with IMplicit constraints (SVORIM). The following principle is used: instead of considering only the errors from the samples of adjacent categories, samples in all the categories are allowed to contribute errors for each hyperplane. It is shown that ordinal inequalities on the thresholds are implicitly satisfied at the optimal solution.

A previous study empirically found (Chu & Keerthi, 2007) that SVOREX performed better in terms of accuracy (with a more local behaviour), and SVORIM performed better in terms of absolute deviations in number of classes or MAE (with a more global behaviour). This was justified theoretically by the loss minimised for each method. Consequently, it is not clear which method should be selected for a given dataset, and we evaluate both in the experimental section of this paper.

4. Experimental study and results

4.1. Dataset description and experimental design

We use a dataset of 54 countries to be classified according to their progress toward KE and for obtaining a classifier to perform such classification. Because of data availability of the selected variables (until 2009 in most official databases), 162 items were selected considering each country and year for 2007–2009 as a single composed item (country–year observation). The countries selected were all European countries (except Monaco, Andorra, Serbia and Montenegro),[4] adding the remaining non-European members of the OECD, as well as China, Hong Kong (due to their socioeconomic and political relevance in the international arena) and Tunisia (due to its Free Trade Zone Agreement with the EU and its geographic proximity).

To select the variables, we considered related literature, official reports and the World Bank's Knowledge Assessment Methodology (World Bank, 2012). As mentioned in Section 2, the knowledge economy index (KEI) of the World Bank is built as a simple average of four sub-indices, which represent the following four pillars of knowledge economy (see Table 1 for details).

These variables attempt to track the overall performance of the economy, which 'illustrate how well an economy is actually using knowledge for its overall economic and social development' (World Bank, 2012). Data were obtained from various sources: the World Bank Database, the SCImago Journal and Country Rank (in the case of scientific publications per country) and the Heritage Foundation for Economic Regime pillar variables. They were adjusted to per 1000 people when necessary and an imputation has been carried out for missing data (occurring in up to 18 cases). The variables and their descriptive statistics are described in Table 1.

In order to assess the ability of the models in an out-of-time dataset, the data have been split into two time periods, holding the later period for evaluation (test) of the model only, the first period being used for training. Moreover, considering longitudinal data is more realistic and useful in these kinds of socioeconomic problems (they can measure change, so we may observe the effect of the variables employed in KE performance of countries and can evaluate the effect of a specific policy by looking at the classification of the country before and after the policy was introduced). Consequently, the training information dataset consisted of observations from 54 countries in 2007–2008 with a total of 108 patterns. The test or generalisation information dataset consisted of data from 54 countries described by the same variables in the year 2009.

For the SVM classifiers (SVOREX and SVORIM), a Gaussian kernel was considered, as this is the most common choice and it usually leads to good performance. The main parameter values associated to these algorithms were two: the cost or regularization parameter (C) for balancing the importance of minimising the slack values and the width of the kernel (σ). SVM classifiers are very sensitive to these parameters and using a wrong parameter configuration can result in over-fitting or under-fitting. In this way, they were tuned by considering a nested 5-fold cross-validation process over each training set, selecting those values yielding the lowest mean absolute error. All the possible combinations were considered on a grid with the following values for the two parameters: $\{10^{-3}, 10, \ldots, 10^{3}\}$.

[4] Monaco and Andorra were not selected due to their small population; Serbia and Montenegro due to the lack of statistics because of their independence in 2006.

4.2. Classification metrics

The two evaluation metrics selected quantify the accuracy of n predicted labels for a given dataset $\{y_1^*, y_2^*, \ldots, y_n^*\}$ with respect to the true targets $\{y_1, y_2, \ldots, y_n\}$. These measures evaluate two different facets to be taken into account in ordinal classification problems (Chu & Keerthi, 2007): if the patterns are correctly classified (*CCR*) and if the classifier obtains a class as close to the true class as possible (*MAE*).

The first, the Correct Classification Rate or accuracy (*CCR*), is the fraction of correct predictions in individual samples:

$$C = \frac{1}{N}\sum_{i=1}^{N} I(y_i^* = y_i) \tag{3}$$

where $I(\cdot)$ is the zero-one loss function and n is the number of patterns in the evaluated dataset (generalisation set). A good classifier achieves the highest possible *CCR* in a given problem.

The second metric is the mean absolute error (*MAE*) that takes into account the degree of misclassification. It is the average deviation in absolute value of the classification from the true target,

$$MAE = \frac{1}{n}\sum_{i=1}^{N} \left|\mathcal{O}(y_i^*) - \mathcal{O}(y_i)\right| \tag{4}$$

where $\left|\mathcal{O}(y_i^*) - \mathcal{O}(y_i)\right|$ is the distance between the true and obtained labels.

4.3. Results and discussion

The authors of SVORIM and SVOREX methodologies (Chu & Keerthi, 2007) provide publicly available versions of their algorithms (available at http://www.gatsby.ucl.ac.uk/~chuwei/svor.htm). The POM algorithm was implemented by using the *mnrfit* function of Matlab software.

The set of experiments performed in this paper includes all the methods considered in Section 3. Table 3 compares the accuracy of the three ordinal methods in performing the classification of the stage or progress toward KE for a given country due to its assigning in the corresponding cluster. Our objective was to test if ordinal classifiers are accurate enough as was expected due to the ordinal nature of the problem, and consequently to choose the best model for classification.

The results for the two different evaluation measures (CCR_G and MAE_G; see Eqs. (3) and (4) are included in Table 3. Based on these values, the performance of each method can be analysed. The first conclusion is that high accuracies (CCR_G) are obtained in ordinal classifiers, what shows that considering the problem as an ordinal task can provide accurate information regarding the stage of transition toward KE. Namely, the SVORIM performs best (CCR_G = 87.04%). MAE_G values are also low and the lowest is in the SVORIM model. A MAE_G value of 0.130 means that the model's classifications are, on average, 0.130 categories lower or higher than the target ones with seven misclassified country–year patterns in 2009, which is the corresponding year of the generalisation set.

Table 3
CCR_G and MAE_G for the generalisation set (year 2009) of the ordinal methods evaluated.[a]

Classifier	CCR_G	MAE_G
Metrics		
POM	85.19%	0.148
SVOREX	79.63%	0.204
SVORIM	**87.04%**	**0.130**

[a] The best result is in bold face.

One could argue that the second ordinal classification stage is not necessary, as classification could be done by using the results of the clustering process. In order to better justify the necessity of the ordinal classification stage, we have performed an additional experiment where, first, all centroids are obtained based on the results of the hierarchical clustering and the intervention of the expert, and then each test pattern is assigned the class of the nearest centroid. The CCR_G and MAE_G obtained using this method are 80.33% and 0.204, respectively. These results are much worse than those obtained by SVORIM (87.04% and 0.130), especially from the point of view of MAE values. This improvement of the results justifies the second ordinal classification stage.

The ordinal classifier SVORIM is chosen because it is more accurate. However, although one could say that the differences are not excessive, it would be interesting to take additional considerations into account. First, in real situations like the one observed in this study, stakeholders and policy-makers need ordered classes both to monitor specific policies and strategies and to carry out benchmarking practices. For policy practice, the variety of indicators could be a problem since policy makers demand an aggregate index that can be easily interpreted and communicated to the general audience. Thus, the identification of ordered clusters might aid the decision maker to identify profiles or patterns to be used later on in ordinal problems for monitoring, benchmarking and offering a new perspective in a world where more and more rankings are coming into play (De Smet, Nemery, & Selvaraj, 2012). Further, once we have the groups of countries, treatment as an ordinal classification problem is advantageous since it only requires modelling of the decision boundaries, and so is likely to be less difficult than predicting a concrete index value. This methodology also allows the easy extension of the model to predict an expert analyst's assessment on an ordinal scale (in a similar way, see Mathieson, 1995). Clustering of countries is used in related issues as Innovation, Human Development, Competitiveness, and so on.

In addition, there are various arguments in favour of the selection of the ordinal classifiers:

1. By performing an ordinal classification, we project observations/patterns into one dimension, while discriminant analysis has multiple projections. To interpret the model and study the projections, it is much better and more natural to use a single projection to order the patterns.
2. The correct classification rate (*CCR*) measures the proportion of correctly classified objects on a sample of data. For the ordinal problem, this might not always be the most appropriate performance criterion. The *CCR* tacitly assumes equal misclassification costs for all incorrectly classified objects. The ordinal nature of the problem, however, implies that a one adjacent cluster misclassification is not equally costly as a two adjacent clusters error in the classification. For example, the correct classification rate for cluster 1 is 92% (23 patterns correctly classified out of 25; see Table 4), which is the cluster that groups the highest number of patterns (46.29% of the patterns in the generalisation set), and only two patterns are not correctly classified in the adjacent cluster 2.

It is worth mentioning that these clusters have produced a similar classification to that of the World Bank. This is an interesting finding for a number of reasons. First it implies that to a large extent the variables selected are fairly good proxies of key KE aspects; an inference that is relevant for policy making decisions. Second, it allows for evaluating non-parametric approaches. Stated otherwise, the proposed approach might complement the classification of countries into KE clusters providing thus new insights in KE analysis.

Table 4
Confusion matrix for the best ordinal classifier (SVORIM).

#Patterns	Actual class	Predicted class (SVORIM)[a]			
		Cluster 1	Cluster 2	Cluster 3	Cluster 4
25	Cluster 1	23	2	0	0
10	Cluster 2	0	10	0	0
7	Cluster 3	0	1	4	2
12	Cluster 4	0	0	2	10

[a] The seven misclassified country–year observations are: Belarus-09, Bosnia-09,Chile-09, Latvia-09, Poland-09, Russia-09 and Slovenia-09.

Table 5
CCR_G and MAE_G, using year 2007 for generalisation and years 2008 and 2009 for training, of the different ordinal methods evaluated.[a]

Classifier	Generalisation set = 2007	
	CCR_G	MAE_G
POM	92.59%	0.0741
SVOREX	96.30%	0.0370
SVORIM	94.44%	0.0556

[a] The best result is in bold face and the second best result in italics.

These clustering results precisely confirm the similarity between the countries belonging to each cluster and the World Bank Knowledge Assessment Methodology, thus it could validate our original purpose to formalise and automate the expert's opinions. An expert could anticipate these results, but what is new is to see less directly understandable similarities for other countries as in cluster 1 (for example, Latvia, Malta or Slovenia along with Japan, Finland, etc.). The characteristics of the various clusters are not always obvious and can be used to evaluate the current stage of KE of a country and set its profile. Another interesting fact reviled was that the size of the country does not matter when it comes to the most competitive knowledge economies in the world (Al Shami et al., 2012).

The selected SVORIM model correctly classified most of the patterns (87.04%). After analysing the errors committed in the classification test set, we observe that the errors correspond to the following seven countries in 2009: Belarus, Bosnia, Chile, Latvia, Poland, Russia and Slovenia. Latvia, Slovenia and Poland have been misclassified in all methods. Belarus, Chile, Latvia and Slovenia are classified into a lower category than in our cluster analysis, while Bosnia, Russia and Poland were classified in a higher category.

These differences may be because the quantitative data employed in our model provide information on the historical data of the selected pillars and on their fundamental structural features. They are essentially backward looking, while in social sciences and at the macro-level, there could arise unexpected situations; a model is not expected to embody all possibilities of reality. Examination of future events by experts must be a supplementary tool to test the vulnerability of a country to a variety of shocks generated both internally and externally. Thus, qualitative and judgmental aspects of analysis are unavoidable, even in the interpretation of quantitative indicators.

Continuing in this vein, these misclassifications could be explained with the help of the literature and experts' opinions. Thus, the first reason could lie in the fact that five out of the seven countries are post-socialist countries and three are new member states (NMS). Fairly little importance is given to regions and regional policy in most of the new member states. Following EU accession, it seems more difficult for EU NMS to build knowledge-based economies; the gaps in knowledge and innovation-related activities will widen, partly because of their weak technological capability (Kolodko, 2001; Orłowski, 2000). Poland is a special case in that its transition to capitalism was facilitated by the fact that a small number of private firms had been tolerated throughout the communist reign (Boettke, Coyne, & Leeson, 2008).

Latvia moves from cluster 2 to cluster 1 in 2009, but classification models still consider it to belong to cluster 2. The analysis of variables shows a significant increase in the performance of the ICT pillar along with a high gross rate of enrolment in tertiary education, which led to its classification in cluster 1. Nevertheless, a significant decrease in the number of schools is observed in previous years because of the demographic situation. This characteristic does not correspond to a country with a high educational level and a strong increase in the penetration of ICT in 2009. This is a special situation that our model considers atypical and is not decisive enough for a cluster change.

The same is observed for Slovenia: it moves from cluster 2 to cluster 1, but the best classifier still considers it to belong to the second cluster. It shows a good level of patent applications and scientific publications as well as an overall increase in the rate of enrolment at all educational levels, which may be due to structural changes, the transition to post-industrial societies, the growth of service sector and the shift towards, a knowledge-based economy, innovation and competition (Altbach, Reisberg, & Rumbley, 2009); it has also continuously improved its ICT penetration rate. This is the reason for the shift from cluster 2 to cluster 1. However, this increase has not been enough to consider it as a first cluster member in our model.

Russia belongs to cluster 4 in all the years under study, but the selected ordinal classifiers obtained that Russia would belong to cluster 3 in 2009. This could also be due to its special transition from a communist to a capitalist economy and its regulatory framework.

Finally, in order to test both the possible differences in time and the impact of the economic crisis on the performance of the models, an additional experiment has been carried out. In this experiment, the performance has been tested with the following dataset: the training set has been built with two years that could reflect the impact of the crisis and the generalisation set with the remaining data, which is not expected to reflect such impacts. Thus, we train the model using the 2008 and 2009 years and validate it with the 2007 year.

Table 5 includes the results of the experiment. As can be observed, the performance of the models has been improved (considering the generalisation set to be 2007). Thus, the model's performance is better than in the case of considering longitudinal data.

The reason for such better performance is perhaps due to the crisis itself, because the first model has been trained using a set, not including the data from 2009 (included in the generalisation set), which presents the worst impacts from the crisis and contains correspondingly unusual data.

5. Conclusions

Knowledge economy has become the major global trend in international society in the globally competitive twenty-first century. Exploring the characteristics of knowledge economy and establishing an appropriate economic paradigm to accelerate technological innovation are priority tasks for governments. In this context, main stakeholders need useful tools to support decision-making, such as the number of composite indicators and indices provided by institutions like the OECD, the World Bank or the EU.

As a first objective, this study performed a hierarchical clustering to obtain homogeneous clusters in which countries could be grouped according to the variables that are usually employed to characterise the KE. Then, to achieve our second aim, we examined three ordinal classification methods from the machine-learning field to obtain a model for the classification of countries according to their stage of transition to a knowledge economy (KE) as a monitoring and benchmarking tool for policy makers and stakeholders. From this comparison we were able to test the ordinal nature of this economic issue.

Thus, a dataset of 162 patterns formed by 54 countries was employed (European Countries and non-European OECD members plus China and Hong Kong were included as relevant economies) in the period 2007–2009 and four clusters were obtained: advanced, followers, moderate and early Knowledge Economies.

Once the clusters were obtained, we compared the accuracy of the proposed ordinal methods from machine learning. The results of this preliminary research show that ordinal classifiers yield very good performance and that the best performing method for the classification of countries was the SVORIM, with high *CCR* and low *MAE* values. Thus, this work has also confirmed how the ordering of clusters yield accurate classifications and supports the initial assumption of the ordinal nature of clusters defined by hierarchical clustering. Complementary economic and methodological reasons were provided to support this claim, as these clusters have produced a similar classification with that of the World Bank.

Monitoring is a relevant task for shedding light on the progress of a knowledge economy. The classification of these countries according to their advances toward KE can also be useful as an additional tool to assess their position compared with other countries with respect to reaching competitiveness and sustainability objectives in the long-term. The clusters obtained and the ordinal classification model selected reflect a global picture of the KE stage of countries, which could enrich and complement the judgment of stakeholders more than a single indicator score value or merely attempting to establish the KE readiness of a country through separate indicators.

The idea of proposing an alternative methodology for composite indicators or indices[5] could be motivated by: (a) indices summarise too much and in doing so, with loss of information, communicate less than the description of the homogeneous characteristics of a cluster; (b) results are sensitive to the arbitrary choices that have to be made on the method of the index's construction and the selection of the variables.

The risks of summarising too much and communicating little are obvious ones. First, it could be the merging of data on separate dimensions of knowledge economy. It could be the case that there is little change in the value of an index from one year to the next. This might be despite a worsening in one pillar. Similarly, two countries may appear to have similar levels of KE, as summarised by an index, but have different levels of the four pillars of considerable interest that would be hidden by focusing on an index's value.

Therefore, the methodology presented can be interesting as: (i) a managerial tool for supporting decision making because it allows a stage in the transition to KE to be identified and analyses the evolution of a country according to its change of cluster; (ii) a benchmarking tool to compare results with other reference countries that belongs to the same group and to learn from their best practices. The classifier could aid experts and main stakeholders for the classification of countries according to their stage of transition to the KE in future years or that of other countries, in a similar manner as scores, rankings, composite indices and other methodologies with publicly available country-data. It could even simulate scenarios that could help to anticipate the classification of the country in a case where one or another variable increases or decreases.

As appropriate data become available, the implications of our approach would appear to be worthwhile for future comparative research in cross-country performance. Additionally, it would be of interest to amplify this research with data from after 2009 because the period analysed in the present study may not fully capture the impact of the economic and financial crisis because of delay in economic impacts and data availability.

Finally, the choice of variables selected could bias both the results in clustering and in classification methods, as they are proxy measures. The selection of alternative, additional and/or a combination of variables for each pillar is proposed for further research.

[5] We employ the term 'index' for both index and composite indicator.

Acknowledgment

This work was partially subsidised by the Spanish Inter-Ministerial Commission of Science and Technology under Project TIN2011-22794, the European Regional Development fund, and the "Junta de Andalucía", Spain, under Project P11-TIC-7508.

References

Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *The Journal of Economic History, 46*, 385–406.
Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica, 60*, 323–351.
Al Shami, A., Lotfi, A., & Coleman, S. (2012). Unified knowledge based economy neural forecasting map. In *WCCI 2012 IEEE world congress on computational intelligence (IJCNN)* (pp. 2008–2015). Brisbane, Australia.
Altbach, P. G., Reisberg, L., & Rumbley, L. E. (2009). Trends in global higher education: Tracking an academic revolution. In *UNESCO 2009 world conference on higher education*. France: UNESCO.
Boettke, P. J., Coyne, C. J., & Leeson, P. T. (2008). Institutional stickiness and the new development economics. *American Journal of Economics and Sociology, 67*, 1536–7150.
Booysen, F. (2002). An overview and evaluation of composite indices of development. *Social Indicators Research, 59*, 115–151.
Carter, A. P. (1996). Measuring the performance of a knowledge-based economy. In D. Foray & B-Å. Lundvall (Eds.), *Employment and growth in the knowledge-based economy* (pp. 61–68). Paris: Organisation for Economic Cooperation and Development.
Chen Derek, H. C., & Dahlman C. J. (2006). The knowledge economy, the KAM methodology and world bank operations. *Policy research working paper* (pp. 1–24). The World Bank.
Chen, Y.-L., Hsu, W.-H., & Lee, Y.-H. (2006). TASC: Two-attribute-set clustering through decision tree construction. *European Journal of Operational Research, 174*, 930–944.
Chu, W., & Keerthi, S. S. (2007). Support vector ordinal regression. *Neural Computing, 19*, 792–815.
Corne, D., Dhaenens, C., & Jourdan, L. (2012). Synergies between operations research and data mining: The emerging use of multi-objective approaches. *European Journal of Operational Research, 221*, 469–479.
Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*, 273–297.
Cowan, R., David, P., & Foray, D. (2000). The explicit economics of knowledge codification and tacitness. *Industrial and Corporate Change, 9*, 211–253.
David, P., & Foray, D. (2002). An introduction to the economy of the knowledge society. *International Social Science Journal, 54*, 9–24.
De Smet, Y., Nemery, P., & Selvaraj, R. (2012). An exact algorithm for the multicriteria ordered clustering problem. *Omega, 40*, 861–869.
Doumpos, M., Kosmidou, K., Baourakis, G., & Zopounidis, C. (2002). Credit risk assessment using a multicriteria hierarchical discrimination approach: A comparative analysis. *European Journal of Operational Research, 138*, 392–412.
Drucker, P. (1969). *The age of discontinuity*. New York: Harper and Row.
Drucker, P. (1993). *Postcapitalist society*. New York: HarperBusiness.
Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from national systems and mode 2 to a triple helix of university–industry–government relations. *Research Policy, 29*, 109–123.
Foray, D. (2004). *The economics of knowledge*. Cambridge: Cambridge University Press.
Fouad, S., & Tino, P. (2012). Adaptive metric learning vector quantization for ordinal classification. *Neuronal Computing, 24*, 2825–2851.
Freeman, C. (1987). *Technology policy and economic performance*. London: Pinter.
Fullerton, A., & Xu, J. (2012). The proportional odds with partial proportionality constraints model for ordinal response variables. *Social Science Research, 41*, 182–198.
Grossman, G. M., & Helpman, E. (1991). *Innovation and growth in the global economy*. Cambridge, MA: The MIT Press.
Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations News, 11*, 10–18.

Hanley, S., & Malafsky, G. (2002). A guide for measuring the value of Km investments. In C. W. Holsapple (Ed.), *Handbook on knowledge management* (pp. 369–394). Heidelberg: Springer-Verlag. Vol. 2.
Herbrich, R., Graepel, T., & Obermayer, K. (1999). Support vector learning for ordinal regression. In *Ninth international conference on artificial neural networks, ICANŃ99* (pp. 97–102). Edinbourgh, UK.
Howitt, P. (1996). On some problems in measuring knowledge-based growth. In P. Howitt (Ed.), *The implications of knowledge-based growth for micro-economic policies* (pp. 9–29). Calgary: University of Calgary Press.
Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multi-class support vector machines. *IEEE Transaction on Neural Networks, 13*, 415–425.
Hsu, G. J. Y., Lin, Y.-H., & Wei, Z.-Y. (2008). Competition policy for technological innovation in an era of knowledge-based economy. *Knowledge Based Systems, 21*, 826–832.
International Bank for Reconstruction and Development (2007). Building knowledge economies: Advanced Strategies for development. World Bank Institute, Knowledge for Development Program, Available at: <http://go.worldbank.org/851HK6EUH0> (last accessed on 28 October 28 2012).
Johannessen, J. A., & Olsen, B. (2010). The future of value creation and innovations: Aspects of a theory of value creation and innovation in a global knowledge economy. *International Journal of Information Management, 30*, 502–511.
Kavida, V., & Sivakoumar, N. (2008). Corporate governance in knowledge economy – the relevance of intellectual capital. research project, Pondicherry, India. Available at SSRN: <http://ssrn.com/abstract=1152892> or http://dx.doi.org/10.2139/ssrn.1152892.
Kolodko, G. W. (2001). The "New Economy" and the old problems. Prospects for fast growth in postsocialist countries. *Zagreb International Review of Economics & Business, 2*, 71–90.
Leydesdorff, L. (2005). The triple helix model and the study of knowledge-based innovation systems. *International Journal of Contemporary Sociology, 42*, 12–27.
Leydesdorff, L. (2006). The triple helix, quadruple helix,...,and an N-tuple of Helices: Explanatory models for analysing the knowledge-based economy? *Journal of the Knowledge Economy, 3*, 25–35.
Leydesdorff, L. (2012). *The knowledge-based economy: Modeled, measured, simulated*. Boca Raton: Universal Publishers.
Leydesdorff, L., & Meyer, M. (2006). Triple helix indicators of knowledge-based innovation systems (Introduction to the Special Issue), (Vol. 35, pp. 1441–1449). Research Policy.
Lin, H.-Y. (2013). Feature selection based on cluster and variability analyses for ordinal multi-class classification problems. *Knowledge Based Systems, 37*, 94–104.
Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22*, 3–42.
Lundvall, B.-Å. (1992). *National systems of innovation*. London: Pinter.
Machlup, F. (1962). *The production and distribution of knowledge in the United States*. Princeton: Princeton University Press.
Malik, K. (2013). Human Development Report 2013. The rise of the South: Human progress in a diverse world. New York: United Nations Development Programme.
Manzoni, A., & Islam, S. M. (2009). Performance measurement in corporate governance: DEA modelling and implications for organisational behaviour and supply chain management. Berlin: Heilderberg: Physica Verlag.
Mathieson, M. J. (1995). Ordinal models for neural networks. In A. Refenes, Y. Abu-Mostafa, & J. Moody (Eds.), *Neural networks in financial engineering* (pp. 523–536). Singapore: World Scientific.
McCullagh, P. (1980). Regression models for ordinal data. *Journal of Royal Statistical Society Series B-Methodological, 42*, 109–142.
Meyer, P., & Olteanu, A.-L. (2013). Formalizing and solving the problem of clustering in MCDA. *European Journal of Operational Research, 227*, 494502.
Mimis, A., & Georgiadis, T. (2013). Economic clustering of countries in the Asia-Pacific region. *International Journal of Social Economics, 40*, 355–366.
Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovanini, E. (2005). *Handbook on constructing composite indicators: Methodology and user guide.* OECD Statistics Working Papers. Paris: OECD Publishing.
Nelson, R. R. (1993). *National innovation systems: A comparative study*. Oxford, New York: University Press.
Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Cambridge, MA: The Belknap Press of Harvard University Press.
OECD (1996). The knowledge Economy. Paris: OECD.
Orłowski, W. (2000). Knowledge economy and knowledge-based growth: Some issues in a transition economy. In A. Kukliński (Ed.), *The knowledge-based economy: The European challenges of the 21st century* (pp. 89–96). Warsaw: Oxford University Press. State Committee for Scientific Research of the Republic of Poland.
Petty, R., & Guthrie, J. (2000). Intellectual capital literature review: Measurement, reporting and management. *Journal of Intellectual Capital, 1*, 155–176.
Powell, W. W., & Snellman, K. (2004). The knowledge economy. *Annual Review of Sociology, 30*, 199–220.
Pyon, Y. S., & Li, J. (2009). Identifying gene signatures from cancer progression data using ordinal analysis. In *BIBM '09 Proceedings of the 2009 IEEE International Conference on Bioinformatics and Biomedicine* (pp. 136–141). Washington, DC: IEEE Computer Society.
Questier, F., Walczaka, B., Massarta, D. L., Bouconb, C., & de Jong, S. (2002). Feature selection for hierarchical clustering. *Analytica Chimica Acta, 466*, 311–324.
Romer, P. M. (1990). Endogenous technical change. *Journal of Political Economy, 98*, 71–102.
Shapira, P., Youtie, J., Yogeesvaran, K., & Jaafar, Z. (2006). Knowledge economy measurement: Methods, results and insights from the Malaysian knowledge content study. *Research Policy, 35*, 1522–1537.
Shashua, A., & Levin, A. (2003). Ranking with large margin principle: Two approaches. *Advances in Neural Information Processing Systems, 15*, 937–944.
Sousa, R., Yevseyeva, I., Pinto da Costa, J. F., & Cardoso, J. S. (2013). Multicriteria models for learning ordinal data: A literature review. In X.-S. Yang (Ed.), *Artificial intelligence, evolutionary computing and metaheuristics, studies in computational intelligence* (pp. 109–138). Heidelberg: Springer-Verlag.
Thurow, L. (1999). *Building wealth: The new rules for individuals, companies, and nations in a knowledge-based economy*. New York: HarperBusiness.
Von Krogh, G., Ichijo, K., & Nonaka, I. (2000). *Enabling knowledge creation: How to unlock the mystery of tacit knowledge and release the power of innovation*. New York: Oxford University Press.
World Bank (2012). Knowledge for Development (K4D), Knowledge assessment methodology 2012 (last access: November, 2012; www.worldbank.org/kam).
Wu, J., Xiong, H., & Chen, J. (2009). Towards understanding hierarchical clustering: A data distribution perspective. *Neurocomputing, 72*, 2319–2330.
Xu, X., Zhou, C., & Wang, Z. (2009). Credit scoring algorithm based on link analysis ranking with support vector machine. *Expert Systems with Applications, 36*, 2625–2632.
Zopounidis, C., & Doumpos, M. (2002). Multicriteria classification and sorting methods: A literature review. *European Journal of Operational Research, 138*, 229–246.