Technological Forecasting & Social Change 79 (2012) 1731–1745

Contents lists available at SciVerse ScienceDirect

Technological Forecasting & Social Change

Non-linear multiclassifier model based on Artificial Intelligence to predict research and development performance in European countries

Mónica de la Paz-Marín [a,*], Pilar Campoy-Muñoz [a], César Hervás-Martínez [b]

[a] Department of Management and Quantitative Methods, ETEA, Business Administration Faculty, 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:
Received 23 July 2011
Received in revised form 14 April 2012
Accepted 3 June 2012
Available online 24 June 2012

Keywords:
R&D performance
Neural networks
Evolutionary algorithms
k-means clustering
Multiclassification
European Union

ABSTRACT

This paper deals with one of the most important keys for economic growth: scientific knowledge and innovation, following the linear Research and Development (R&D) model. Patents, scientific publications and expenditure in R&D as well as the personnel involved in these activities are taken into account as proxy indicators, together with variables related to education and economy in order to classify R&D performance in 25 European Union (EU) Member States. This study classifies these countries using a set of variables which characterize them from 2005 to 2008 and analyses the most relevant ones for this classification. The Multilayer Perceptron Model (MLP) and the Product-Unit Neural Network (EPUNN) models, both trained by evolutionary algorithms (EA), were used to classify yearly country observations in clusters previously defined by employing unsupervised algorithm *k*-means clustering, obtaining four different classes of national R&D performance: low, moderate, high and innovation driven economies. Finally, our methodology is compared to other classification methods normally used in machine learning. The results show that while various methods of classification exist, our methodology obtains models with a significantly lower number of coefficients without decreasing their accuracy in predicting the classification of other European countries or in these countries in the following years.
© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Research, Development (R&D) and innovation, especially at this time of economic downturn and budget constraint, have become both the key for economic and smart growth in a knowledge-based society as well as a driving force for national competitive advantage [1]. Europe's competitiveness, our future standard of living, depends on its ability to incorporate innovation into products, services, business and social processes and models. The dynamics of modern economies have increased through globalization, greater competition and rapid technological change; consequently, R&D and innovation were first placed at the heart of the European Lisbon Strategy (2000), then re-launched in the Barcelona Council (2002) and today are found in the Europe 2020 Strategy and the European Research Area.

In general, the R&D effort is a very complex structure with multiple factors to resource allocation strategies and to convert them into innovations. In response to international competitive pressure, firms' survival and competitive advantage rely upon R&D ability, and hence, upon innovation.

However, although the design and estimation of models for efficiency (performance) assessment have made an important progress [2], the problem of defining a model to measure the R&D efficiency at the country level does not seem to have been resolved yet in spite of the various attempts in the field. Scholars and experts started investigating the topic from different, albeit complementary, perspectives. Prior studies generally test and evaluate the efficiency/performance of R&D investments mainly using statistical indicators and non-parametric techniques like Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis, at macro

* Corresponding author. Tel.: +34 957 218404; fax: +34 957 218030.
E-mail addresses: mpaz@uco.es (M. de la Paz-Marín), mpcampoy@etea.com (P. Campoy-Muñoz), chervas@uco.com (C. Hervás-Martínez).

0040-1625/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.techfore.2012.06.001

[3,4] and micro level (R&D programs, projects and firms [5][1]). It is evident that the purposes and the subjects of measurement efforts, and the context which these concepts are suited for, can be quite different [6].

The context of statistical indicators has been a field of application for composite indicators to measure national Science, Technology and Innovation (STI) performance and comparisons among countries have been a field of application for such composite indicators. Academics, international organizations (with their well-known efforts of normalization of STI indicators and concepts thanks to the OECD, the National Science Foundation or NSF and the EU) and policymakers largely adopt this approach for recommendations in this field [7]. These indicators share two similar characteristics: i) they are "composite" indicators[2] made up of various other sub-indicators, and (ii) they provide a synthesis of sub-indicators by aggregating them through using arithmetic mean (weighted or un-weighted) [8].

At European country level, a leading example is the study of R&D and innovation performance or, more widely, of the competitiveness and the state of the knowledge-based economy contained in the Innovation Union Scoreboard with respect to the countries under study. The Scoreboard examines differences in efficiency in R&D policies by assuming that efficiency is the ratio of outputs over inputs. Here, it is measured by comparing the ratio between a composite indicator score for one or more input dimensions and one or more output dimensions. The first edition of the Innovation Union Scoreboard in 2010 (based on the previous European Innovation Scoreboard) and the second edition in 2011 include innovation indicators and trend analyses for all 27 EU Member States and others [9]. Thus, the methodology captures more dimensions and variables of a country's innovation performance than in this study.

Based on their Summary Innovation Index (SII) score and the growth rate of the SII, the countries included in the analysis are also divided in four groups or clusters: average performance is measured using a composite indicator building on data for 24 indicators going from a lowest possible performance of 0 to a maximum possible performance of 1. This average in the 2010 report reflects performance in 2008/2009 and the 2011 report refers to the years 2009/2010 due to a lag in data availability. The methodology used for calculating this composite innovation indicator goes through seven steps ranging from identifying and replacing outliers to calculating the composite innovation indexes.

Finally, a wide range of evaluation methodologies and indicators for assessing the socio-economic impacts of R&D have also been presented as a toolbox document by the European Commission [2]. The report presents evaluation from a user perspective and highlights the expectations of different participants describing aspects of evaluation in four broad policy areas, i.e. R&D financing, the provision of R&D infrastructures, technology transfer and its legal framework. Eleven evaluation methodologies are reviewed with their descriptions, requirements, applications and good practice examples and outline the synergies among these evaluation methods.

With respect to the methodology employed in this paper, in the last two decades Artificial Neural Networks (ANNs) have been successfully used in economics for detecting and predicting banking crises [10], modeling financial time series [11], bankruptcy prediction [12] and identifying market structure [13]. ANNs have many advantages over conventional methods. They have the ability to analyze complex patterns with a high degree of accuracy and make no assumptions about the nature of the distribution of the data. Compared with an econometric model, it could be said that it is easier to use ANNs where a forecast needs to be obtained in a shorter period of time. One of the applications of neural networks is pattern recognition[3]. Like the DEA, a ANN is a non-parametric technique that make weaker assumptions and, from the point of view of application in R&D performance analysis, both are able to detect non-linear functional relationships that are hidden in case data and are able to apply these relationships to new data; finally, the problem can be described by a few independent variables, resulting in low dimensionality or less complexity [14].

In spite of this, to date hardly any works related to R&D performance have been based on Artificial Neural Networks (ANNs) and, basically, these are focused on R&D management activities [15]. We find that Data Mining and Cluster Analysis techniques have been used in a similar way dealing with Regional Innovation Systems [16] and studying the competitiveness of nations [17]. In [16], i.e., the authors use Regional Innovations Systems instead of countries. Thus, a wide range of 265 EU regions at NUTS 2 level are grouped into 9 clusters by employing 11 variables related to economy, education and R&D.

Regarding the neural network models available, the most popular could be the Multilayer Perceptron (MLP) due to its simple architecture yet powerful problem-solving ability, where neurons are grouped in layers and only forward connections exist [18]. However, alternatives to MLP have arisen in the last few years: Product Unit Neural Network (PUNN) models are an alternative to MLPs and are based on multiplicative neurons instead of additive ones. They correspond to a special class of feed-forward neural networks introduced by Durbin and Rumelhart [19].

In many cases, neural networks that use sigmoid unit basis functions[4] and product unit basis functions (PUNNs) are trained by using evolutionary algorithms (EAs), obtaining with this method significant advantages with respect to traditional training approaches [15]. One of the most important aspects of employing evolutionary programming (EP) algorithms as a modeling methodology is their ability to make the complexity of the network more flexible during training through the use of structural mutation operators that add and delete nodes and add and delete connections.

The aim of this work focuses on predicting the classification of R&D and innovation performance throughout 25 European countries thanks to their assignment into clusters, which will help to monitor European strategies for R&D and innovation and some key features related to the EU innovation policy, in general. First, *k*-means clustering (an unsupervised algorithm) is applied to detect behavioral patterns in 25 EU Member States between 2005 and 2008. As a result of clustering, a number of classes will be set to define the characteristics of each one. Then a multiclass classifier is built to assign each country–year observation to its corresponding cluster

[1] For a survey, see V. Ojanen and O. Vuola (2003) [20].

[2] Reviews and contribution about the methodology of the composite indicators can be found on the European Commission-Joint Research Centre, website: http://composite-indicators.jrc.ec.europa.eu/, last access March 6th, 2011.

[3] For a book of neural networks see S. Hayken (1994) [21] and R.P. Lippman (1989) [22].

[4] Also called MLP or SUNN in this paper.

متن کامل مقاله ← دریافت فوری

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات