Determination of the Effective Economic and/or Demographic Indicators in Classification of European Union Member and Candidate Countries Using Partial Least Squares Discriminant Analysis

Authors : Esra Polat

Abstract : Partial Least Squares Discriminant Analysis (PLSDA) is a statistical method for classification and consists a classical Partial Least Squares Regression (PLSR) in which the dependent variable is a categorical one expressing the class membership of each observation. PLSDA can be applied in many cases when classical discriminant analysis cannot be applied. For example, when the number of observations is low and when the number of independent variables is high. When there are missing values, PLSDA can be applied on the data that is available. Finally, it is adapted when multicollinearity between independent variables is high. The aim of this study is to determine the economic and/or demographic indicators, which are effective in grouping the 28 European Union (EU) member countries and 7 candidate countries (including potential candidates Bosnia and Herzegovina (BiH) and Kosova) by using the data set obtained from database of the World Bank for 2014. Leaving the political issues aside, the analysis is only concerned with the economic and demographic variables that have the potential influence on country's eligibility for EU entrance. Hence, in this study, both the performance of PLSDA method in classifying the countries correctly to their pre-defined groups (candidate or member) and the differences between the EU countries and candidate countries in terms of these indicators are analyzed. As a result of the PLSDA, the value of percentage correctness of 100 % indicates that overall of the 35 countries is classified correctly. Moreover, the most important variables that determine the statuses of member and candidate countries in terms of economic indicators are identified as 'external balance on goods and services (% GDP)', 'gross domestic savings (% GDP)' and 'gross national expenditure (% GDP)' that means for the 2014 economical structure of countries is the most important determinant of EU membership. Subsequently, the model validated to prove the predictive ability by using the data set for 2015. For prediction sample, %97,14 of the countries are correctly classified. An interesting result is obtained for only BiH, which is still a potential candidate for EU, predicted as a member of EU by using the indicators data set for 2015 as a prediction sample. Although BiH has made a significant transformation from a war-torn country to a semi-functional state, ethnic tensions, nationalistic rhetoric and political disagreements are still evident, which inhibit Bosnian progress towards the EU.

Keywords : classification, demographic indicators, economic indicators, European Union, partial least squares discriminant analysis

Conference Title : ICEMS 2017 : International Conference on Environmental Mathematics and Statistics
Conference Location : Tokyo, Japan
Conference Dates : May 28-29, 2017