Determination of Voting Tendencies in Turkey through Data Mining Algorithms

Ali Bayır, Informatics Department, Istanbul University, Istanbul, Turkey
Sebnem Ozdemir, Informatics Department, Istanbul University, Istanbul, Turkey
Sevinç Gülseçen, Informatics Department, Istanbul University, Istanbul, Turkey

ABSTRACT

Political elections can be defined as systems that contain political tendencies and voters’ perceptions and preferences. The outputs of those systems are formed by specific attributes of individuals such as age, gender, occupancy, educational status, socio-economic status, religious belief, etc. Those attributes can create a data set, which contains hidden information and undiscovered patterns that can be revealed by using data mining methods and techniques. The main purpose of this study is to define voting tendencies in politics by using some of data mining methods. According to that purpose, the survey results, which were prepared and applied before 2011 elections of Turkey by KONDA Research and Consultancy Company, were used as raw data set. After Preprocessing of data, models were generated via data mining algorithms, such as Gini, C4.5 Decision Tree, Naive Bayes and Random Forest. Because of increasing popularity and flexibility in analyzing process, R language and Rstudio environment were used.

KEYWORDS
Data Mining, Political Elections, R Language, RStudio, Voter’s Tendency

1. INTRODUCTION

Political tendencies and the final decisions of voters are determined and ruling administrators are decided in a result of elections. In this context, voter preferences can be seen as an activity which indirectly defines public policy as a result (Dalgıç, 2006). Nevertheless, this activity has a nature of fluidity under different complex effects. Political tendencies of individuals are shaped according to wide variety of factors including, but not limited to, family structure, beliefs, education level, household income, living area, local socio-economic level, communication opportunities, geographic location, ethnicity and communities joined (Gülmen, 1979; Doğan ve Göker, 2010; Temizel, 2012). Kalaycioğlu (1977) specified that feeling close to a specific political party or regime or sympathy to a leader and supporting a political institution also have influence on voting behavior. In Figure 1, the formation of political behavior is shown. (Saribay ve Öğün, 2006).

DOI: 10.4018/IJEA.2017010105

Copyright © 2017, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
Although it seems to be easy to take the decision for election, there have been a lot of independent factors that affect the preferences. These are: individual’s familial and political background, level of loyalty to the political party and leaders, expectations, identifications that defines the candidates’ images such as sexuality, age, marital status, level of education, level of income, profession and political identity etc. (Damlapınar ve Balcı, 2005).

Before the election, there have been public opinion polls for determining how to shape the voting behavior. Several aims for implementing the pools are: guiding and influencing the voters (Atar, 2006), determining election strategies of political parties (Kaban, 1995), making assessments while collecting scientific and objective data through the selected sample that represents the universe (Sezgin, 2003) etc. When viewed from this aspect, it is possible to say that; information and election forecasting that have been submitted to the society as a result of these public opinion polls have crucial and critical importance in terms of elections.

As a matter of fact, Kalender (1998) referred to the influences of the pools on the voters such as; tending to the popular side, supporting the weak side, being raised to a higher pitch, esteem of voters whose party’s votes seem to be low to the ballot box, recognizing the existence of small parties to vote, giving/not giving support to the parties which might not cross the election threshold etc. Realization of all of these effects in compliance with the ethical rules is connected with the neutral and correct presentation of the outcomes driven from the pools and also with ability and influence of human factor on the steps taken for collecting, analyzing and commenting the data. Data Mining Methods and Technics provide non-manipulative, efficient decision rules and results dealing with these complexities thanks to advantages revealed by data analysis and interpretation.

Data mining method has been used in various areas such as; finance, insurance, biomedical therapy, internet customer analysis, medicine, customer services, biology, astronomy (Özekes, 2003; Shi and others., 2015), astrophysics, hindering terrorism, indicating the fraud, entertainment (Özkan,
Facilitating the Egyptian Uprising: A Case Study of Facebook and Egypt’s April 6th Youth Movement
Mariam F. Alkazemi, Brian J. Bowe and Robin Blom (2013). Teaching Cases Collection (pp. 256-282).
www.igi-global.com/chapter/facilitating-egyptian-uprising/73061?camid=4v1a

The Significance of Marketing in Library and Information Science
Monday Obaidjevwe Ogbomo (2012). Library and Information Science in Developing Countries: Contemporary Issues (pp. 70-81).
www.igi-global.com/chapter/significance-marketing-library-information-science/60797?camid=4v1a