Data Mining and Analysis of the Citizens' Behavior towards the Source Separation of Waste Project by Applying C4.5 Algorithm of Decision Tree

Mohammd Reza Kohansal
Professor of Agricultural Economics, Ferdowsi University of Mashhad, Mashhad, Iran
Ali Firoozzare[7]
PhD Candidate of Agricultural Economics. Ferdowsi University of Mashhad, Mashhad, Iran
Mohammad Mehdi Baradaran
PhD Candidate of Geography and Urban Planning, International Unit, Ferdowsi University of Mashhad, Mashhad, Iran

1. Introduction
Nowadays, in the developing and some developed countries the source separation of the waste project is conducted intensively because of the shortage of keeping abnormal waste places and landfills along with the economic considerations. Hence, the recycling and conversion material organizations of municipalities have done some projects which resulted in increasing the participation of families in the separation of valuable waste. Investigation of the current state of dry waste source separation schemes in Mashhad shows that, currently, about six projects in the field of source separation of waste are executing in Mashhad. It is clear that the success or failure of these kinds of projects depends on the amount of citizens attendance and their participation in these projects; therefore, taking the importance of social, environmental, and economic aspects into consideration along with its special role in the urban development issues, it is necessary to detect the behavioral algorithms of citizens in terms of their assistance for the source waste separation project. As a result, it would be possible to determine the most important variables influencing the citizens' behavior in this field and to enhance the possibility of adopting the targeted policies in order to improve the citizens' assistance level for the source separation of the waste project as well as increasing the level of the cost efficiency of the adopted policies.

2. Theoretical Framework
The rational choice theory says that social phenomena are the results of community members' actions. Humans unlike other creatures act based on the knowledge and planning and their actions have reasons and wise evaluations. It is possible to interpret many social phenomena as in terms of the target actions of the social

---

7- Corresponding author: Email: alifiroozzarea@gmail.com

activists. Three groups of personal factors, situational factors, and the action space will influence the activists' social behavior

3. Methodology
This research presents decision-making algorithms of Mashhad citizens with regards to the source separation of the waste project by applying the new knowledge of data analysis and C4.5 algorithm of the decision tree. The data have been collected from 145 citizens through simple random sampling method in the year 2012. C4.5 algorithm is an augmented version of ID3 algorithm which is developed by Quinlan in order to form the decision tree. The main idea of this algorithm is that the small trees are better than the bigger trees. . In this study, in order to determine the socio-economic factors affecting the behavior of Mashhad citizens in the source separation of waste in the year 2011, the questionnaires were administered as a tool of data collection. In order to design a questionnaire, firstly, the literature of the source separation of the waste was reviewed by identifying some of the influential factors of citizens' behavior in the source separation of the waste. Subsequently, the scale was constructed under the close supervision of the experts.

4. Results & Discussion
Based on the research findings, it can be said that the fitted tree has a high classification accuracy as it could make a distinction between about 88 percent of the citizens who separate their waste from source and 87.5 percent of the citizens who do not participate in the source separation of waste projects. On the whole, it could classify 87 percent of the citizens properly. Besides, the results show that the factor of Satisfaction of the Project is the most important factor influencing the citizens' assistance. The other variables in order of influence, are the mother's employment status, the value estimation of the project, the awareness of the environmental, economic, and social advantages of the project along with the education level.

5- Conclusion & Suggestions
The findings can help the policymakers adopt the right policies in terms of the urban management scope which would result in the enhancement of the citizens' welfare level as the most important goal of all urban organizations. Given that one of the most important variables influencing citizens' behavior is their satisfaction from the manner of implementation, it is essential that executors of this project pay special attention to the better implementation of the projects in the city in order to attract more participation from the citizens. In addition, due to the importance of the citizens' high consideration of the separated waste, it is required that the authorities take the effects of waste separation more seriously into account with

regards to the economic benefits of the projects as well as providing the citizens with the minimal tangible economic benefits. It is also recommended to run more practical programs that can increase the working women's assistance for with the project along with offering special training for the less educated ones.

 Key words: Data mining, Decision tree, C4.5 Algorithm, Source separation of waste.

References

1. Austin, J., Hatfield, D., Grindle, A., & Bailey, J. (1993). Increasing recycling in office environments: The effects of specific informative cues. *Journal of Environmental Systems*, *6*, 355-368.
2. Corporation, T. C. (1999). *Introduction to data mining and knowledge discovery* (2nd ed.). *Michigan State University.* Two Crows Corp.
3. Daniel, T., & Larose, J. (2005). *Discovery knowledge in data: An introduction To data mining* (2nd ed.). New Jersey: Wiley.
4. Department of Statistics and Information. (2007). *Statistical yearbook of Khorasan Razavi.* Mashhad: Astan Qods Razavi. (in Persian)
5. Fitzgerald, R. W., & Lees, B. G. (1994). Assessing the classification accuracy of multisource remote sensing data. *Remote Sensing of the Environment, South-Eastern USA, 47*, 362-368.
6. Gamba, R., & Oskamp, S. (1994). Factors influencing community residents' participation in commingled curbside recycling programs. *Environment and Behavior*, *26*, 587-612.
7. Geetha, S., Ishwarya, N., & Kamaraj, N. (2010). Evolving decision tree rule based system for audio stego anomalies detection based on Hausdorff distance statistics. *Information Sciences, 180*, 2540-2559.
8. Ghaderi Asli, B. (1997). *The history of economic thought*. Tehran: Tehran University Press. (in Persian)
9. Han, J., & Kamber, M. (2006). *Data mining: Concepts and techniques* (2nd ed.). Burlington: Morgan Kaufman.
10. Hand, D. J. (1998). Review of data mining. *The American Statistician*, *52*, 112-118.
11. Hashemi, S. Z., & Javadi Yeganeh, M. (2005). The conflict between individual interest and collective interest (social choice) and its influencing factors. *Social Sciences Journal*, *26*(26), 141-174. (in Persian)
12. Kantardzic, M. (2003). *Data mining: Concepts, models, methods, and algorithms* (2nd ed.). New Jersey: John Wiley.

13. Kazemi Khybari, Kh. (2007, May). *Waste separation at source and mechanization systems for collecting recyclable materials according to the management model and executive approaches in Mashhad.* Paper presented at the National Conference on Waste Management, the Country's Municipalities and Village Administrations, the Environmental Protection Agency, Tehran, Iran.(in Persian)

14. Kusiak, A., & Smith, M. (2007). Data mining in design of products and production systems. *Annual Reviews in Control*, *31*, 147- 156.

15. Lansana, F. (1992). Distinguishing potential recyclers from non-recyclers: A basis for developing recycling strategies. *Journal of Environmental Education*, *23*, 16-23.

16. Marban, O., & Menasalvas, E. (2008). A cost model to estimate the effort of data mining projects (DMCoMo). *Information System*, *33*(1), 59-74.

17. Margai, F. (1997). Analyzing changes in waste reduction behavior in a low-income urban community following a public outreach program. *Environment and Behavior*, *29*, 769-792.

18. Nyamwange, M. (1996). Public perception of strategies for increasing participation in recycling programs. *The Journal of Environmental Education*, *27*, 19-22.

19. Outlook of Mashhad. (2009). *Mashhad municipality department of planning and development, the information and communication technology* . Retrieved from https://amar.mashhad.ir (in Persian)

20. Owens, J., Dickerson, S., & Macintosh, D. (2000). Demographic covariates of residential recycling efficiency. *Environment and Behavior*, *32*, 637-650.

21. Papoly Yazdi, M., & Vosughi, F. (2004). Solid waste recycling industry organization in Mashhad and the need to create a recycle zone. *Geography and Development*, *2*(3), 147-166. (in Persian)

22. Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning, 1*, 81-106.

23. Quinlan, J. R. (1996). Improved use of continuous attributes in C4.5. *Journal of Artificial Research*, *4*, 77-90.

24. Sadeqalhosaini, M., & Ranjbar, M. (2011, April 21). Rational expectations. *The World Economy Newspaper,* p. 2.(in Persian)

25. Sarvar, R. (2012). Social impact assessment of the project of waste separation from source in urban areas (Case study: District 21 of Tehran municipality). *Journal of Geography, 10*(33), 49-69. (in Persian)

26. Scott, D. (1999). Equal opportunity, unequal results. *Environment and Behavior*, *31*, 267-290.

27. Sugumaran, V., & Ramachandran, K. I. (2007). Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing. *Mechanical Systems and Signal Processing*, *21*, 2237-2247.

28. Sugumaran, V., Sabareesh, G. R., & Ramachandran, K. I. (2008). Fault diagnostics of roller bearing using kernel based neighborhood score multi-class support vector machine. *Expert Systems with Applications, 34*, 3090-3098.

29. Thogersen, J. (1994). A model of recycling behavior with evidence from Danish source separation programs. *International Journal of Research in Marketing*, *11*,145-163.

30. Valle, P., Reis, E., Menezes, J., & Rebelo, E. (2004). Behavioral determinants of household recycling participation: The Portuguese case. *Environment and Behavior*, *36*, 505-540.

31. Vivian, T. (2006). Evaluations of existing waste recycling methods: A Hong Kong study, *41*, 1649-1660.

32. Werner, C., & Makela, E. (1998). Motivations and behaviors that support recycling. *Journal of Environmental Psychology*, *18*, 373-386.