JEAS
Journal of Economics and Administrative Sciences
E-ISSN: 2148-1792 P- ISSN: 1302-2024
Volume 3, Supplement Issue 3
www.jeasweb.org

Neural network and regression comparison in forecasting municipal solid waste generation

Peyman Ajorloo
M.Sc. in Industrial Engineering, Arak Azad University, Arak, Iran.

ABSTRACT

This study aimed to compare neural network and regression in forecasting municipal solid waste generation. The research method used in this article is descriptive and purpose-based. In order to identify the effective factors in forecasting waste production, due to the nature of the topic, research questions and research objectives, first a list of variables of municipal waste production has been prepared by reviewing the research background. Then, by using Delphi technique and survey of academic experts as well as the managers active in the waste industry, effective factors on municipal waste production has been identified. Also by studying the research literature on methods of multivariate regression and neural network analysis, comprehensive and complete information was collected and data collection tools of standard questionnaire related to survey to identify effective factors are performed by using the Delphi technique. The team of experts in this study are related to waste specialists and academics as well as university professors who have been used to survey factors affecting municipal waste production, which are a number of 15 people. After collecting the data using Delphi technique, the required variables were identified, and then were analyzed by using multivariate regression and artificial neural network method, which MATLAB, EXCEL, and SPSS software have been used for this purpose. According to the results, the artificial neural network model has better capability to forecast waste production and considering MAPE and MSE of both models, the presented model using neural network in this study has better performance in forecasting waste production than linear regression.
Keywords: Waste, Forecast, Artificial Neural Networks

Introduction

Municipal waste is the natural result of human activities. Collection, transportation and disposal of municipal solid waste is one of the urban problems worldwide. Mass waste generation in cities has created health, environmental, socio-economic, transport and traffic problems. The location of landfills and waste recycling industries is one of the major and complex issues in cities. If the proper management system is not used, these materials will cause a great deal of environmental pollution and endanger human health (Chang, 2010). It is very difficult to create such a system because of the complexity and the very heterogeneous nature of waste generation. One of the most important factors in the proper operation of such a complex system is recognizing the quantity of the generated waste in it, as the quantity of production is effective in the volume of investment for machinery, on-site containers, transmission stations and disposal capacity, proper organization is effective. (Abdoli et al., 2007) Estimating the amount of the produced waste due to high production volatility and various factors affecting it is one of the most difficult tasks in waste management. Since municipal waste management systems should be designed for a long period of time, awareness of the amount of waste on the horizon is essential. Determining the amount of waste produced in the long term by definitive and undefinitive methods can provide managers and designers with information about designing and planning of municipal solid waste management system (Fallahnejad and Abdoli, 2014).

The important point is that linear variable methods are typically used to model the linear relationship between a dependent variable and one or more independent variables. On the other hand, nonlinear and dynamic methods such as fuzzy systems, artificial neural networks and various combinations of these tools are required for fluctuating trends and are necessary to construct an accurate predictive model.

Therefore, according to the contents stated, the purpose of the present study, due to the necessity of waste management, forecasting of municipal solid waste production, and evaluating the performance of neural networks and regression in forecasting and comparing these two models is to show which one of the two models are more efficient.

Literature review

In his study, Cocadel (2015) stated that forecasting solid waste production volumes plays an important role in solid waste master plans. In his model, the Bayesian function in neural network is used to forecast solid waste and shows that the solid waste produced is influenced by factors such as socio-economic and population and the results show that the performance of neural network is at an acceptable level. Patel (2013) estimated a model to forecast the amount of paper waste. This model includes 8 independent variables: number of night stands per person, purchasing power index per household, purchasing power index per person, number of housewives, number of employees in offices and industry, number of employees in services, number of agriculture companies and the number of paper collection locations. The estimated R2 of the model was 48%. Chong et al. (2010) first examined the advantages and disadvantages of time series and multi-factor models, and then forecasted the amount of waste produced by 2036 using a multi-factor analysis model. In their study, GDP per capita, population, and housing number were considered as independent variables and the amount of daily waste produced per capita was considered as the dependent variable and was estimated as the best auto regression model. Then, by forecasting the input variables, forecasted the amount of industrial and domestic waste produced by 2036. Fallahnejad and Abdoli (2014) stated that waste generation in human societies is an everyday routine and natural. In this study, artificial neural network has been used as an efficient tool for modeling the amount of waste produced in Mashhad city. The results show that by making preprocesses on the raw data input to the model, more accurate results can be obtained. Three different modes were investigated and the best pre-processing including logarithm, process deletion and data standardization was selected. The network architecture was obtained with two hidden layers, each with 5 neurons, MAPE of 0.06, MSE of 0.46 and correlation coefficient of 0.86. Karimi Jashani et al. (2012) stated that the lack of efficient waste management system, rapid population growth and consequently ever-increasing waste generation has made solid waste one of the most important environmental problems today. In their study, they used feedforward neural network and linear regression model and the obtained results were compared. In order to carry out these studies, the amount of waste produced in 74 villages of Fars province in 2010 were used and neural network power and regression model in forecasting waste generation of these locations were evaluated which results indicates the superiority of the neural network.

Methods

The basis of any science is its method of recognition, and the validity and value of the laws of any science is based on the methodology in which science is used. In fact, it can be said that the effectiveness of a research task rests on the correct choice of the research method suitable for that particular type of research. As mentioned, this study, in terms of purpose is "applied". in the sense that when researching with the aim of obtaining the results of research findings to solve existing problems of the organization, it is considered as applied research. On the other hand, the research method is "descriptive" because it uses library and documentary studies and library studies have led us to well describe anything that's here. In this research, descriptive method was used and in terms of type it is placed in applied research group. After explaining the theoretical foundations of the research, formulating the assumptions in a way compatible with the aims of the research, formulating the appropriate research method, determining the population and statistical sample, and finally gathering the data and information needed for the research using reliable sources and tools, It is time to examine the relationships between variables. At the analysis stage, the researcher makes judgments about whether to accept or reject the relationships based on the findings from the analysis of the relationships among the variables and thus answers the research questions. In the previous chapter, issues related to the data collection methods and information needed, how to calculate and measure variables, and how to examine the relationships between variables using this information were presented.

In this paper, the relationship between the variables of the present study was investigated using MATLAB and SPSS software. Initially, the data were analyzed by linear regression technique and then by using neural network prediction of municipal solid waste production. Since municipal solid waste production depends on the climatic conditions, in this research, in order to eliminate different climatic conditions of different regions of the country, Arak city is considered. Following are 9 variables identified according to the research literature, which are evaluated by experts as the variables that scored above 0.8 and are the variables used in this study.

Table 4-1: variables of problem

signal	name of variables	Score
A1	population	0.86
A2	minimum temperature	0.85
A3	maximum temperature	0.91
A4	geographical location	0.63
A5	season of the year	0.62
A6	Annual household income	0.81
A7	Annual rainfall	0.75
A8	Technology level	0.87
A9	Number of housing	0.79

Evaluating the hypothesis of normality of the variables

Since the normality of the variable depends on the normality of the model residuals, it is necessary to check its normality before fitting the model.

Table 4-2 : Kolmogorov - Smirnov test (K-S) related to the dependent Variable (Waste production)

Significance level	Kolmogorov Smirnov	Standard deviation	average	Number	Variable
0.836	0.621	22662.84	71507	21	Waste production

The null hypothesis and the alternative hypothesis of the normality test are as follows:

The distribution of data (abnormal returns) is normal: H0

Distribution of data (abnormal returns) is not normal: H1

The Kolmogorov-Smirnov test was used to test the above hypothesis. In this test, when the significance level is less than 5%, the null hypothesis at 95% of certainty is rejected; Based on the presented values (Table 4-2) since the significance level values for waste generation in the model are more than 5%, therefore the null hypothesis, i.e. hypothesis of normality of the variables, is not rejected.

Table 4-3: Correlation coefficient, Adjusted determination coefficient and Durbin - Watson Test

Durbin-Watson	Estimated standard error	Adjusted coefficient of determination	Coefficient of determination	Correlation coefficient	Model
2/526	609/73	0/999	0/999	1.00	1

Table 4-3 shows the values of the correlation coefficient and the adjusted coefficient of determination. The adjusted coefficient of determination indicates how much of the dependent variable, i.e. municipal solid waste production, can be explained and forecasted by the independent variables. In this example, the independent variables can explain and forecast 99% of the dependent variable, i.e. municipal solid waste generation, which is a significant amount.

Table 4-4: analysis of variance

Significance level	F statistics	average of squares	Degrees of freedom	sum of squares	Model
0.000	5521.99	2.05E9	5	1.027E10	**Regression**
		371840.72	15	5577610.89	**residual**
			20	1.027E10	**Total**

Table (4-4) shows the analysis of variance between the variable of forecast error of profit per share as an independent variable and the abnormal return as dependent, according to this output, the overall significance of the regression model is tested by ANOVA table using the following statistical hypotheses:

There is no linear relationship between two variables: H0

There is a linear relationship between two variables: H1

Since sig (significance level) is less than 5%, the hypothesis of linearity of the two variables is confirmed.

Table 4-5: results of multiple regression model

	Not standardized coefficients		Standardized coefficients	T statistics	Significance level
	B	Std.error			
Constant value	18305.140-	10059.122		1.960-	049.
population	157.	018.	488.	8.895	000.
Minimum temperature	254.754-	209.264	008.-	1.217-	242.
maximum temperature	70.108-	71.199	006.-	985.-	340.
Annual household income	23.734	2.985	327.	7.952	000.
Technology level	1468.890	178.731	197.	8.218	000.

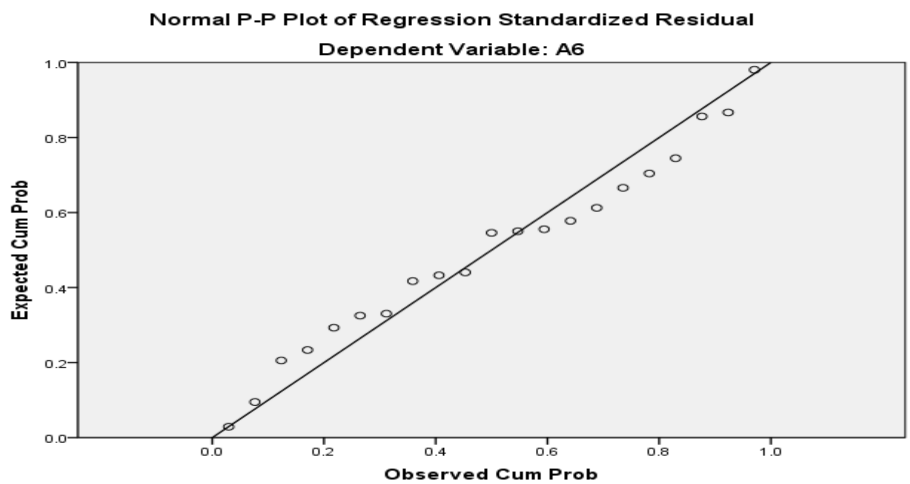

Diagram 4-1: scatter between variables

diagram 4-1 The presence or absence of a relationship between two variables can be shown in a diagram called the scatter diagram. Where in this diagram each of the variables are on one axis and the points obtained on the diagram, Shows their correlation.

Table 4-6: error values in regression model

	MAPE	MSE	RMSE
Regression model	0.132	0.198	0.44

It should be noted that the method of calculating the error in this study is the mean absolute error percentage and mean squared error which is explained below and the results of neural network implementation are presented in the table.

$$MAPE: \frac{\sum_{t=1}^{n}\left|\frac{x_t - x'}{x_t}\right|}{n}$$

$$\mathrm{MSE} = \frac{1}{\mathrm{N}}\sum_{t=1}^{n}(x_t - x')^2$$

MAPE (Mean Absolute Percentage Error)

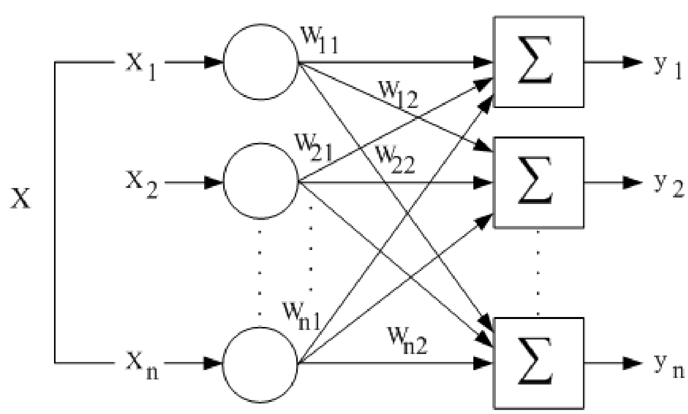

Figure 4-2: Two layered neural network structure

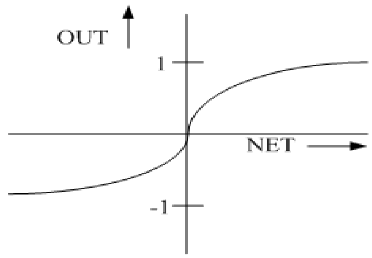

Figure 4-3: Hyperbolic Tangent Function

Table 4-7: neural network error rate in training mode

Model type	The number of repetitions	Training Algorithm	arrangement	Transfer function		MAPE	MSE
				first layer	hidden layer		
Perceptron 1	1000	Levenberg Marquardt	14	Hyperbolic Tangent	Linear	0.131	0.124
Perceptron 2	1000	Descending Gradient	12	Hyperbolic Tangent	Linear	0.124	0.176
Perceptron 3	1000	Descending Gradient	5	Hyperbolic Tangent	Linear	0.118	0.128
Perceptron 4	1000	Levenberg Marquardt	9	Hyperbolic Tangent	Linear	0.117	0.145
Perceptron 5	1000	Descending Gradient	8	Hyperbolic Tangent	Linear	0.156	0.168
Perceptron 6	1000	Descending Gradient	11	Hyperbolic Tangent	Linear	0.125	0.145
Perceptron 7	1000	Levenberg Marquardt	13	Hyperbolic Tangent	Linear	0.119	0.183
Perceptron 8	1000	Levenberg Marquardt	20	Hyperbolic Tangent	Linear	0.113	0.118

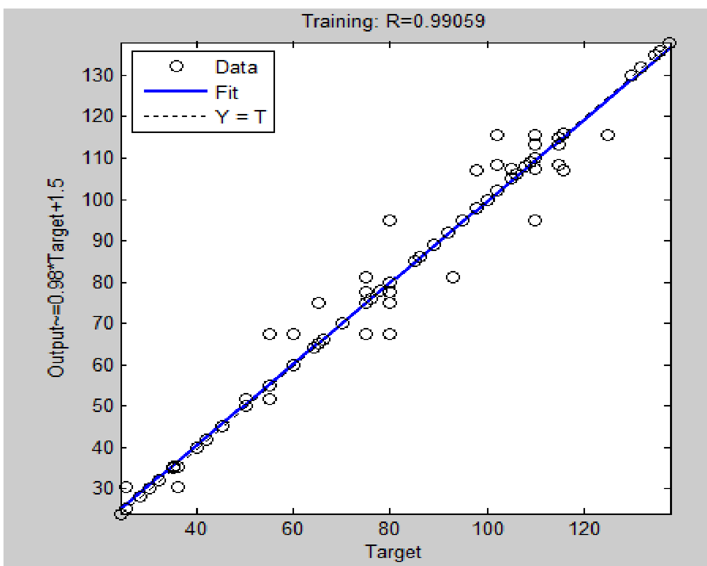

Diagram 4-2: data scatter in training mode

Table 4-8 : neural network error rate in test mode

Model type	The number of repetitions	Training Algorithm	arrangement	Transfer function		MAPE	MSE
				first layer	hidden layer		
Perceptron1	1000	Levenberg Marquardt	14	Hyperbolic Tangent	Linear	0.138	0.195
Perceptron2	1000	Descending Gradient	12	Hyperbolic Tangent	Linear	0.122	0.163
Perceptron3	1000	Descending Gradient	5	Hyperbolic Tangent	Linear	0.104	0.165
Perceptron4	1000	Levenberg Marquardt	9	Hyperbolic Tangent	Linear	0.105	0.206
Perceptron5	1000	Descending Gradient	8	Hyperbolic Tangent	Linear	0.126	0.210
Perceptron6	1000	Descending Gradient	11	Hyperbolic Tangent	Linear	0.105	0.165
Perceptron7	1000	Levenberg Marquardt	13	Hyperbolic Tangent	Linear	0.089	0.101
Perceptron8	1000	Levenberg Marquardt	20	Hyperbolic Tangent	Linear	0.126	0.130

Studying Table 4-8 shows that the best case of artificial networks is when the first transfer function is hyperbolic tangent and the second transfer function is linear and the learning function is Levenberg Marquardt and the number of neurons in the first hidden layer is 13 which In this case the error rate of the best network is 8.9%.

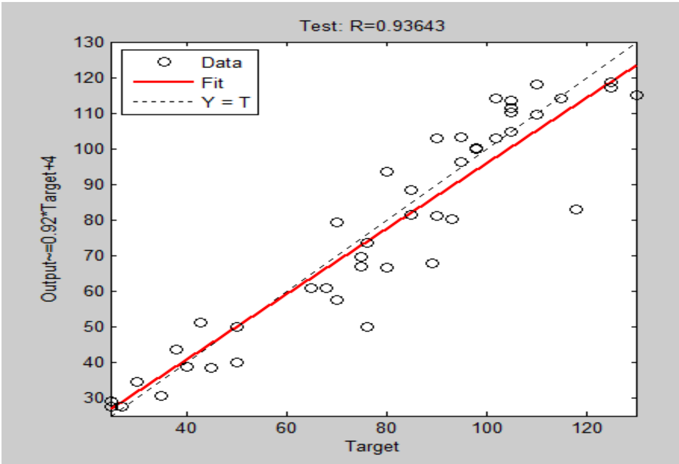

Diagram 4-3 data scatter in test mode

This diagram shows the relationship between the forecasted values and the actual values in the training mode, as it's to be seen, the variable R refers a simple correlation between the two variables, i.e. the intensity of the correlation between the two variables.

Table 4-9 : error values in the neural network model

	MAPE	MSE	RMSE
Neural network model	0.089	0.101	0.31

Results

As you can see The regression model has conditions and assumptions for forecasting waste generation Therefore, the regression model can be used to forecast municipal waste production. The results also illustrate that in order to forecast municipal waste production by regression only variables of population, household income level and technology level can be used and the rest of the variables do not have an appropriate significance level and overall error values of MAPE are 13.2% and MSE error is 19.8% for the regression model which is acceptable. Also neural network model is capable of forecasting municipal waste production using 85% of data for training and 15% for test, which in this neural network model error values of MAPE are 8.9% and MSE error is 10.1% which is acceptable and Indicates the capability of this model.

Table 5-2: results of comparison of neural network model and regression

model	MAPE	MSE
Neural network	0.089	0.101
Linear regression	0.132	0.198

Based on the discussions above, it can be seen that artificial neural network model has better capability of forecasting waste production and regarding MAPE and MSE of each of two models, the presented model using neural network in this research has better performance in forecasting waste generation than linear regression.

References

[1] Pajoheshfarr, Peyman, (2010), "Presenting a New Algorithm to Reduce Artificial Neural Networks Error in Customer Credit Scoring Models", BSc thesis.

[2] Hosseini, Seyyed Saman;Saberi, Morteza;Salarvand, Farhad and Pajoheshfarr, Peyman "Presenting A New Algorithm for Reducing Artificial Neural Networks Error in Bank Customer Credit Scoring Models" Twelfth Conference of Intelligent Systems, Bam, Iran, 2012

[3] Khorasani, Nematollah; Naqibzadeh, Seyed Shahram and Salman Mahini, Abdolrasoul (2010), Forecasting of municipal Solid Waste Production Using Dynamic System Modeling, Fifth Waste Management Conference, Mashhad, Iran, 2010.

[4] Abdoli, Mohammad; Noori, Rouhollah; Jalil, Mehdi and Salehian, Ehsan (2007) "forecasting Tehran Waste generation Using Artificial Neural Network and Multivariate Statistical Methods", Third National Conference on Waste Management.

[5] Azizi Danalo, Samaneh, 2014, "The Role of Municipal Waste Recycling in Creating New sustainable municipal revenue sources" MSc Thesis, Islamic Azad University, Central Tehran Branch

[6] "landfill locating using the MADM methods" 2016 akbarian, mojtaba; Qaffari, salane, teymoori, fateme; First International Conference on Industrial Engineering, Management and Accounting, University of Applied Science

[7] qafari, majid (2015)forecasting natural gas consumption by artificial neural network MSc Thesis in Industrial Management, Arak Azad University

[8] Fatehnia Alireza, 2011 study of condition of solid municipal waste separation at source And providing solutions for public participation increase in the 5th region Municipality of Tehran, ph.d thesis in urban health management (mhp),School of Public Health, Tehran University of Medical Sciences

[9] fallahnejad, malihe; Abdoli, Mohammad (2014) Investigating the Effect of Data Preprocessing on the Accuracy of municipal Waste Production Modeling Results Using Neural Network, Amirkabir Scientific and Research Journal,Civil and Environmental Engineering,vol. 46 ,no. 2, Winter 2014, pages 29-37

[10] Karimi Jeshani, Ayub; Bidokhti, Nasser; Azadi, Sama; Amiri, Hamid and Minosepehr, Mohammad (2012), Evaluation of Artificial Neural Network Power in forecasting Rural Waste Production in Fars Province », 6th National Conference and First International Conference on Waste Management, 2012.

[11] majlesi monireh, damanafshan hojjat (2009) locating municipal landfills in dezful using geographic information system (GIS), 12th national Conference on environmental health of iran, shahid beheshti university of medical sciences, pages 436-441

[12] Mohammad Nouri, Sepehr; Omidvari, Manouchehr and Tahmasebi, Sima, 2014,Evaluation of succesful factors of ISO 14001 system in waste managment of tehran hospitals using Analytic Hierarchy Process (AHP) , Scientific and Research Journal of alburz medical sciences university, vol. 3 , no. 3 , summer 2014 , pages 181-190

[13] Mirzadeh qasab, fereshteh, 2012 ,Evaluation of amount of desirability of garbage Disposal Methods Using Multi Criteria Decision Making Techniques, MSc Thesis in management, bahonar university of kerman.

[14] Mirzaee, Mojgan; Salman Mahini, Abdolrasul; Mirkarimi, Seyyed Hamed, 2012, locating landfills using the Analytic Hierarchy Process and topsis method, Case study: golpaygan city, natural environment journal, natural resources of iran (vol. 67 , no. 1 , spring 2014 , pages 105-119)