Neural Network Regression with Support Vector Regression for Land-Use Growth Prediction

by Herlawati Herlawati

Neural Network Regression with Support Vector Regression for Land-Use Growth Prediction

Herlawati
STMIK Bina Insani
Bekasi, Indonesia
herlawati@binainsani.ac.id

Rahmadya Trias Handayanto
Information System Computer Engineering
Universitas Islam 45 Bekasi
Bekasi, Indonesia
rahmadya.trias@gmail.com

Solikin
Information System
STMIK Bina Insani
Bekasi, Indonesia
Solikin2004@gmail.com

***Abstract*—Growth prediction in a developing cities is important for the city planners to prevent environmental degradation, social and health problems, improper land-use location, etc. Bekasi City in Indonesia, is experiencing rapid urban growth and as a "post-suburbia" city faced many problems related to the land use management. This paper proposed a combination of Nonlinear Autoregressive with External Input (NARXNET) and the Support Vector Regression (SVR) to predict the land-use growth in the study area. The result showed the proposed method was able to predict the number of land-use in Bekasi City several years ahead.**

***Keywords*— land-use growth, non-linear regression, postsuburbanization, Bekasi City**

I. Introduction

Cities worldwide, especially near the metropolitan area, are experiencing rapid urban growth with postsuburbanization phenomenon. This phenomenon is characterized by higher land-use growth and more important than before compared to the central city [1], [2]. Predicting the urban growth of this "post-suburbia" city is useful for the city planners to handle the urban growth effects, e.g. environmental degradation, slums area, social and health problems, etc. [3].

Research on urban growth prediction often used a multilayer perceptron neural network (MLPNN) as a prediction tools in IDRISI software [4]–[6]. Neural network with the backpropagation learning method have been widely used [7]– [11]. One of the neural network regression methods is Nonlinear Autoregressive Neural Network with External Input (NARXNET) with the external input (exogenous variables) as intervention indicators. However, another method, i.e. Support Vector Regression (SVR), has been widely studied [12], [13]. Both MLPNN and SVR have similar capability in handling non-linear data.

This paper proposed the integration between NARXNET and SVR. NARXNET was used as initial prediction before land-use prediction through SVR. Before land-use prediction, both NARXNET and SVR were validated using the existing data. Both Mean Average Percent Error (MAPE) and Mean Square Error (MSE) were used in accuracy calculation.

The rest of the paper is organized as follows. After discussing neural network and support vector prediction methods and comparison, the proposed method predicted the land-use growth in Bekasi City from 2015 to 20130 based on the previous data. The results are discussed and concluded.

II. Methods

There are two kinds of regression: linear and nonlinear regression. Nonlinear regression has an advantage in handling nonlinear data that occur in real life. Both NARXNET and SVR can be used to predict the nonlinear data.

A. NARXNET

Neural network mimics the brain works with the neuron to transfer the information. Weights and biases were set to the neurons by a learning mechanism. The famous learning method is backpropagation which reset the weight and bias backward if there is some errors compared to the target [9].

Inputs in neural network were variables of a system. In projection, the variables were the previous prediction. NARXNET used previous prediction and another variable called external input/intervention indicators. In this study the external input was population. Fig 1 shows the NARXNET structure with the window mechanism.

External input was used to guide the network in predicting whether the prediction is increasing or decreasing. However, the number of land use still become the main part of the prediction model and the model should not use a lot of external input since the problem is projection based on previous data.

In this study a NARXNET function in Matlab software was used with the simple GUI as user interface. Nine neurons were used, sigmoid function in hidden layer and purelin function in output. Iteration are 1000 epoch and other parameters such as learning rate and goals are 0.001 and 0.00000001 respectively.

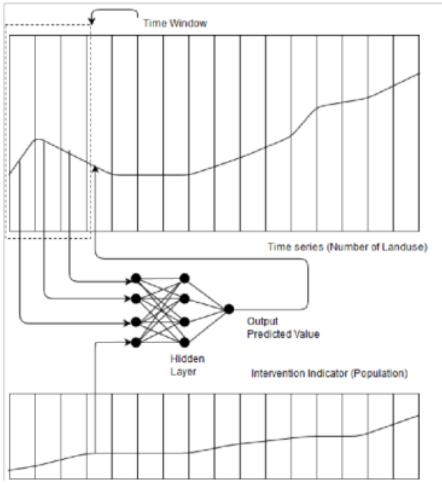


Fig. 1. NARXNET Structure

The previous data were used for prediction. The prediction results were used as input in next prediction. With the external input (in this study used population data), this method predicted number of land use several years ahead.

B. SVR

SVR uses Support Vector Machine (SVM) as a prediction method. Points as support vector were find as classifier. SVR as kernel-based machine learning was applied to tasks, e.g. function approximation and regressive parameter estimation [12]–[15]. SVR run iteratively to find the weights and biases as follows:

$$\mathbf{W} \cdot \mathbf{X} + b = 0 \quad (1)$$

$$H_1: w_0 + w_1x_1 + w_2x_2 \geq 1 \text{ for } y_i = +1 \quad (2)$$

$$H_2: w_0 + w_1x_1 + w_2x_2 \leq 1 \text{ for } y_i = -1 \quad (3)$$

where W is weight, b as bias, and X as variables. A hyperplane was found to separate H_1 and H_2 according to equation (2) and (3). Two function in Matlab were used, i.e. "svmtrain" for training and "svmclassify" for classification. The gauss function was used as kernel function.

C. Data

Land-uses in Bekasi City were captured using historical data menu in Google Earth Pro. Ten Land-use classes were used, i.e. commercial, industrial, elementary school, middle school, college, sport, medical, park, high density residential, and low density residential.

Table 1 shows the current and the historical data of landuse in Bekasi City (shown as a map in Fig 2). Some land-uses have a tendency to increase, except the park class. The Bekasi City population increased up to 3 million in 2018.

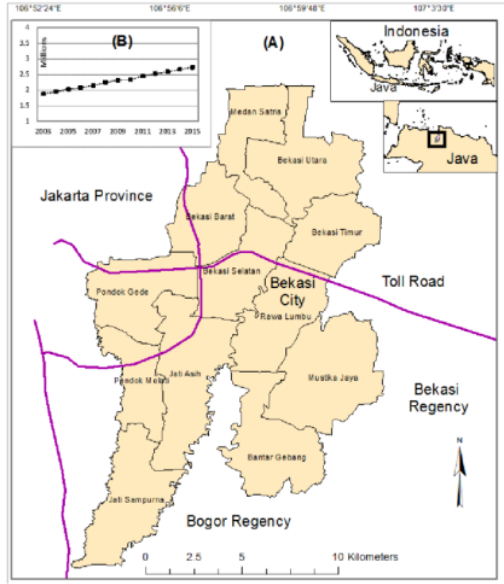


Fig. 2. Bekasi City as Study Area

Table 1. Land Use Data

LU Class	2003	2006	2009	2012	2015	2018
1. Commercial	245	276	309	325	344	355
2. Medical	87	92	101	107	118	121
3. Sport	19	23	24	24	26	27
4. HD Res	2122	2217	2262	2282	2326	2379
5. Industrial	118	131	139	149	154	167
6. LD Res	1446	1481	1580	1539	1537	1362
7. College	35	35	37	38	38	40
8. Middle School	202	203	203	206	207	210
9. Elementary School	270	270	272	272	273	271
10. Park	1904	1904	1887	1858	1772	1805
Population (millions)	1.88	2.08	2.32	2.52	2.73	3.01

D. Proposed Method

Both methods showed good prediction performance using data in Table 1. Mean Average Percent Error (MAPE) was used for error calculation:

$$\text{MAPE} = \frac{1}{n} \sum abs\left(\frac{actual - prediction}{actual}\right) \quad (4)$$

SVR outperformed NARXNET with 3.79% against 8.98% MAPE. As comparison, another methods, i.e. linear regression had MAPE about 21.15% error. However, SVR need input data before prediction that can be supplied by NARXNET with its ability to predict only with external input (population). The algorithms of the proposed methods are as follows:

1. Use NARXNET with the historical data to predict land uses from 2003 to 2030 and its population with 9 neurons.
2. Check whether the Mean Square Error (MSE). If MSE greater than 0.3, stop the training and start prediction.
3. Use NARXNET prediction result and the population as the input data of SVR.
4. Start SVR module

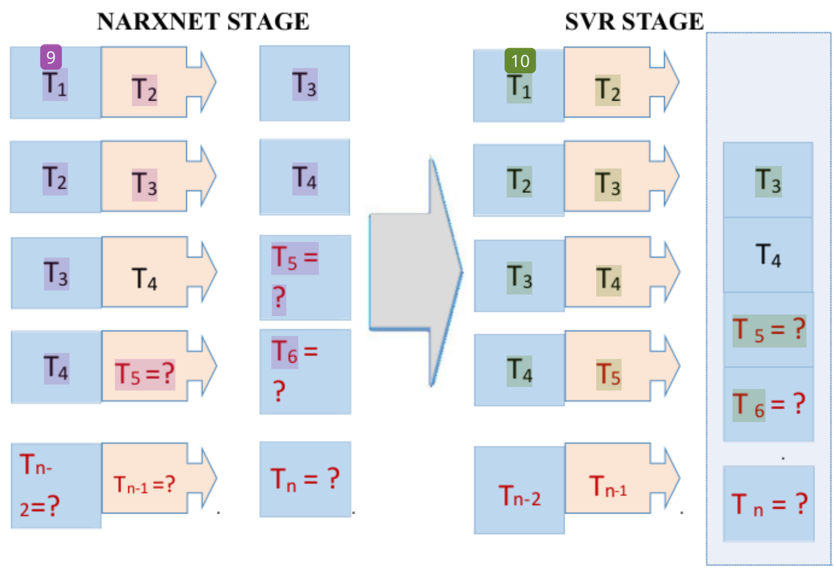


Fig. 3. NARXNET and SVR combination for LU Prediction

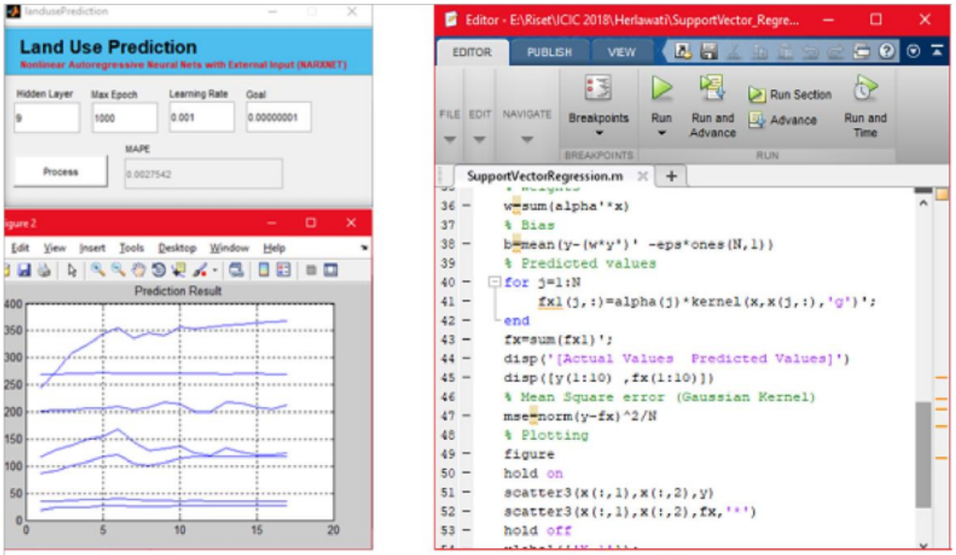
Fig 3 shows an example to predict T_5 to T_n using NARXNET based on population growth (left) and refine the result through the use of SVR to predict T'_5 to T'_n (right). Four data (T_1, T_2, T_3, and T_4) and external input (population growth from T_1 to T_n) were used to create a prediction model. These data were separated into some groups which are called window. First tuple T_1 and T_2 were used to predict T_3, Second tuple T_2 and T_3 were used to predict T_4 and continue until the last T_{n-2} and $T_{(n-1)}$ were used to predict T_n. Both the window and the prediction result from NARXNET then were predicted through the use of SVR with the result T'_5 to T'_n as the revision from NARXNET's result. The predictions are shown in red color in Fig 3.

E. Validation

Validation is an important phase before a model is used for land-use prediction. The actual value is compared with the prediction from the model. In Fig 3 the actual value of T_5 is compared with the prediction (in NARXNET and SVR). In this study, number of land use in 2018 was used for validation through MAPE calculation in equation 4. Table 2 shows the comparison of three methods, i.e. NARXNET, NARXNET with SVR, and linear regression. NARXNET with SVR (8.59% of MAPE) outperformed both NARXNET (8.98% of MAPE) and linear regression (21.15% of MAPE). The MAPE of NARXNET with SVR showed the accuracy of 91.41% that was good enough for land-use prediction.

Table 2. Comparison with other methods

Land Use Class	**MAPE**		
	NARXNET	**NARXNET with SVR**	**Regression**
Commercial	14.63%	17.55%	50.58%
Medical	35.15%	39.92%	11.74%
Sport	17.10%	25.43%	133.33%
HD Res	6.12%	3.15%	32.01%
Industrial	24.79%	11.65%	56.12%
LD Res	16.29%	13.46%	12.73%
College	17.28%	9.69%	25.60%
Middle School	0.19%	1.92%	1.08%
Elementary School	1.43%	1.51%	3.52%
Park	10.63%	13.10%	11.73%
Total	8.98%	8.59%	21.15%



(a) (b)

Fig. 4. NARXNET GUI (a) and SVR Script (b)

A Graphic User Interface (GUI) of NARXNET prediction was created to supply input data for SVR module. Both NARXNET and SVR module are from Matlab 2013 software.

III. Result and Discussion

With two data input as window a NARXNET GUI generated 80 rows prediction. From 2018 to 2030 (in threeyear period) NARXNET used external input and previous prediction result to predict the land use. Fig 4 shows the running GUI of NARXNET (a) and SVR script (b).

Prediction results from NARXNET were used as the input of SVR. NARXNET module shown in Fig 4.a was run several numbers until finding the best MAPE score.

SVR module then refined the NARXNET result with new predictions as shown in Fig 5 and Table 3 in which most of the land use increased except the park. Therefore, city planners need a proper land-use plan to handle this rapid landuse growth because of the limited area of Bekasi City for land use.

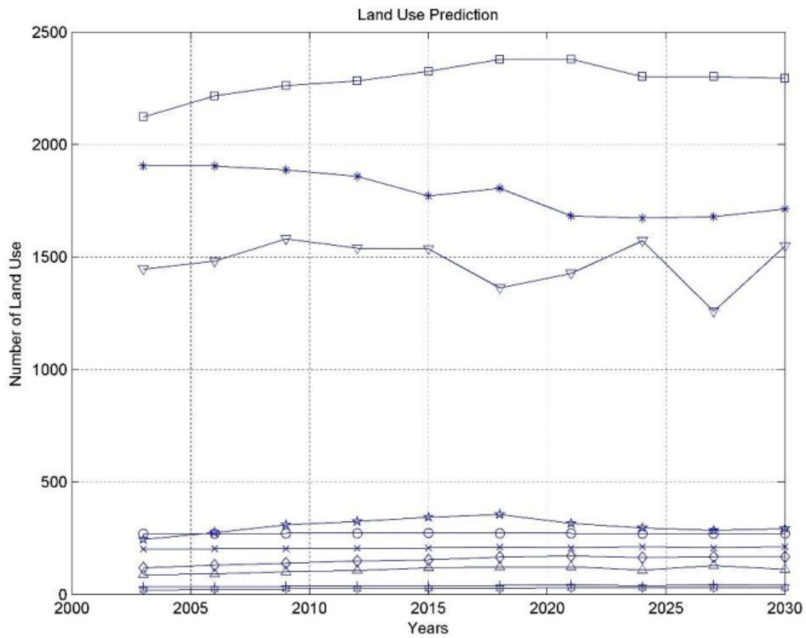


Fig. 5. Land Use Prediction

Table 3. Land Use Prediction

No	LU Class	2003	2006	2009	2012	2015	2018	2021	2024	2027	2030
1	Commercial	245	276	309	325	344	355	316.35	295.82	285.51	292.58
2	Medical	87	92	101	107	118	121	122.83	108.01	128.22	112.06
3	Sport	19	23	24	24	26	27	29.70	29.33	29.97	29.24
4	HD Res	2122	2217	2262	2282	2326	2379	2379.52	2301.77	2301.30	2295.42
5	Industrial	118	131	139	149	154	167	171.17	164.46	167.32	167.87
6	LD Res	1446	1481	1580	1539	1537	1362	1427.76	1572.17	1258.15	1548.12
7	College	35	35	37	38	38	40	42.63	38.05	43.40	39.73
8	Middle School	202	203	203	206	207	210	207.43	212.99	207.69	212.92
9	Elementary School	270	270	272	272	273	271	272.53	270.53	270.40	271.56
10	Park	1904	1904	1887	1858	1772	1805	1682.95	1671.80	1678.66	1713.43

IV. Conclusions

Planners need a land-use growth prediction tool to manage the current and future land use as well as to avoid the potential negative effects from rapid urban growth. Results showed that the combination of NARXNET and SVR were able to work together in predicting the land-use growth for several years ahead. The accuracy of the proposed method also outperformed the other methods, i.e. NARXNET and linear regression. The prediction result can be used as a base for land use optimization as well as land use simulation where exact land use location are predicted. The result can also be used as a warning for local city government to control the growth of land use and the population.

Acknowledgment

The authors thank to Research, Technology and Higher Education Department (RISTEK-DIKTI) of Indonesia, local government of Bekasi City, Higher Education in Informatics and Computer (APTIKOM), STMIK Bina Insani, and Universitas Islam 45 Bekasi. Also for the reviewers for the comments and valuable suggestions.

References

[1] R. T. Handayanto, N. K. Tripathi, S. M. Kim, and S. Guha, "Achieving a sustainable urban form through land use optimisation: Insights from Bekasi City's land-use plan (20102030)," *Sustain.*, 2017.

[2] T. Firman, "New town development in Jakarta Metropolitan Region : a perspective of spatial segregation," vol. 28, pp. 349–368, 2004.

Neural Network Regression with Support Vector Regression for Land-Use Growth Prediction

ORIGINALITY REPORT

16%	13%	11%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.scilit.net Internet Source	5%
2	repository.unismabekasi.ac.id Internet Source	2%
3	Rahmadya Trias Handayanto, Sohee Minsun Kim, Nitin Kumar Tripathi, Herlawati. "Land use growth simulation and optimization in the urban area", 2017 Second International Conference on Informatics and Computing (ICIC), 2017 Publication	2%
4	Teddy Mantoro, Rahmadya Trias Handayanto, Media Anugerah Ayu, Jelita Asian. "Prediction of COVID-19 Spreading Using Support Vector Regression and Susceptible Infectious Recovered Model", 2020 6th International Conference on Computing Engineering and Design (ICCED), 2020 Publication	2%

journal.ummgl.ac.id

5	Internet Source	1%
6	core.ac.uk Internet Source	1%
7	Abiodun Ayodeji, Yong-kuo Liu. "SVR optimization with soft computing algorithms for incipient SGTR diagnosis", Annals of Nuclear Energy, 2018 Publication	1%
8	Endang Retnoningsih, Fata Nidaul Khasanah, Rahmadya Trias Handayanto, Herlawati, Syahbaniar Rofiah, Solikin. "Knowledge Management System for Supporting the Small Medium Enterprise (UMKM) in Bekasi City", 2019 Fourth International Conference on Informatics and Computing (ICIC), 2019 Publication	1%
9	dokumen.tips Internet Source	1%
10	pdffox.com Internet Source	1%
11	www.mdpi.com Internet Source	<1%
12	Syahbaniar Rofiah, Dadan Sukma, Rahmadya Trias Handayanto, Herlawati, Solikin, Didik Setiyadi. "Mobile-based Augmented Reality	<1%

for Historic- Building Lessons in West Java", 2019 Fourth International Conference on Informatics and Computing (ICIC), 2019
Publication

13	Rahmadya Trias Handayanto, Sumanta Guha, Nitin Kumar Tripathi, Herlawati. "Genetic Algorithms with Variable Length Chromosomes for High Constraint Problems in Spatial Data", 2018 Third International Conference on Informatics and Computing (ICIC), 2018 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		