東京大学　関本研究室 / Sekimoto Lab. IIS, the University of Tokyo.

# Rent Estimation Model and Visualization of Area Potential by Using Deep Neural Network —A Focus on Attributes of Building and Area—

M. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama

## Background

**1.** There is a need for the a system to estimate the prices rent paid by tenants as rent and grasp consider the features of living areas from several perspectives.

**2.** In accordance with the trend for open data, a variety of data which owned by private companies and local governments own have been easily accessible.

**3.** For both price estimation and an analysis of factors affecting price, the hedonic approach has been widely used. On the other contrary, machine learning methods have become more and more common with an increase of in the amount of open data.

The goal is to build a system to estimate the rent paid by tenants, and analyze the features of the relevant areas to enable an assessment of the impact of development projects therein by using open data and machine learning methods.
As the first step, we built a price estimation model with both building and area attributes.

## Dataset

At Home Real Estate Data Library Nationwide 1999–2013 dataset, which is a joint use data belonging to the Center for Spatial Information Science at the University of Tokyo, was used for estimation

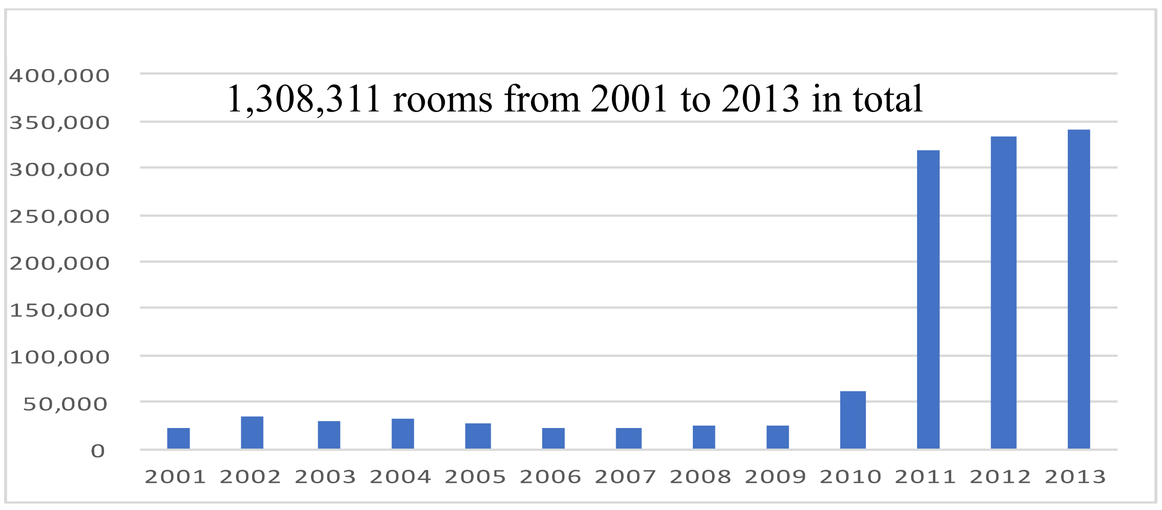


Figure 1. Number of registered rooms in At Home Dataset in each year in the Tokyo 23 Wards

Table 1. Data items

| | | | |
|---|---|---|---|
| 1. Room Number | 2. Use | 3. New/Second-hand | 4. Prefecture |
| 5. Location1 | 6. Location2 | 7. Railway Line | 8. Station |
| 9. Walk (min) | 10. Bus（min） | 11. Walk to Bus Stop（min） | 12. Bus Stop |
| 13. Rent | 14. Land Area | 15. Exclusive Area | 16. Balcony Area |
| 17. Land Ownership | 18. Building Coverage | 19. Floor Area Ratio | 20. No. of Above-ground Stories |
| 21. No. of Underground Stories | 22. Floor | 23. Year Built | 24. Month Built |
| 25. Date of Registration | 26. Building ID | | |

## Rent Estimation Model

Table 2. Explanatory Variables

| Attributes | Data Type |
|---|---|
| 1. Office for Rent | Binary |
| 2. Shop for Rent | Binary |
| 3. Office and Shop for Rent | Binary |
| 4. Area (m²) | Real |
| 5. No. of Above-ground Stories | Integer |
| 6. No. of Underground Stories | Integer |
| 7. Floor | Integer |
| 8. Year of Building | Integer |
| 9. Neighborhood Dummy Value | Binary |

Building Attributes

Area Attributes
939 neighbor hoods in Tokyo 23 wards

**1. Deep Neural Network**
No. of perceptron:1000
No. of layers:5
Learning rate:0.001
Trained for 10,000 times
Repeated this for 30 times

**2. Random Forest Regression**
Even though over-fitting might occur, we set the depth of the decision trees to 100 since the number of explanatory variable is 937. We trained the model and estimate rent for 10 times.

**3. Hedonic Approach**

$$y = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + a_9x_9 + \ldots + a_{947}x_{947}$$

$$x_4y = x_4(a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + a_9x_9 + \ldots + a_{947}x_{947})$$

$x_1, x_2, \ldots, x_8$ are consistent with the eight building attributes in Table 2, whereas $x_9, \ldots, x_{947}$ are neighborhood binary dummy values showing the neighborhoods of the 939 neighborhoods.

## Comparison

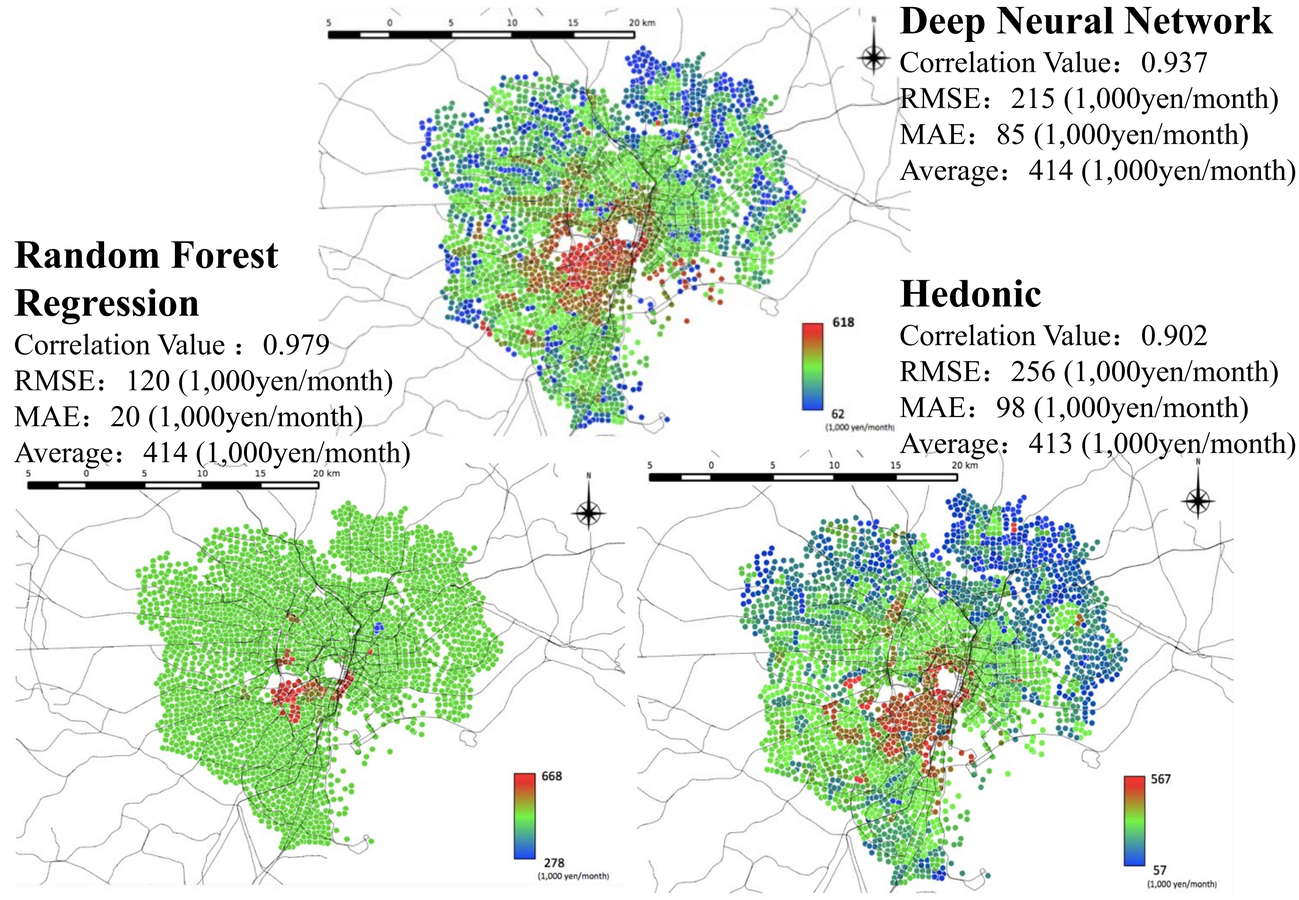


Figure 2. Comparison of three methods

## Conclusion

Deep Neural Network is suitable for both price estimation with both building and area attributes.
Challenges for the future are as follows:

- Bayesian optimization should be applied to estimate hyperparameters.
- The same analyses and trials should be conducted using data for other residences.