Abstract: Employing a large dataset (at most, the order of n = 10^6), this study attempts enhance the literature on the comparison between regression and machine learning (ML)–based rent price prediction models by adding new empirical evidence and considering the spatial dependence of the observations. The regression–based approach incorporates the nearest neighbor Gaussian processes (NNGP) model, enabling the application of kriging to large datasets. In contrast, the ML–based approach utilizes typical models: extreme gradient boosting (XGBoost), random forest (RF), and deep neural network (DNN). The out–of–sample prediction accuracy of these models was compared using Japanese apartment rent data, with a varying order of sample sizes (i.e., n = 10^4, 10^5, 10^6). The results showed that, as the sample size increased, XGBoost and RF outperformed NNGP with higher out–of–sample prediction accuracy. XGBoost achieved the highest prediction accuracy for all sample sizes and error measures in both logarithmic and real scales and for all price bands (when n = 10^5 and 10^6). A comparison of several methods to account for the spatial dependence in RF showed that simply adding spatial coordinates to the explanatory variables may be sufficient.