

# Noise Surrogate Model for Urban Planning

Authors: Deepika Dittakavi, Lois Dankwa, Tyler Gmerek

Date: June 10, 2020

Industry Project: Data Science & Analytics Certificate Bootcamp with Stack Education

## Executive Summary

**Background:** Architects, engineers, real estate developers, and municipalities represent some of the organizations charged with the responsibility of ensuring that building mass in yellow zone comply with code and regulations. One goal of Spacemaker AI is to explore the use of surrogate models to mimic the behavior of simulation models including noise due to the benefits of being computationally cheaper to evaluate and higher turnaround in rapid iterations.

**Goal:** To analyze statistical models to best predict the building mass in yellow zone.

**Data & Method:** Data provided by Spacemaker AI consisted of 4500 non-specific noise simulations from 9 different noise scenarios, 250 specific noise simulations from the test scenario which together were used to train and test models and 250 noise simulations from the test site scenario were used to apply the final model. Each dataset included noise sources and building configurations along with fraction yellow zone. Since the fraction yellow zone, the dependent variable to be predicted, is a continuous variable Regression analysis has been used by measuring root mean square error for comparison. Features engineered from the configurations were analyzed and then applied to various Machine Learning concepts.

**Key Insights:** Based on data and feature exploration, the fraction yellow zone is lower for lower noise sources and is greater for lower building mass coverage. Also, for greater average distance of buildings to road the fraction was lower. Feature variation importance plot showed that Noise coverage, average distance to noise source and building coverage and height at specific distance from noise source have been significant. Out of all the models trained, Regression Trees, Random Forest and Support Vector Machines resulted in lower RMSE values and hence were used together as an ensemble model on the test site data resulting in a significant decrease in the RMSE value to 0.047 from baseline model RMSE of 0.169.

To further decrease the value Ensemble models with stacking and blending and Neural Networks concepts can be evaluated.

**Conclusion:** With about 95% prediction accuracy of the fraction yellow zone, clients would be well informed and will be able to use the prediction of the fraction yellow zone for their intended proposal of building and noise configuration and thereby be able to comply with code and regulations.