Team 30
A LSTM-DNN Model for Gentrification Prediction

Ramunas Genys, Deivis Banys, Chris Liu, Richard Xia

Highlights

- A LSTM-DNN hybrid model for predicting the home value for next year with significant result
- Genrating features using momentum, volatility, mean and begin/end change ratio on time-series to make prediction
- 0.64 correlation in prediciton the gentrification process
- we aare not only tell whether the tract is gentrified, but also tell you it has been gentrified for how many percentages.

Background

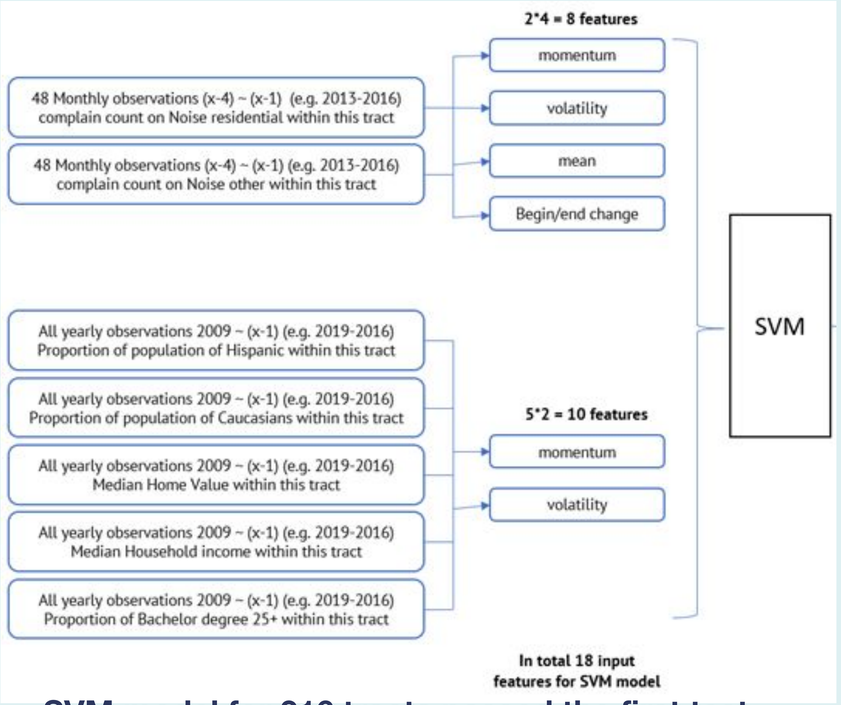
In this project, we are going to forecast the gentrification process and the residential home value movement of one interested tract for a specific year based on historical observation on home value, educational level, 311 call complains, races composition and household income.

Data

311 call data: the dataset that contains all complaints calls in different tracts

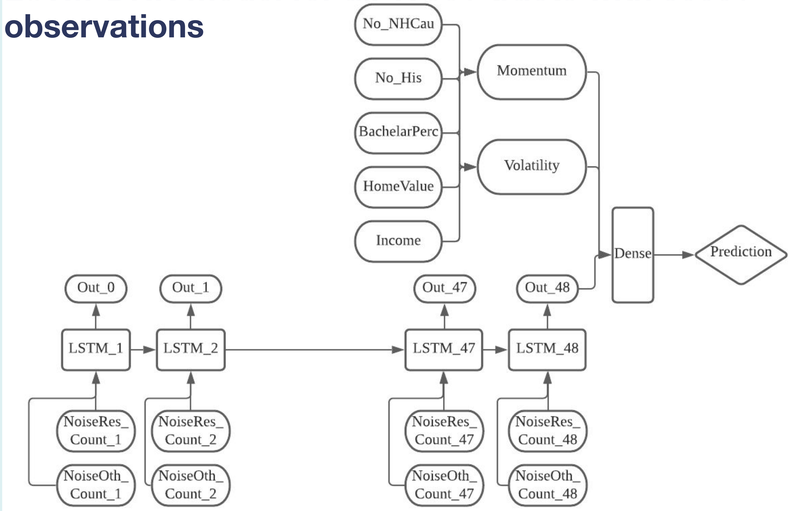
census data: population, races, homevalue and education for all tracts

ShapeFile: For identify the location of each tarct and visualize



SVM model for 316 tracts passed the first test

LSTM-DNN model for all 1700+ tracts with 8000+ observations



Quantify process of gentrification

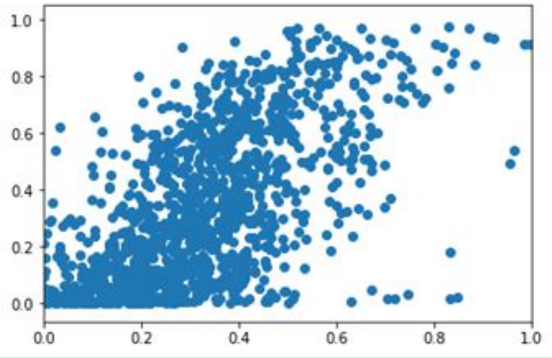
Rather than binary result, we define threshold k as the gentrification property for each geoid, the minimum threshold for the tract to pass the second test

$$min_{0 \le k \le 1} F_k(Geoid) = 1$$

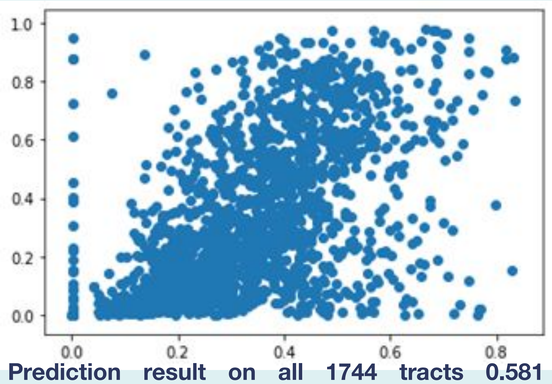
Model

There are only 316 tracts passed the first test of gentrification process, hence, we only hae 800 total observations for gentrification prediction. We use SVM on this dataset to predict gentrification process

Because K also represent the percentile of homevalue this tract hasin the NYC area, predict on k for whole 1700+ tracts is a prediction on home property value. We have 8500+ obervations for model training, hence, I use LSTM-DNN model to predict on k for each tract



Prediction result on 316 tracts pass first test with 0.636 correlation on test dataset



Prediction result on all 1744 tracts 0.581 correlation on test dataset for home value prediction

We have significant result on gentrfication process prediction for next year with 0.636 correlation and 0.581 correlation for home value prediction in next year