Model Card - Housing Price Predicitons

Model Details

- Developed by Thomas Moore (thomashokie@vt.edu), Ben Ha (benh04@vt.edu), and Ben Post (benjpost@vt.edu), 2023
- Random Forest Regression model used to predict the pricing of houses using variables such as, square feet, bedrooms, bathrooms, the neighborhood, and year built.

Intended Use

- Intended purpose for easily estimating prices of homes, as well as use by scientists/data collectors, not designed for exact pricing of houses rough estimate of house depending on variables.

Factors

- Takes into account physical factors of the house, square feet, bedrooms, bathrooms, year built, etc.
- Also takes geographical factors such as surrounding neighborhood, and proximity to the ocean.

Metrics

- Our first Linear Regression Model obtained an accuracy score of -1825193204.1768553, prompting a switch to the Random Forest Regression Model
- Random Forest Regression Model had an accuracy score of 51% with its first prediction of pricing.

Evaluation Data

- Data originally from *"Sparse spatial autoregressions." Statistics & Probability* by Pace, R. Kelly, Ronald Barry
- Data processed by original programmrer

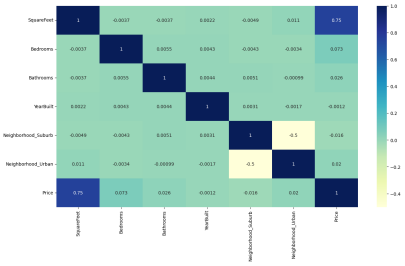
Training Data

- Models uses variables that take into account the square feet, # of bedrooms, # of bathrooms, year the house was built, as well as the price.

Caveats and Recommendations

- Original Linear Regression model was too inaccurate to be used, so the Random Forect Regression Model was used instead
- Model does not take into account furnshings/decorations

Quantitative Analysis



Ethical Considerations

- Ethical concerns that may arise with this kind of machine learning include bias, price gouging/manipulation
- Other consideration arise with machine learning as a whole, regarding pay and aspects of treatment of employees

Piano, Samuele Lo (2020). *Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward.* https://www.nature.com/articles/s41599-020-0501-9