

PREDICT THE HOUSE PRICES FOR RESIDENTIAL HOMES IN AMES, IOWA

ASIYA SHAKEEL

DATA SCIENCE CAPSTONE PROJECT JUL 2023



THANKS TO SPRINGBOARD MENTOR

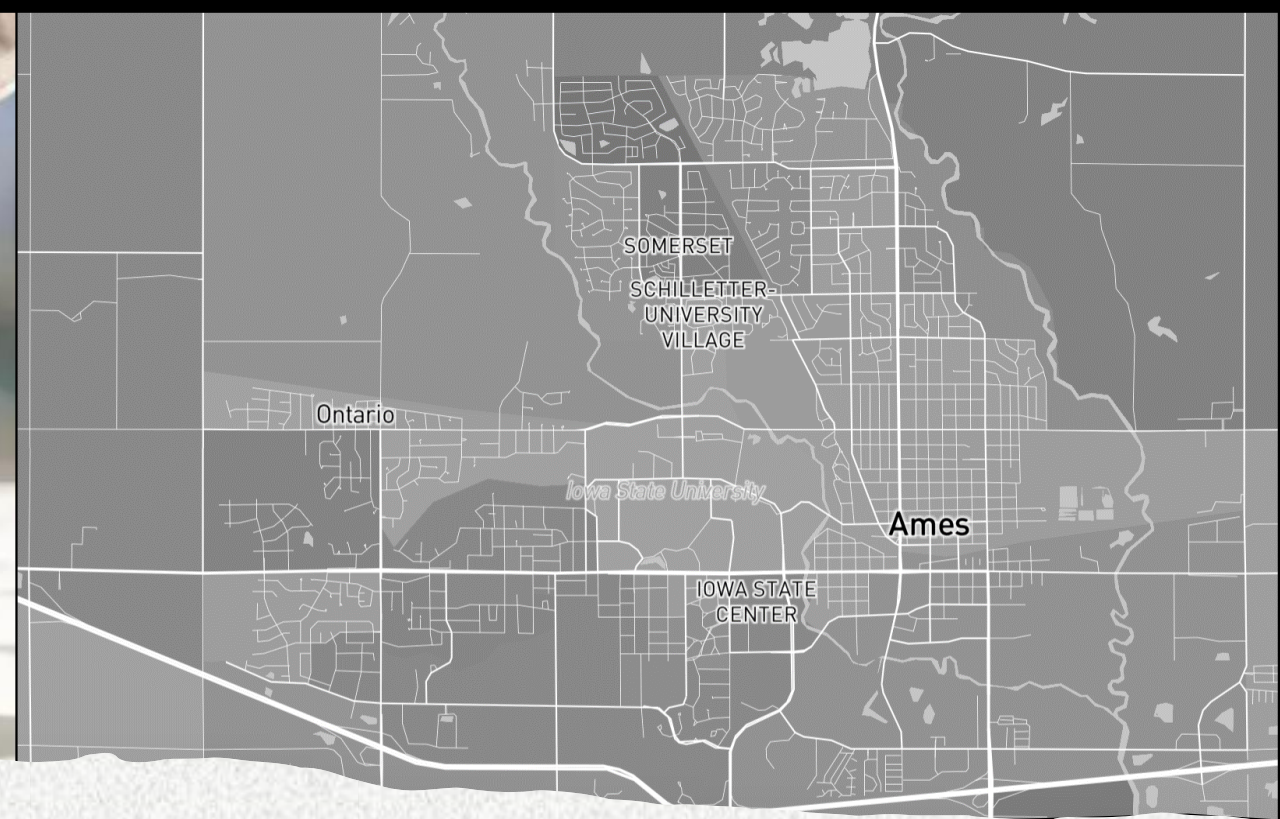


AJ Sanchez, Ph.D.

GOALS:

- Accurately Predict House Prices.
- Identify Key Features that Enhance House Value.





The intended stakeholders are Real Estate Agents, Brokers, Homeowners, Investors and prospective buyers.

DATA EXPLORATION: 25 neighborhoods, 1460 houses and 80 features

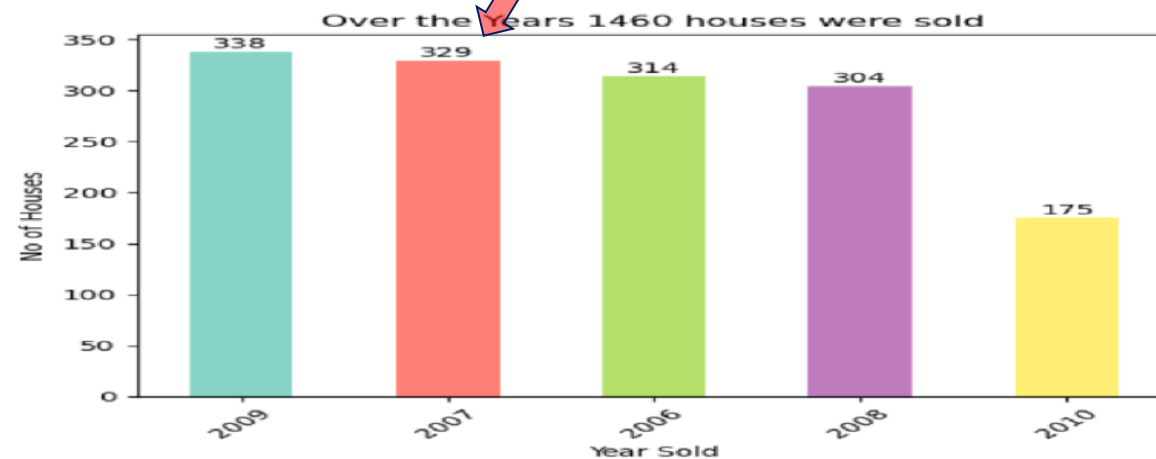
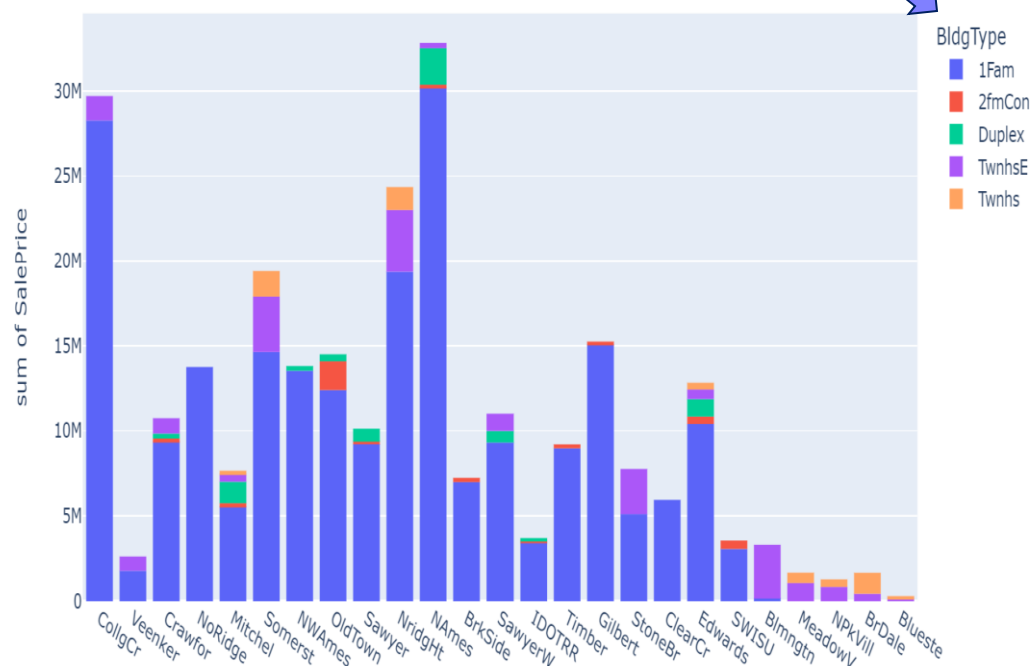
1st Floor square feet

price range \$50,000 to \$800,000

over 300 houses per year have been sold

ACTIVE HOUSING MARKET

Sales Prices of Houses in all the Neighborhood with BldgType



MACHINE LEARNING PIPELINE

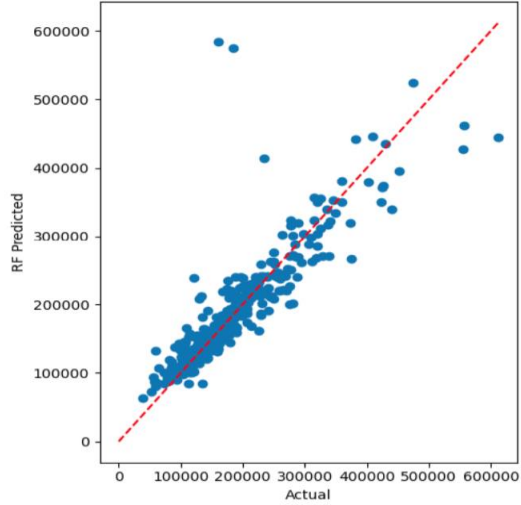
Model	MAE	MAPE
LGBMRegressor	\$17,976	11%
XGBRegressor	\$18,707	12%
RandomForestRegressor	\$19,679	12%
GridSearchCV (RandomForestRegressor)	\$19,455	12%
GridSearchCV (LGBMRegressor)	\$16,758	11%
GridSearchCV (XGBRegressor)	\$16,718	11%
RandomSearch (RandomForestRegressor)	\$19,814	12%
RandomSearch (LGBMRegressor)	\$18,152	11%
RandomSearch (XGBRegressor)	\$17,261	11%

1. We started with an initial model Linear Regression.
2. We tried different models like RandomForestRegressor, LGBMRegressor, and XGBRegressor to find the best one.
3. We fine-tuned the models using techniques called GridSearchCV and RandomizedSearchCV to improve their performance.

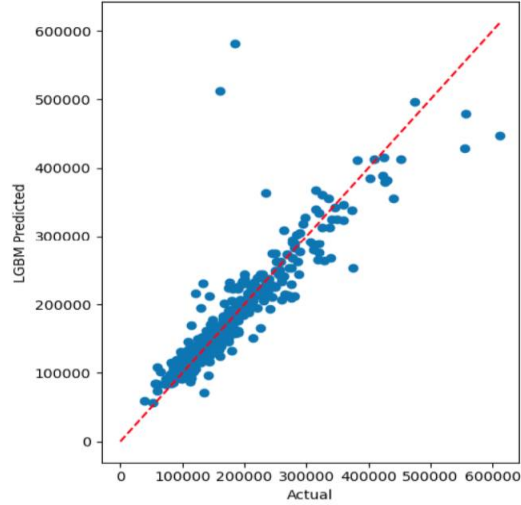
VISUALIZING THE IMPROVED PREDICTIONS

BEFORE

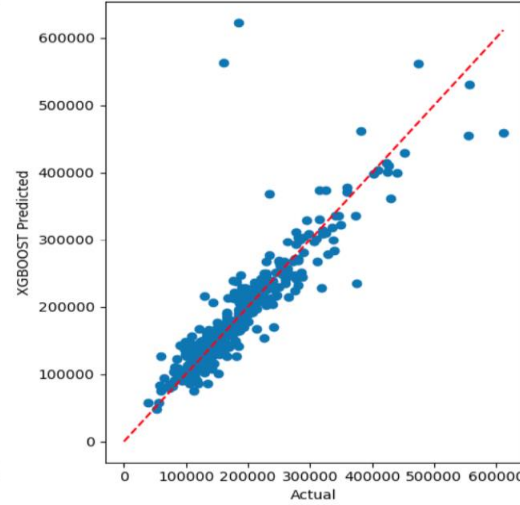
RF Predicted



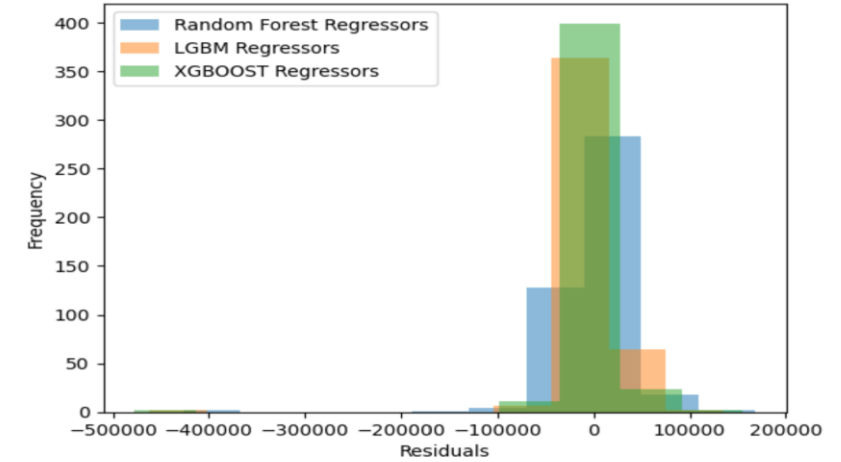
LGBM Predicted



XGBOOST Predicted

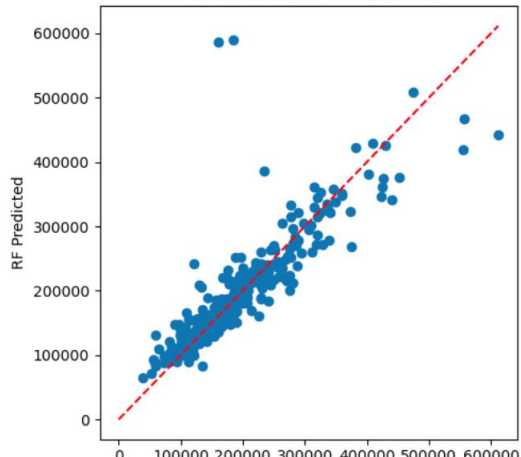


Histogram of Residuals with GridSearchCV

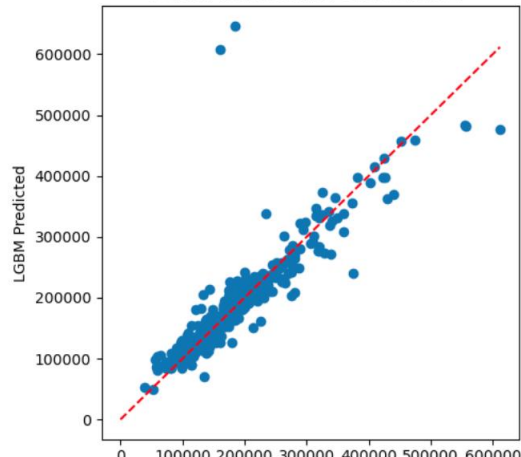


AND AFTER MODEL TUNING

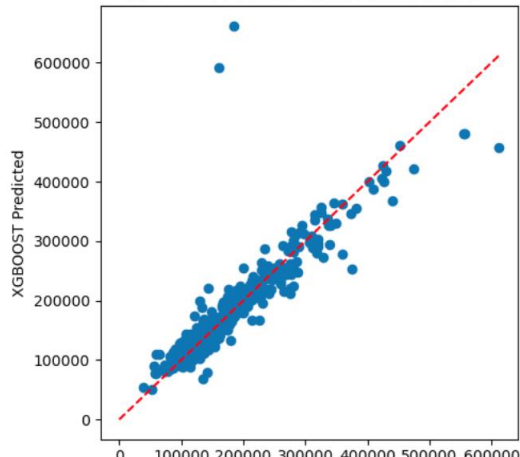
RF Predicted with GridSearchCV



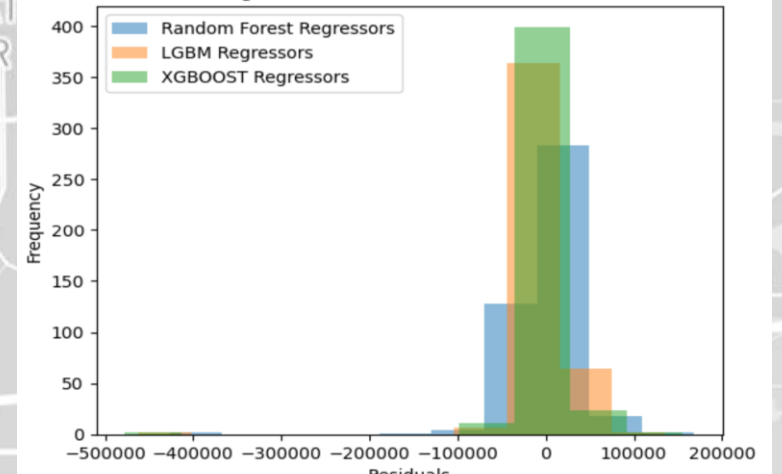
LGBM Predicted with GridSearchCV



XGBOOST Predicted with GridSearchCV



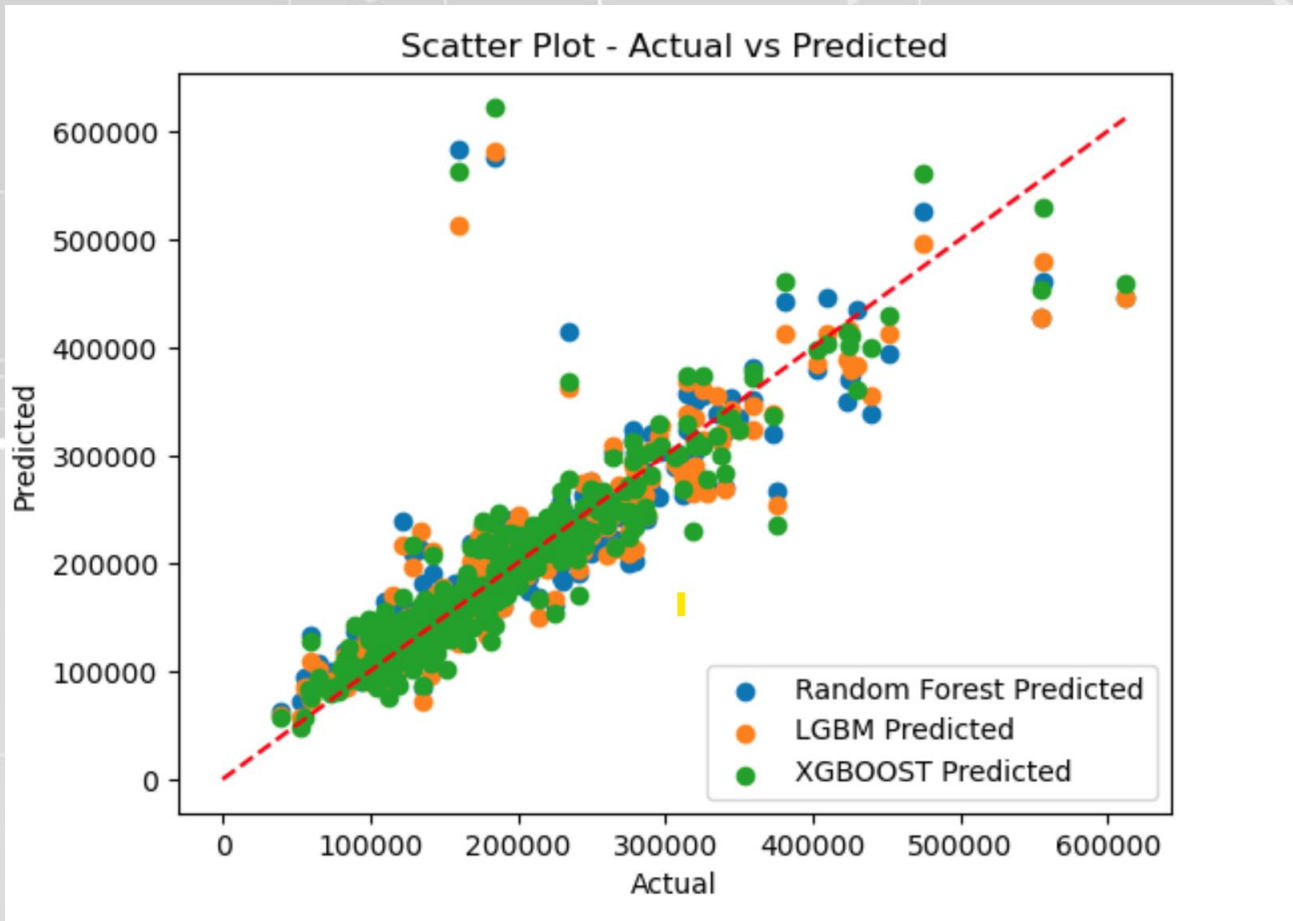
Histogram of Residuals with GridSearchCV



THE BEST MODEL

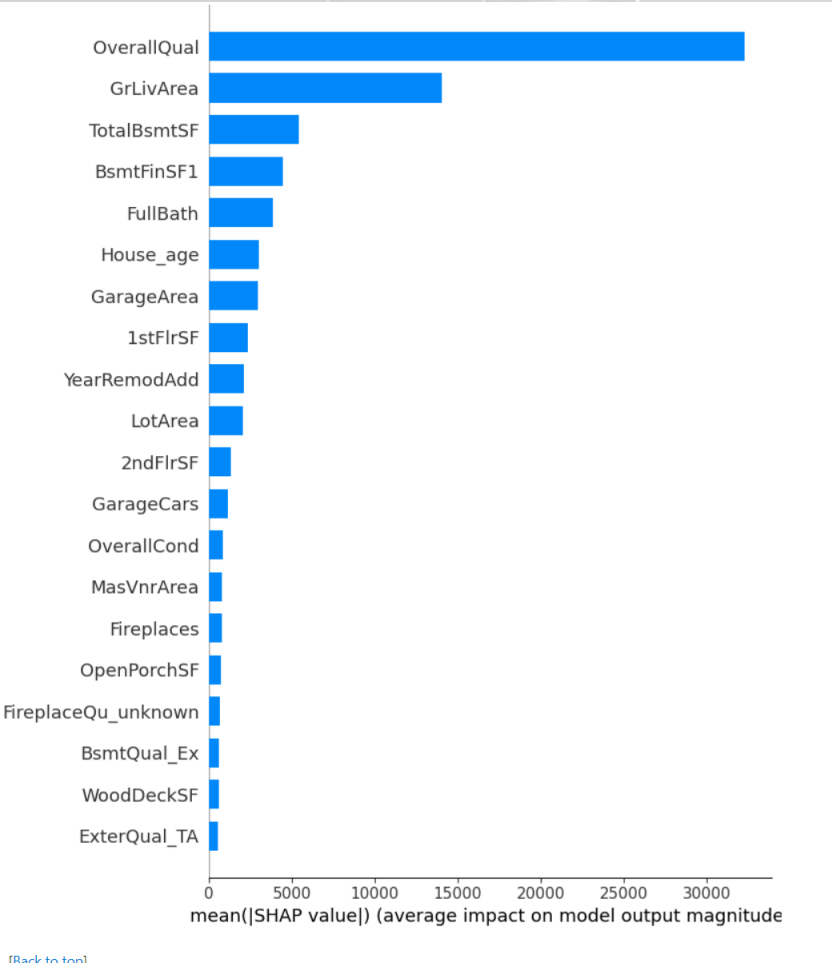
XGBRegressor with GridSearch tuning

With deviation of **\$16,718** between predicted and actual house prices.



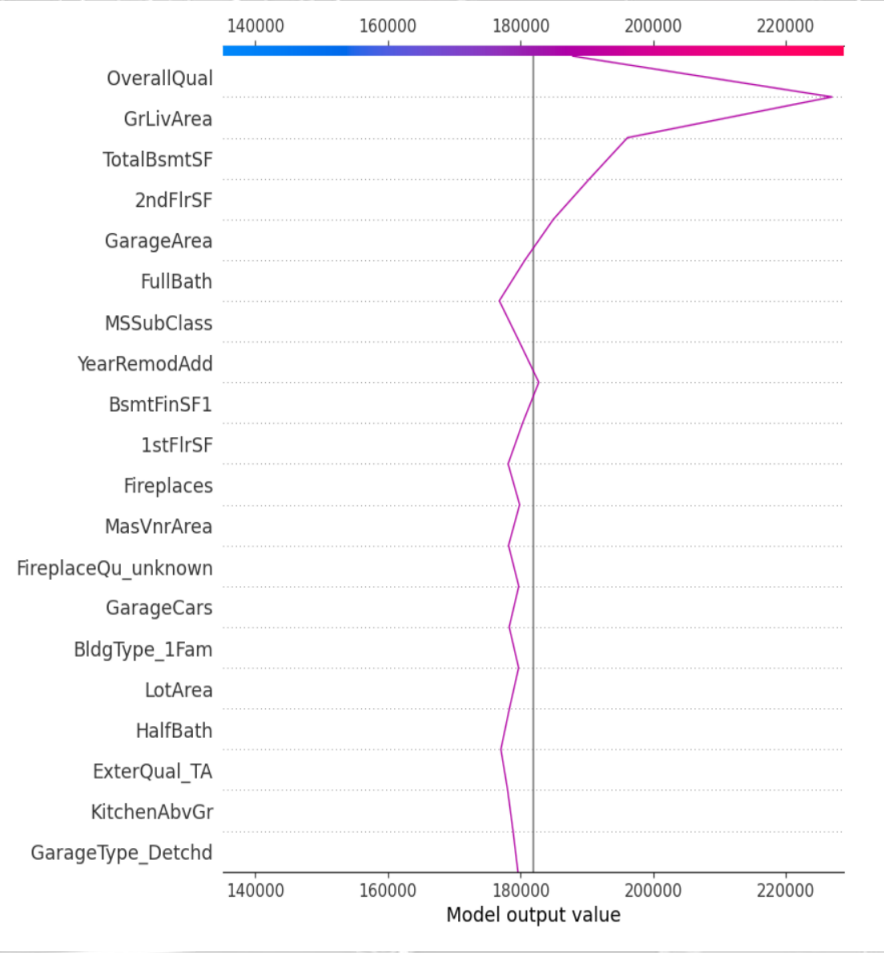
KEY FEATURES THAT ENHANCE HOUSE VALUE.

SHAP (SHapley Additive exPlanations)



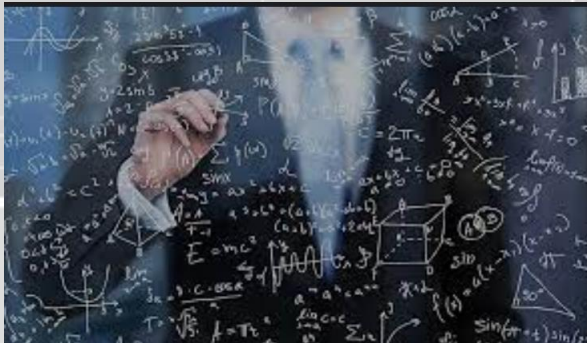
AFFORDABLE WAYS TO INCREASE HOUSE VALUE:

- **ADDING A GARAGE AREA.**
- **INSTALLING A FIREPLACE.**



FUTURE WORK

- More Data about Economic indicators, crime rates, school ratings
- Advance feature engineering techniques
- Annual updates



SUMMARY

- Analyzed three models (RandomForestRegressor, LGBMRegressor, XGBRegressor) to predict house prices.
- XGBRegressor with GridSearch tuning performed the best (MAE: \$16,718, MAPE: 11%).
- Valuable insights for real estate analysis and decision-making; SHAP force plot visualizes feature impact on prices.