Jasmeen Pabla

Student# 500578338

Supervisor Name: Tamer Abdou

Project Title: House Prices Prediction

Date 20th October 2023

Introduction:

The realm of housing price prediction has become a focal point in real estate research and investment, offering valuable insights for prospective homebuyers, property developers, and policy makers. As housing markets continue to evolve, driven by economic, social, and environmental factors, accurate predictions of property values have never been more critical. Traditional methods, such as the House Price Index (HPI), have served as foundational tools for assessing housing market trends, but their limitations are increasingly apparent. In response to these challenges, a series of studies and projects have emerged, employing various machine learning models and techniques to enhance the accuracy and scope of housing price predictions. This introduction will delve into the problem at hand, namely, the need for more robust and comprehensive housing price prediction methods, and will outline the steps taken by researchers and developers to address this issue. From exploring advanced machine learning models like Random Forest, Support Vector Regression (SVR), and Gaussian Processes (GPs) to innovative mobile applications, these initiatives aim to provide more accurate, location-specific, and flexible predictions to meet the dynamic demands of the housing market.

https://www.sciencedirect.com/science/article/abs/pii/S1051137712000228

Based on various information about the house, predict the sales price of the house. This

problem falls under Regression. The problem I am solving are What are the factors that most / least affect the sales price of a house? How does the price change based on the five most important factors? I am using the House Prices dataset that is publicly available from Kaggle.The link to the dataset is: https://www.kaggle.com/competitions/house-prices-advanced-regression-techniques

The house prices dataset contains the following fields SalePrice - the property's sale price in dollars. MSSubClass: The building class. MSZoning: The general zoning classification. LotFrontage: Linear feet of street connected to property. LotArea: Lot size in square feet Street: Type of road access. Alley: Type of alley access. LotShape: General shape of property. LandContour: Flatness of the property. Utilities: Type of utilities available. LotConfig: Lot configuration. LandSlope: Slope of property. Neighborhood: Physical locations within Ames city limits. Condition1: Proximity to main road or railroad. Condition2: Proximity to main road or railroad (if a second is present). BldgType: Type of dwelling. HouseStyle: Style of dwelling. OverallQual: Overall material and finish quality. OverallCond: Overall condition rating. YearBuilt: Original construction date. YearRemodAdd: Remodel date. RoofStyle: Type of roof. RoofMatl: Roof material. Exterior1st: Exterior covering on house. Exterior2nd: Exterior covering on house (if more than one material). MasVnrType: Masonry veneer type. MasVnrArea: Masonry veneer area in square feet. ExterQual: Exterior material quality. ExterCond: Present condition of the material on the exterior. Foundation: Type of foundation. BsmtQual: Height of the basement. BsmtCond: General condition of the basement. BsmtExposure: Walkout or garden level basement walls. BsmtFinType1: Quality of basement finished area. BsmtFinSF1: Type 1 finished square feet. BsmtFinType2: Quality of second finished area (if present). BsmtFinSF2: Type 2 finished square feet. BsmtUnfSF: Unfinished square feet of basement area. TotalBsmtSF: Total square feet of basement area. Heating: Type of heating. HeatingQC: Heating quality and condition. CentralAir: Central air conditioning. Electrical: Electrical system. 1stFlrSF: First Floor square feet. 2ndFlrSF: Second floor square feet. LowQualFinSF: Low quality finished square feet (all floors). GrLivArea: Above grade (ground) living area square feet. BsmtFullBath: Basement full bathrooms. BsmtHalfBath: Basement half bathrooms. FullBath: Full bathrooms above grade.HalfBath: Half baths above grade.Bedroom: Number of bedrooms above basement level. Kitchen: Number of kitchens.KitchenQual: Kitchen quality. TotRmsAbvGrd: Total rooms above grade (does not include bathrooms). Functional: Home functionality rating. Fireplaces: Number of fireplaces. FireplaceQu: Fireplace quality. GarageType: Garage location. GarageYrBlt: Year garage was built. GarageFinish: Interior finish of the garage. GarageCars: Size of garage in car capacity. GarageArea: Size of garage in square feet. GarageQual: Garage quality. GarageCond: Garage condition. PavedDrive: Paved driveway. WoodDeckSF: Wood deck area in square feet. OpenPorchSF: Open porch area in square feet. EnclosedPorch: Enclosed porch area in square feet. 3SsnPorch: Three season porch area in square feet. ScreenPorch: Screen porch area in square feet. PoolArea: Pool area in square feet. PoolQC: Pool quality. Fence: Fence quality. MiscFeature: Miscellaneous feature not covered in other categories. MiscVal: $Value of miscellaneous feature. MoSold: Month Sold. YrSold: Year Sold. SaleType: Type of sale. The dataset contains 1460 rows and 163 columns. I will be using python to work on my project. Since my task is to predict the house sales price I will be using regression techniques in machine learning. I will be using libraries like matplotlib for visualization, scikit-learn for building machine learning models, and pandas to read and manipulate my input data. I will be using root mean squared error

(RMSE), mean absolute error (MAE), to measure the performance of my model. I will

start with Linear Regression, Random Forest, Naive Bayes, and if needed an ensemble of the models. I will be looking at the feature importance graph from Random forest to determine the most important feature that determines the sale price of the house. Based on these features, I will carry out further analysis.

Summary of Article 1:

Housing Price Prediction via Improved Machine Learning Techniques

This paper explores the application of various machine learning models for predicting housing prices, emphasizing the significance of moving beyond traditional models and evaluating the performance of both popular and complex methods. It highlights the use of the House Price Index (HPI) as a common measure for housing price changes in different countries. The study considers models such as Random Forest, XGBoost, LightGBM, Hybrid Regression, and Stacked Generalization Regression for predicting housing prices using the "Price in Beijing" dataset. The analysis involves data preprocessing and feature engineering, including handling missing data, attribute transformation, and outlier detection. The results suggest that while Random Forest exhibits the lowest error on the training set, it is prone to overfitting, whereas XGBoost and LightGBM offer competitive accuracy with higher time complexities. Hybrid Regression, due to its generalization, outperforms these methods, but the Stacked Generalization Regression stands out as the best choice when accuracy is paramount, even though it has a more complex architecture. This paper provides valuable insights into optimizing housing price prediction models by considering various factors, models, and performance metrics. Rewrite this paragraph (Quang Truong, Minh Nguyen, Hy Dang, Bo Mei. (2020))[1].

Summary of Article 2

Predicting property prices with machine learning algorithms

In this study, led by Winky Ho and published in the Journal of Property Research in 2020, three machine learning algorithms, namely support vector machine (SVM), random forest (RF), and gradient boosting machine (GBM), were employed to assess property prices. This research utilized a dataset of approximately 40,000 housing transactions spanning over 18 years in Hong Kong, comparing the performance of these algorithms in predicting property prices. The findings revealed that RF and GBM exhibited superior predictive power and outperformed SVM in terms of key performance metrics, including mean squared error (MSE), root mean squared error (RMSE), and mean absolute percentage error (MAPE). However, the study also highlighted that SVM remains a valuable tool for producing reasonably accurate predictions under stringent time constraints. Ultimately, the research underscores the promise of machine learning as an alternative technique in property valuation and appraisal research, particularly in the context of property price prediction (Winky K.O. Hoa, Bo-Sin Tangb and Siu Wai Wong. (2020))[2].

Summary of Article 3

This research paper extensively investigates the application of various machine learning models and techniques for housing price prediction. It underscores the crucial role of the House Price Index (HPI) as a tool for measuring price changes in residential housing across different countries, emphasizing the need for additional factors beyond HPI to achieve accurate individual housing price predictions. The study introduces traditional and advanced machine learning models, including Random Forest, XGBoost, and LightGBM, for the prediction task. Notably, it explores the Hybrid Regression model, which combines multiple regression models, and Stacked Generalization Regression, an ensemble technique that leverages the predictions of other models. The paper highlights the performance differences and trade-offs among these models, considering factors such as training accuracy, time complexity, and generalization. Ultimately, it recommends selecting models based on specific priorities and provides avenues for future research toce housing price prediction models. This research significantly contributes to the understanding of further enhancing effective machine learning techniques for predicting housing prices (Wai Wong. (2022))[3]

Summary of Article 4

House Price Prediction using a Machine Learning Model: A Survey of Literature

This study conducted by Nor Zulkifley and colleagues in their 2020 paper "House Price Prediction using a Machine Learning Model: A Survey of Literature" delves into the application of data mining in the real estate market to predict house prices and identify key housing attributes. Through an analysis of 14 articles, the study suggests that locational attributes are crucial in predicting house prices, with Support Vector Regression (SVR), Artificial Neural Network (ANN), and XGBoost emerging as the most efficient models for this purpose. The authors emphasize the significance of these models in accurately predicting house prices and highlight the potential of a spatial consideration model when using the artificial neural network algorithm. They recommend further research to validate their findings in practical models, which could be of substantial benefit to housing developers and researchers seeking to enhance their understanding of house price determinants and predictive modeling in this field. The paper provides a comprehensive analysis of factors influencing house prices and the data mining techniques used for house price prediction. It emphasizes the significance of locational attributes, structural attributes, and neighborhood factors in determining house prices, with locational attributes being particularly crucial. The study reviews various prediction models, including Support Vector Regression (SVR), Artificial Neural Network (ANN), and XGBoost, and finds that these models, especially when focused on locational attributes, are the most efficient for predicting house prices. The research highlights the potential economic impact of accurate house price predictions and suggests that these models can aid investors, homebuyers, and developers in making informed decisions. However, further research is recommended to validate these findings in practical models, offering valuable insights for the real estate market. The introduction highlights the importance of housing as a fundamental human need and the increasing demand for houses. The study reviews 14 articles that examine attributes influencing house prices, with a focus on factors such as square footage, the number of bedrooms and bathrooms, and the impact of floor space on house value. The analysis emphasizes the significance of locational attributes, structural features, and neighborhood factors, revealing that locational attributes like proximity to shopping malls and hospitals play a pivotal role in determining house prices. The study also delves into various prediction models, including Multiple Linear Regression, Support Vector Regression (SVR), Artificial Neural Network (ANN), and XGBoost. It concludes that SVR, ANN, and XGBoost are the most efficient models, particularly when focused on locational attributes. These models hold great potential for predicting house prices accurately and can benefit investors, homebuyers, and developers. The paper underscores the need for further research to validate these findings and develop practical models for the real estate market. (Nor Hamizah Zulkifley, Shuzlina Abdul Rahman, Nor Hasbiah Ubaidullah, Ismail Ibrahim. (2020)).[4]

Summary of Article 5

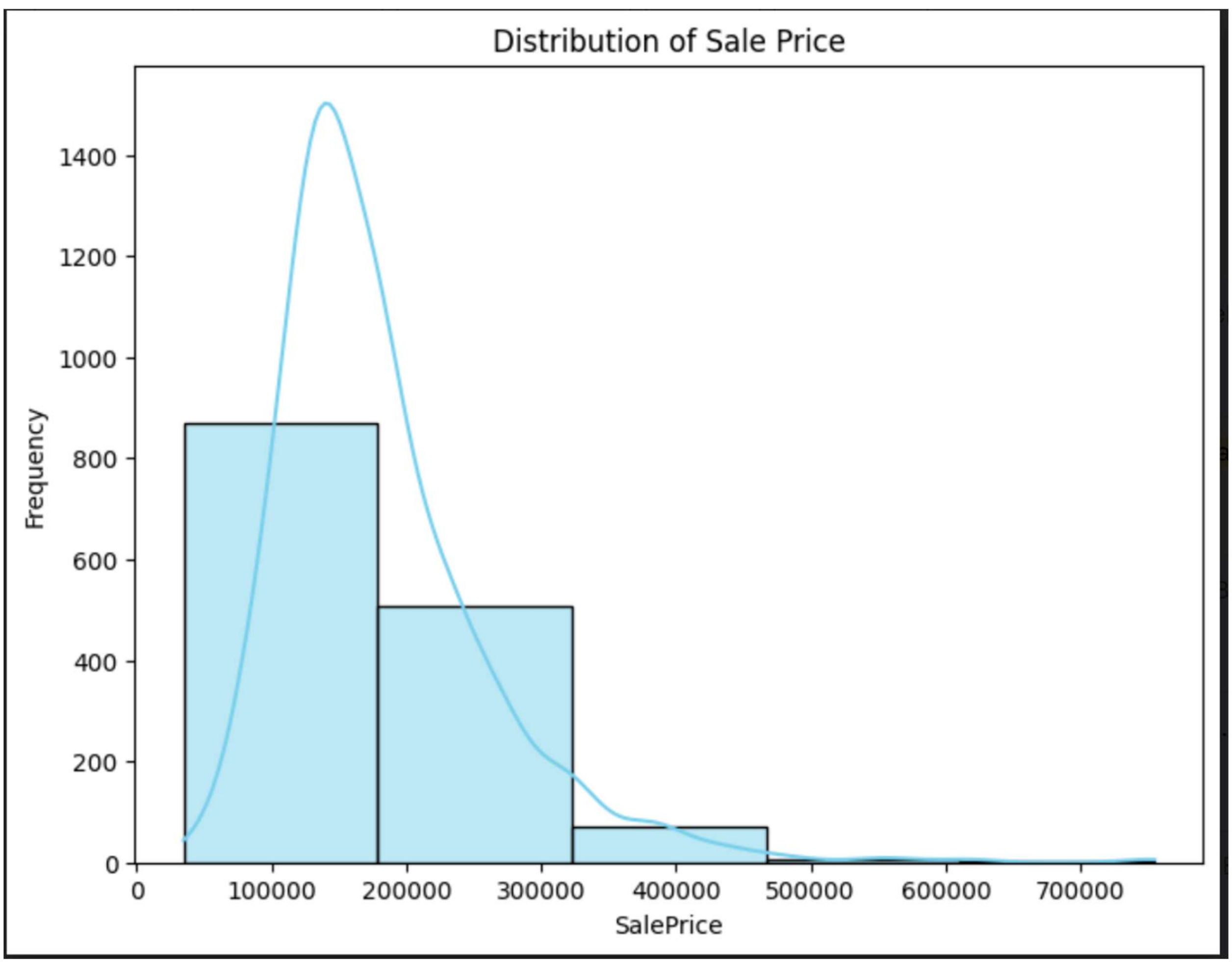
House Price Prediction using Random Forest Machine Learning Technique

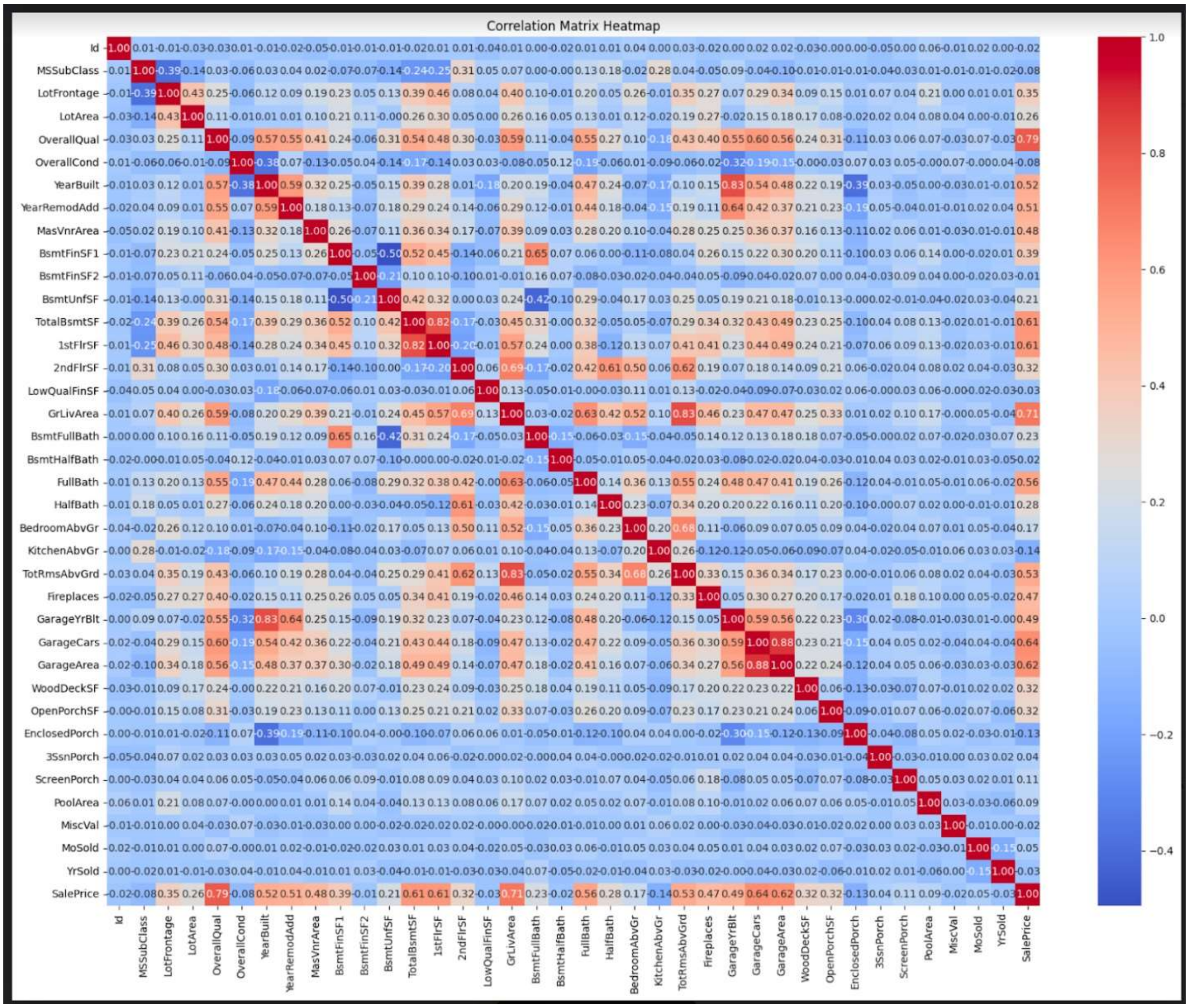
This study explores the use of the Random Forest machine learning technique for predicting house prices, addressing the limitations of traditional methods such as the House Price Index (HPI). The research emphasizes the importance of predicting price variances rather than specific values and highlights that housing prices are influenced by various factors, including location and population, which necessitate more comprehensive information for accurate predictions. By applying the Random Forest model to the UCI Machine Learning Repository's Boston housing dataset, which contains 506 entries and 14 features, the study achieves promising results. The model's predictions show an acceptable margin of error, with prices deviating by approximately ±5 from actual values, demonstrating the effectiveness of the Random Forest approach in improving house price predictions (Abigail Bola Adetunji, Oluwatobi Noah Akande, Funmilola Alaba Ajala, Ololade Oyewo, Yetunde Faith Akande, Gbenle Oluwadara. (2022))[5].

Summary of Article 6

Machine Learning for a London Housing Price Prediction Mobile Application

The project's primary objective was to create a mobile application for predicting London housing prices, employing Gaussian Processes (GPs) for regression, which offer probabilistic and flexible modeling capabilities. It managed a vast dataset of property transactions by using distributed GPs, allowing the prediction system to consider spatial structures and offload computational tasks to a server-side component. The Android-based mobile application offers location-specific predictions, heatmaps, and graphs, making it user-friendly. While successfully meeting its initial objectives, the project received positive feedback for its innovative approach but also identified areas for improvement, particularly in prediction speed and rendering quality. Overall, this project exemplifies a comprehensive and innovative approach to predicting London housing prices through a combination of advanced machine learning techniques and a user-friendly mobile application design (Aaron Ng. (2015))[6].





The correlation matrix shows that OverallQual, GrLivArea, GarageCars, GarageArea are highly positively correlated with SalePrice

As GrLivArea goes up, the SalePrice goes up too

Tentative methodology:

Exploratory analysis: Exploratory analysis of the factors that affect the sales price of a house can provide valuable insights for homeowners, real estate professionals, and investors. Here are some common factors that typically have a significant impact on house prices, from most to least influential:

Most Influential Factors:

Location: The location of a property is often the most critical factor. Proximity to schools, workplaces, public transportation, shopping centers, and desirable neighborhoods can significantly affect a home's price.

Size and Square Footage: The size of the house, including the number of bedrooms and bathrooms, as well as the overall square footage, plays a vital role. Larger homes tend to command higher prices.

Condition and Age: The age and condition of the property, including renovations and upgrades, are essential. Well-maintained and updated homes generally have higher values.

Market Conditions: The overall state of the real estate market, including supply and demand, interest rates, and economic conditions, can impact prices.

Comparable Sales (Comps): Recent sales of similar properties in the same area (comparables) are a strong indicator of a home's value.

Amenities and Features: Special features such as a pool, fireplace, modern kitchen, and energy-efficient appliances can increase a home's value.

Neighborhood Trends: Changes and trends in the neighborhood, such as gentrification or declining property values, can influence prices.

Curb Appeal: Aesthetic factors like landscaping, exterior appearance, and curb appeal can make a significant difference.

School District: Proximity to quality schools can raise property values in many areas.

Historical Sales Data: Examining historical price trends for the specific property and its neighborhood can provide insights into how it has appreciated or depreciated over time.

Least Influential Factors:

Personal Preferences: Factors that cater to individual tastes, such as interior paint colors or decorative choices, have a relatively minor impact on the overall property value.

Furniture and Personal Items: The presence of furniture and personal items, which are typically removed during a sale, doesn't significantly affect the sales price.

Seasonal Changes: Seasonal variations in the real estate market, while existing, tend to have a more modest impact than the factors listed above.

Seller's Attachment: A seller's emotional attachment to the property doesn't usually influence the market value of the house.

Mortgage Interest Rate: While mortgage rates can impact a buyer's affordability, they generally don't directly affect a home's list price.

It's important to note that the relative importance of these factors can vary depending on the local real estate market and specific circumstances. A comprehensive analysis should consider a combination of these factors to understand their collective impact on a house's sales price. Real estate professionals often use statistical methods and regression analysis to quantify the influence of these variables on home prices.

preprocessing: In a preprocessing stage for analyzing the factors that most and least affect the sales price of a house, you typically perform data cleaning, feature engineering, and statistical analysis. Here's how you can preprocess the data to identify the most and least influential factors:

Data Collection:

Gather a comprehensive dataset that includes information about house sales prices and various relevant features like location, size, condition, amenities, and more.

Data Cleaning:

Handle missing data by either imputing missing values or removing rows with missing information.

Address outliers by considering whether they are erroneous data points or genuine extreme cases.

Feature Selection:

Use statistical methods, correlation analysis, or domain expertise to identify potential influential features.

Consider including features that are likely to have a significant impact on house prices based on prior knowledge.

Feature Engineering:

Create new features if they may provide additional insights. For example, you can calculate the price per square foot, the age of the property, or distance to key amenities like schools and transportation.

Data Transformation:

Convert categorical variables into numerical format using techniques like one-hot encoding.

Normalize or scale continuous variables to ensure that they are on a similar scale.

Data Visualization:

Create data visualizations like scatter plots, histograms, and correlation matrices to explore the relationships between potential factors and house prices.

Statistical Analysis:

Perform statistical tests, such as regression analysis or ANOVA, to quantify the impact of different factors on house prices. This will help you rank the factors from most to least influential.

Feature Importance:

Utilize machine learning algorithms like random forests or gradient boosting to determine feature importance scores. This can provide insights into which factors have the most impact.

Dimensionality Reduction:

Apply techniques like Principal Component Analysis (PCA) to reduce the dimensionality of your data while retaining the most critical information.

Data Splitting:

Divide the dataset into training and testing sets to evaluate the predictive performance of your model.

Model Building:

Develop a predictive model, such as a regression model, to estimate house prices based on the selected features.

Model Evaluation:

Assess the model's performance using metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared to understand how well it predicts house prices.

By following these preprocessing steps and conducting thorough analysis, you can determine which factors are the most and least influential in affecting the sales price of a house. The final model can provide insights into the relative importance of each factor and help in making informed decisions related to real estate investments or pricing strategies.

Regenerate: The factors that most and least affect the sales price of a house are vital considerations in the real estate market. Here's a summary of the primary factors that can significantly influence house prices:

Most Influential Factors:

Location: The geographical location of the property is a paramount factor. Proximity to essential amenities, desirable neighborhoods, and low crime rates can lead to higher house prices.

Size and Square Footage: The physical size of the house, including the number of bedrooms, bathrooms, and overall square footage, plays a substantial role. Larger homes often command higher prices.

Condition and Age: The property's condition and age are crucial. Well-maintained and updated homes generally have higher values than older or poorly maintained ones.

Market Conditions: The state of the real estate market, including supply and demand dynamics, interest rates, and economic factors, can significantly impact house prices.

Comparable Sales (Comps): Recent sales of similar properties in the same vicinity (comparable sales) are strong indicators of a home's market value.

Amenities and Features: Special features like a swimming pool, fireplace, modern kitchen, and energy-efficient appliances can increase a home's value.

Neighborhood Trends: Changes and trends in the neighborhood, such as gentrification or deterioration, can influence house prices.

Curb Appeal: Aesthetic factors like landscaping, exterior appearance, and overall curb appeal can make a substantial difference in the perception of value.

School District: Proximity to high-quality schools can positively impact property values in many areas.

Historical Sales Data: Examining historical price trends for the specific property and its neighborhood can provide insights into how its value has changed over time.

Least Influential Factors:

Personal Preferences: Factors related to individual tastes, such as interior design choices and color schemes, typically have a minimal impact on overall property value.

Furniture and Personal Items: The presence of personal furniture and belongings, which are usually removed during a sale, does not significantly affect the sales price.

Seasonal Changes: While there are seasonal fluctuations in the real estate market, they tend to have a relatively minor influence on house prices.

Seller's Attachment: A seller's emotional attachment to the property usually doesn't impact its market value.

Mortgage Interest Rate: While mortgage rates can affect a buyer's affordability, they generally do not directly influence the list price of a house.

It's essential to recognize that the importance of these factors may vary depending on the local real estate market and individual circumstances. A thorough analysis, including statistical modeling, is often necessary to quantify the collective impact of these variables on house prices. Real estate professionals and data analysts commonly use regression analysis and other statistical methods to assess the relative influence of these factors on real estate pricing.

Algorithms:

When analyzing the factors that most and least affect the sales price of a house using algorithms, you can employ various machine learning techniques and models to determine the significance of each factor. Here's how you can approach this using algorithm-based analysis:

Data Preparation:

Gather a comprehensive dataset that includes information on house sales prices and relevant features such as location, size, condition, amenities, and more.

Clean and preprocess the data, addressing missing values and outliers.

Feature Engineering:

Create new features if necessary, such as age of the property, price per square foot, or distance to key amenities.

Encode categorical variables using techniques like one-hot encoding.

Feature Selection:

Utilize feature selection techniques, such as Recursive Feature Elimination (RFE) or feature importance scores, to identify influential features.

Model Selection:

Choose machine learning algorithms suitable for regression tasks. Common choices include Linear Regression, Decision Trees, Random Forests, Gradient Boosting, Support Vector Machines, and Neural Networks.

Model Training:

Split the dataset into training and testing sets to train and evaluate the model's performance.

Model Training and Evaluation:

Train the selected models on the training data and assess their performance using regression evaluation metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared.

Feature Importance Analysis:

For models like Random Forests or Gradient Boosting, extract feature importance scores to understand the relative impact of each feature on house prices.

Data Visualization:

Create visualizations, such as feature importance plots and partial dependence plots, to visualize the relationships between features and house prices.

Interpret Results:

Analyze the algorithm's output to identify the factors that most influence house prices. Features with higher importance scores or coefficients in linear regression models are typically the most influential.

Model Validation:

Validate the model's findings using domain knowledge and real estate expertise to ensure the results are logical and in line with market expectations.

Further Analysis:

Conduct sensitivity analysis to assess the robustness of the model's findings, exploring how changes in specific features affect house prices.

Optimization:

Fine-tune the model and its hyperparameters to improve predictive accuracy and the understanding of feature importance.

By applying machine learning algorithms and regression models, you can quantitatively determine the factors that have the most and least impact on house prices. This data-driven approach provides valuable insights for homeowners, real estate professionals, and investors, enabling them to make more informed decisions in the housing market.

Evaluation:

In evaluating the factors that most and least affect the sales price of a house, you typically rely on statistical analysis and regression models to quantify the impact of various features on house prices. Here's how you can evaluate these factors:

Data Collection and Preparation:

Gather a dataset containing information on house sales prices and relevant features, ensuring data quality and completeness.

Data Exploration:

Use descriptive statistics, data visualizations, and correlation analysis to get a preliminary understanding of the data and relationships between features.

Feature Selection:

Employ statistical techniques or domain expertise to select potential influential features. Factors such as location, size, and condition are commonly considered.

Statistical Analysis:

Utilize regression analysis, such as multiple linear regression, to quantify the relationship between the selected features and house prices. The coefficients of the regression equation indicate the impact of each factor.

Significance Testing:

Perform statistical tests to determine the significance of each feature. Features with low p-values are considered more influential.

Model Evaluation:

Assess the goodness of fit of the regression model using evaluation metrics like R-squared, Mean Absolute Error (MAE), and Mean Squared Error (MSE). Higher R-squared values and lower MAE/MSE indicate better model performance.

Feature Importance:

If using advanced algorithms like Random Forest or Gradient Boosting, extract feature importance scores to understand which factors contribute the most to predicting house prices.

Visualizations:

Create visualizations such as scatter plots, feature importance plots, and partial dependence plots to illustrate the relationships between features and house prices.

Sensitivity Analysis:

Conduct sensitivity analysis to explore how small changes in specific features impact house prices. This helps in understanding the robustness of the model's findings.

Expert Validation:

Consult real estate experts and domain specialists to validate the results and ensure that they align with market expectations.

Cross-Validation:

Implement cross-validation techniques to test the model's performance on different subsets of the data, ensuring that the findings are generalizable and not specific to the training data.

Interpretation:

Interpret the results to identify which factors have the most significant influence on house prices. Features with high coefficients, low p-values, and strong feature importance scores are generally considered the most influential.

Reporting:

Prepare a comprehensive report that summarizes the findings, including a list of the factors ranked from most to least influential in affecting house prices.

Continuous Improvement:

Periodically revisit the analysis as new data becomes available and as market conditions change to ensure the model's relevance.

The evaluation of factors affecting house prices is an iterative process, and the accuracy of your results depends on the quality of the data and the rigor of your analysis. By following these steps, you can quantitatively assess which factors play the most and least significant roles in determining the sales price of a house, providing valuable insights for various stakeholders in the real estate industry.

Link for the GitHub Repository:

https://github.com/pablajas/House_sales_prediction

References:

[1] Quang Truong, Minh Nguyen, Hy Dang, Bo Mei. (2020). Housing Price Prediction via Improved Machine Learning Techniques
https://www.sciencedirect.com/science/article/pii/S1877050920316318?ref=pdf_download&fr=RR-2&rr=81848f7d3da3a246

[2] Winky K.O. Hoa, Bo-Sin Tangb and Siu Wai Wong. (2020). Predicting property prices with machine learning algorithms
https://www.tandfonline.com/doi/epdf/10.1080/09599916.2020.1832558?needAccess=true

[3] Wai Wong. (2022). House Price Pridiction Using Machine Learning
https://d1wqtxts1xzle7.cloudfront.net/98410960/I7849078919-libre.pdf?1675868177=&response-content-disposition=inline%3B+filename%3DHouse_Price_Prediction_Using_Machine_Lea.pdf&Expires=1697401182&Signature=JXP3Mq4qjSDi5bnsbwLWvj8e9FKew4-Ycx~hrH~ySqLuDe9~7mBinJi11ElSkdu9DyI0fobjvI75lkLSdocMKihxmK-khbGd8MMfz9YJavOLN40eZpP4zgS5AByYboGWVqepoZpNXcuMhcuUqG-uJWSrOLSMGLAcmBYnVIR9GXP4dMQ24d0kMhRsM5b60OurADjMv~8M4BZnjC2ccAG3JjDs1kLZGuq~v1HwqwyTkIuHwuujv75WrILAPnvqcvmSbPL2RHUB69bqWu9HKTd1TSM2bJ7uybDdFcPaMgsNLPAj0ulEjSBRgbwzlProolgtwHZoT3Qvk-KX4iI1aqOOPA__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA

[4] Nor Hamizah Zulkifley, Shuzlina Abdul Rahman, Nor Hasbiah Ubaidullah, Ismail Ibrahim. (2020). House Price Prediction using a Machine Learning Model: A Survey of Literature
https://www.mecs-press.org/ijmecs/ijmecs-v12-n6/IJMECS-V12-N6-4.pdf

[5] Abigail Bola Adetunji, Oluwatobi Noah Akande, Funmilola Alaba Ajala, Ololade Oyewo, Yetunde Faith Akande, Gbenle Oluwadara. (2022). House Price Prediction using Random Forest Machine Learning Technique,
https://www.sciencedirect.com/science/article/pii/S1877050922001016?ref=pdf_download&fr=RR-2&rr=8184940a6d9939e1

[6] Aaron Ng. (2015).Machine Learning for a London Housing Price Prediction Mobile Application http://www.doc.ic.ac.uk/~mpd37/theses/2015_beng_aaron-ng.pdf