## A red and white sign Description automatically generated with low confidence

## Artificial Intelligence and Machine Learning

Project Report

Semester-IV (Batch-2022)

**HOUSAGE**

**Supervised By: Submitted By:**

Tushar khitoliya Sir Lovish Bansal,2210992584,G-30  
 Keshav Gupta,2210990505,G-30  
 Kushagra Juneja,2210990533,G-30

Augustya,2210990196,G-30

**Department of Computer Science and Engineering**

**Chitkara University Institute of Engineering & Technology,Chitkara University, Punjab**

**Abstract**

* Develop a predictive model for house prices using machine learning.
* Explore linear regression, decision trees, random forests, and gradient boosting.
* Assess models using mean squared error, mean absolute error, and R-squared.
* Random forests and gradient boosting demonstrate superior performance.
* Identify significant factors like location and property size impacting prices.
* Offer valuable insights for real estate stakeholders to make informed decisions.
* Handle missing values, encode features, and normalize numerical data.

**Table of Contents**

* Introduction
* Problem Definition
* Requirements
* Proposed Design / Methodology
* Results

**Introduction**

* House price prediction is pivotal for buyers, sellers, and agents to make informed decisions in real estate transactions.
* Conventional approaches relying on historical data and expert knowledge may overlook the complexity of market dynamics.
* Machine learning offers advanced analysis of extensive datasets, enabling precise predictions by uncovering intricate patterns and relationships.
* This research aims to leverage machine learning techniques to develop robust predictive models for accurate house price forecasting, providing stakeholders with actionable insights to navigate the real estate market effectively.

**Problem Definition**

* House price prediction is crucial for real estate stakeholders to make informed decisions, such as buying, selling, or investing.
* Traditional methods often rely on historical data and expert opinions, which may not adequately capture the complexities of the housing market.
* Leveraging machine learning techniques can enhance prediction accuracy by analyzing diverse datasets, including property features and market trends.
* Develop a robust predictive model using machine learning to accurately forecast house prices, thereby empowering stakeholders with actionable insights for navigating the real estate market.

**Requirements**

* Access to comprehensive datasets containing relevant information such as property features, location attributes, economic indicators, and historical sales data.
* Ensure data accuracy, consistency, and completeness by addressing issues such as missing values, outliers, and data inconsistencies.
* Identify key features influencing house prices through exploratory data analysis and domain knowledge.
* Evaluate and select appropriate machine learning algorithms such as linear regression, decision trees, random forests, or gradient boosting for prediction tasks.
* Train machine learning models using historical data to learn patterns and relationships between input features and house prices.
* Assess the performance of trained models using metrics such as mean squared error, mean absolute error, and R-squared to ensure prediction accuracy.
* Implement the trained model into a production environment for real-time house price prediction.

**Proposed Design**

* Gather comprehensive datasets containing property features, economic indicators, and historical sales data. Cleanse the data by addressing missing values and outliers, and perform feature engineering to extract relevant features.
* Identify key features influencing house prices and select the most relevant ones for model training. Evaluate and select appropriate machine learning algorithms such as linear regression, decision trees, or gradient boosting for prediction tasks.
* Split the dataset into training and validation sets. Train the selected machine learning model using the training data and optimize model hyperparameters. Evaluate model performance using metrics like mean squared error and R-squared on the validation set to ensure accuracy.
* Deploy the trained model into a production environment for real-time house price prediction. Implement monitoring tools to track model performance and data quality issues. Regularly update the model with new data and retrain as necessary to maintain accuracy over time.

**Results**

* Evaluation metrics such as mean squared error, mean absolute error, and R-squared demonstrate the predictive accuracy of the developed model.
* Compare the performance of different machine learning algorithms tested during the study, highlighting the strengths and weaknesses of each approach.
* Identify the most influential features affecting house prices, providing insights into the factors driving real estate market dynamics.
* Demonstrate the practical utility of the developed model in assisting stakeholders, such as buyers, sellers, and agents, in making informed decisions in the real estate market.