Project Proposal: House Price Prediction in King County, Seattle

Objective:

The primary objective of this project is to develop a model capable of accurately estimating house prices in King County, Seattle, using a relevant set of features. Leveraging machine learning techniques, our aim is to provide valuable insights to empower homebuyers, sellers, and real estate professionals in making informed decisions within the dynamic housing market.

Dataset Overview:

The selected dataset, sourced from OpenML, comprises house sale prices for King County, Seattle, spanning transactions from May 2014 to May 2015. With a comprehensive size of 2.31 MB and dimensions (21,653 x 22), the dataset is structured in ARFF format and includes crucial attributes such as house sale prices, bedrooms, bathrooms, living area size, lot size, and various other pertinent features.

Problem Statements:

- **Prediction Accuracy:** Develop a machine learning model that ensures high prediction accuracy, minimizing the margin of error.
- **Feature Importance:** Identify and understand the most influential features that significantly impact house prices, fostering a deeper comprehension of the underlying factors.
- **Model Interpretability:** Ensure the developed model is interpretable, providing stakeholders with a clear understanding of the rationale behind the predicted house prices.
- **Deployment Consideration:** Address challenges and considerations related to model deployment, ensuring seamless integration into practical real-world scenarios.

Planning and Analyzing stage

Milestone	Tasks	Outcome/Deliverable	Estimated Time
Milestone 1	• Outline project workflow • Select relevant dataset • Write project proposal	• Stakeholders updated	Sunday, 10
Milestone 2	• Data exploration and cleaning	• Data ready for modeling	Wednesday, 13

Constructing and Executing stage

Milestone 3	• Finalize modeling strategy • Build machine learning model • Test models for accuracy	• Machine learning model • Stakeholders updated	Wednesday, 20
Milestone 4	• Finalize results • Share findings with stakeholders • Incorporate feedback	• Visualization • Executive summary • Presentation results	Sunday, 24