

WORKSHOP I



Universidad Distrital Francisco José de Caldas

FACULTY OF ENGINEERING

Authors

Juan David Zárate Moya 20222020184

Jesús Mateo Munevar Mendez 2023202042

Kevin David Rincon Valencia 20232020356

Juan David Romero Morales 20222020102

Teacher

Carlos Andrés Sierra Virguez

Bogotá D.C

2025

Contents

| | | |
|-----|--|---|
| 1 | Competition Overview | 1 |
| 2 | Systemic Analysis | 1 |
| 2.1 | Key Elements | 1 |
| 2.2 | Relationships and Interactions | 1 |
| 2.3 | System Boundaries | 1 |
| 2.4 | Sensitivity and Complexity | 2 |
| 2.5 | Chaos and Unpredictability | 2 |
| 2.6 | Diagram: System Elements and Relationships | 2 |

1 Competition Overview

The Kaggle competition *House Prices: Advanced Regression Techniques* (<https://www.kaggle.com/competitions/house-prices-advanced-regression-techniques>) is designed to predict the sale price of residential houses in Ames, Iowa, using a structured dataset.

The dataset consists of two main files:

- **train.csv**: includes 1460 observations with 79 explanatory variables and the target variable **SalePrice**.
- **test.csv**: includes 1459 observations with the same features but without the target.

The explanatory variables include numerical features (e.g., lot area, ground living area, year built), categorical variables (e.g., neighborhood, house style), and ordinal assessments (e.g., overall quality). These represent structural, spatial, and qualitative aspects of housing.

The evaluation metric is the Root Mean Squared Error (RMSE) on the logarithm of the predicted price versus the true price:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(y_i) - \log(\hat{y}_i))^2}.$$

2 Systemic Analysis

2.1 Key Elements

The competition can be framed as a simplified representation of a real estate system. The dataset is one part of a larger socio-technical environment, where:

- **Dataset**: a curated sample of properties in Ames.
- **Features**: physical attributes (area, rooms, year built), qualitative assessments (quality, condition), and locational aspects (neighborhood).
- **Target**: **SalePrice**, representing market value at the moment of sale.
- **External Factors (not included)**: interest rates, macroeconomic conditions, urban planning policies, school district quality, demographic shifts.
- **Stakeholders**: homeowners, buyers, real estate agents, banks, city planners.

2.2 Relationships and Interactions

The system involves multiple relationships:

- Physical features (e.g., living area, construction year) directly affect the perception of house quality.
- Location features (e.g., neighborhood) are proxies for external socioeconomic conditions, influencing demand and price.
- Qualitative scores (e.g., OverallQual) reflect subjective judgments, introducing variability and bias.
- Market price (**SalePrice**) is an emergent property, influenced not only by physical attributes but also by social and economic feedback loops.

2.3 System Boundaries

The dataset defines a closed system with fixed features. However, the real-world housing system is open and dynamic:

- It interacts with broader economic cycles (e.g., financial crises, inflation).
- It responds to policy decisions (e.g., zoning laws, subsidies).
- It exhibits feedback loops: rising prices attract more investment, which further increases prices.

2.4 Sensitivity and Complexity

The housing system is sensitive to both internal and external variations:

- A small change in interest rates can alter buyer demand significantly.
- Missing values or noisy variables (e.g., LotFrontage) can change model predictions considerably.
- Nonlinear relationships exist: e.g., additional square footage may increase price up to a point, but not indefinitely.

2.5 Chaos and Unpredictability

Unpredictable dynamics arise from:

- Economic shocks (e.g., 2008 housing crisis).
- Behavioral factors (e.g., speculation, buyer psychology).
- Local feedback loops (e.g., a prestigious school district increases demand, further raising prices).

2.6 Diagram: System Elements and Relationships

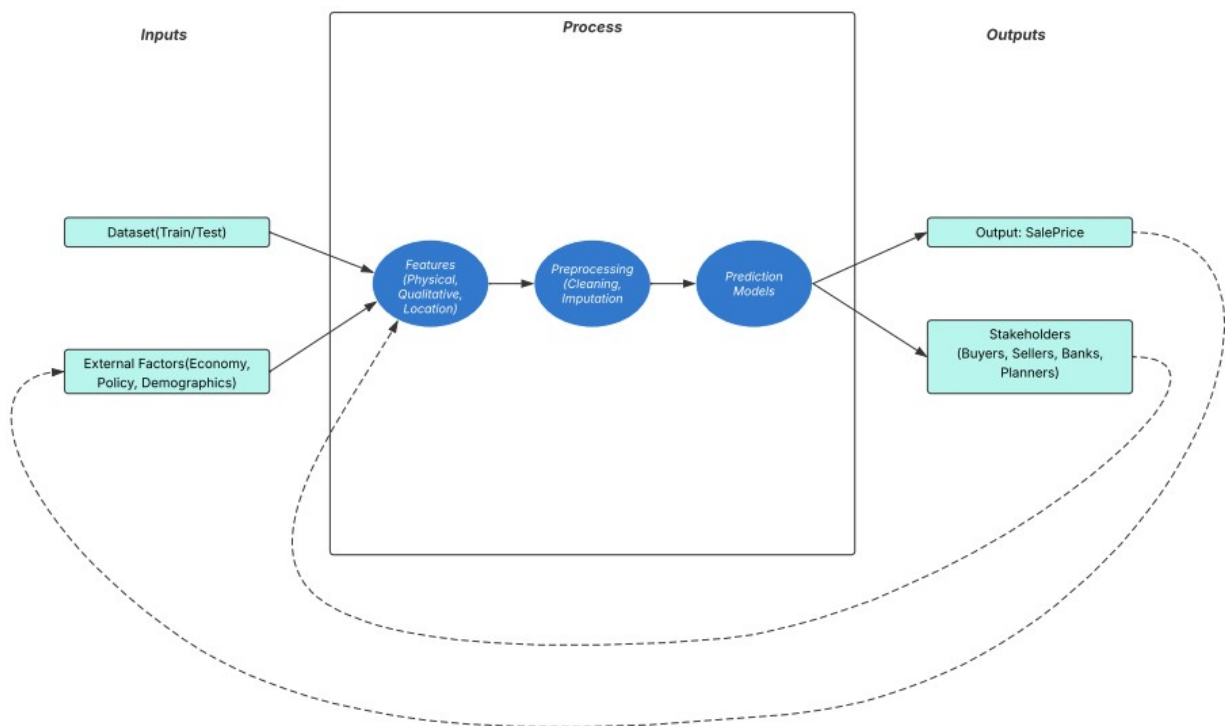


Figure 1: Diagram of system elements, processes, and feedback loops for housing price prediction.

References

- Blanchard, B. S., & Fabrycky, W. J. (2014). *Systems Engineering and Analysis* (5th ed.). Pearson.
- Kaggle. (2021). *House Prices: Advanced Regression Techniques*. Retrieved from <https://www.kaggle.com/competitions/house-prices-advanced-regression-techniques>