



# Visualizing Housing Market Trends: An Analysis of Sale Prices and Features Using Tableau and Machine Learning

Yaragorla Snehalatha<sup>1</sup>, Gantipalli Hemanth<sup>2</sup>, Vemapati Bhargav<sup>3</sup>, Shaik Maneesha<sup>4</sup>,  
K.Sirisha M. Tech<sup>5</sup>

U.G. Student, Department of CSE (Artificial Intelligence & Machine Learning), Kallam Haranadhareddy Institute of  
Technology, Guntur, Andhra Pradesh, India<sup>1-4</sup>

Assistant Professor, Kallam Haranadhareddy Institute of Technology, Guntur, Andhra Pradesh, India<sup>5</sup>

**ABSTRACT:** The real estate industry produces large amounts of data related to property prices, house features, and location characteristics. Understanding these trends requires advanced data analytics and visualization techniques. This project presents a housing market analytics system that integrates Machine Learning models with Tableau visualization dashboards to analyze housing market trends and predict house prices.

Housing datasets containing features such as number of bedrooms, number of bathrooms, house area, lot size, and year built are collected and preprocessed. Tableau dashboards are used to visualize housing market insights such as price distribution, location-based price variations, and property feature comparisons.

A Random Forest Regression model is implemented to predict house prices based on property characteristics. The system is deployed through a Flask-based web application where users can explore interactive dashboards and predict house prices dynamically.

The results show that integrating machine learning prediction with interactive visualization tools significantly improves housing market analysis and supports data-driven decision making in real estate.

**KEYWORDS:** Housing Market Analysis, Tableau, Data Visualization, Machine Learning, Random Forest Regression, Dashboard, Flask Web Application..

## I. INTRODUCTION

The housing market is a major component of the global economy. Property prices are influenced by several factors such as house size, number of bedrooms, location, and construction year. Understanding these relationships is important for buyers, sellers, and real estate investors.

Traditional analysis methods struggle to interpret large housing datasets effectively. Data visualization tools such as Tableau can transform raw housing data into meaningful visual insights that help identify trends and relationships. Machine learning techniques further enhance housing analysis by enabling predictive models that estimate house prices based on property characteristics.

This project proposes a **data-driven housing market analytics system** that combines machine learning prediction with interactive Tableau dashboards. The system allows users to analyze housing trends visually and predict property prices based on housing features.

## II. LITERATURE SURVEY

Many research studies have applied data analytics and machine learning techniques in real estate market analysis. Data visualization tools such as Tableau are widely used to explore large datasets and present insights through interactive dashboards.

Previous research shows that visualization techniques help identify housing market trends such as price distribution, regional price differences, and property demand patterns.

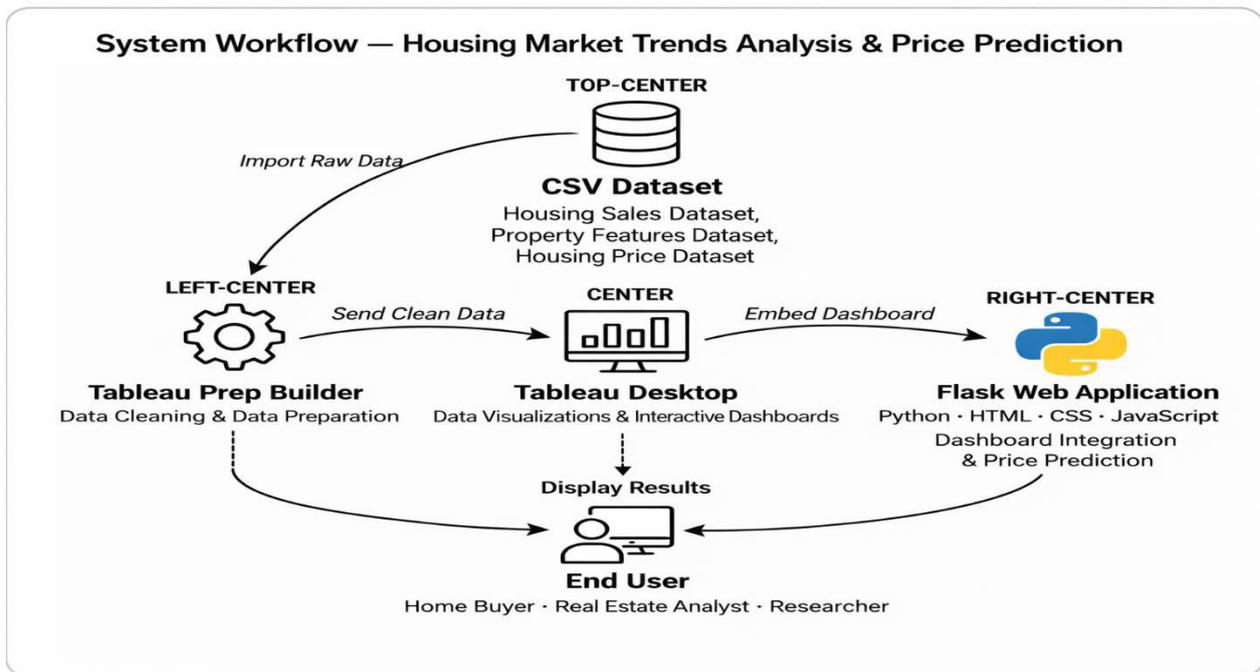
Machine learning algorithms are also commonly used in real estate analysis for predicting house prices. Regression-based models such as **Linear Regression, Decision Trees, and Random Forest** are widely applied for price prediction tasks.

Among these models, **Random Forest Regression** is considered effective because it combines multiple decision trees to improve prediction accuracy and reduce overfitting.

However, many existing systems focus either on predictive modeling or visualization separately. This project integrates **machine learning prediction with visualization dashboards** in a single web-based platform for comprehensive housing market analysis.

### Technologies Used

Technology	Purpose
Python	Core programming language for ML and backend development
Flask	Web framework connecting ML model with user interface
Scikit-learn	Building Random Forest regression model
Tableau Public	Creating interactive dashboards for housing market visualization
Pandas	Data cleaning and preprocessing
NumPy	Numerical computations and feature processing
HTML	Web page structure
CSS	User interface styling
JavaScript	Frontend interactivity
Jinja2	Dynamic HTML rendering
Pickle	Saving trained ML model
CSV Dataset	Storing housing market data
Tableau Prep Builder	Data cleaning and preparation



### III. PROPOSED METHODOLOGY

The proposed system integrates **data analytics, machine learning, and visualization tools** to analyze housing market trends and predict house prices.

#### Data Collection

Housing datasets were collected from publicly available real estate data sources. The dataset contains information about various housing properties including:

- Sale Price
- Number of Bedrooms
- Number of Bathrooms
- House Area (Square Feet)
- Lot Size
- Year Built
- Location

These features are used for both visualization analysis and machine learning prediction.

#### Data Preprocessing

Before performing analysis, the dataset is cleaned and prepared.

Data preprocessing steps include:

- Removing missing values
- Eliminating duplicate records
- Standardizing measurement units
- Handling inconsistent data entries
- Organizing the dataset into structured format

The cleaned dataset is then used for visualization and machine learning modeling.

### Data Visualization

Once the datasets are cleaned, they are imported into Tableau Desktop for visualization. Tableau provides powerful tools for transforming raw data into graphical representations.

Different visualization techniques were used in this project including:

- **Bar Charts** to compare house prices across
- **Line Charts** to analyze housing price trends over
- **Scatter Plots** to study the relationship between house size and price
- **Pie Charts** to show property type distribution

These visualizations help users understand patterns and relationships in the housing market.

### Dashboard Creation

After creating individual charts and graphs, they are combined into interactive dashboards in Tableau. The dashboard integrates multiple visualizations into a single interface that allows users to analyse housing market trends and property price patterns.

The dashboard includes features such as:

- Housing price distribution
- Bedroom vs price comparison
- Location-based price analysis
- House size vs sale price relationship

Interactive filters and navigation options are added to allow users to explore the data dynamically.

### Machine Learning Model

A **Random Forest Regression model** is used to predict housing prices.

The input features include:

- Number of Bedrooms
- Number of Bathrooms
- House Area
- Lot Size
- Year Built

The Random Forest algorithm constructs multiple decision trees and combines their outputs to produce accurate predictions.

### Model Evaluation Metrics

The performance of the machine learning model is evaluated using standard regression metrics:

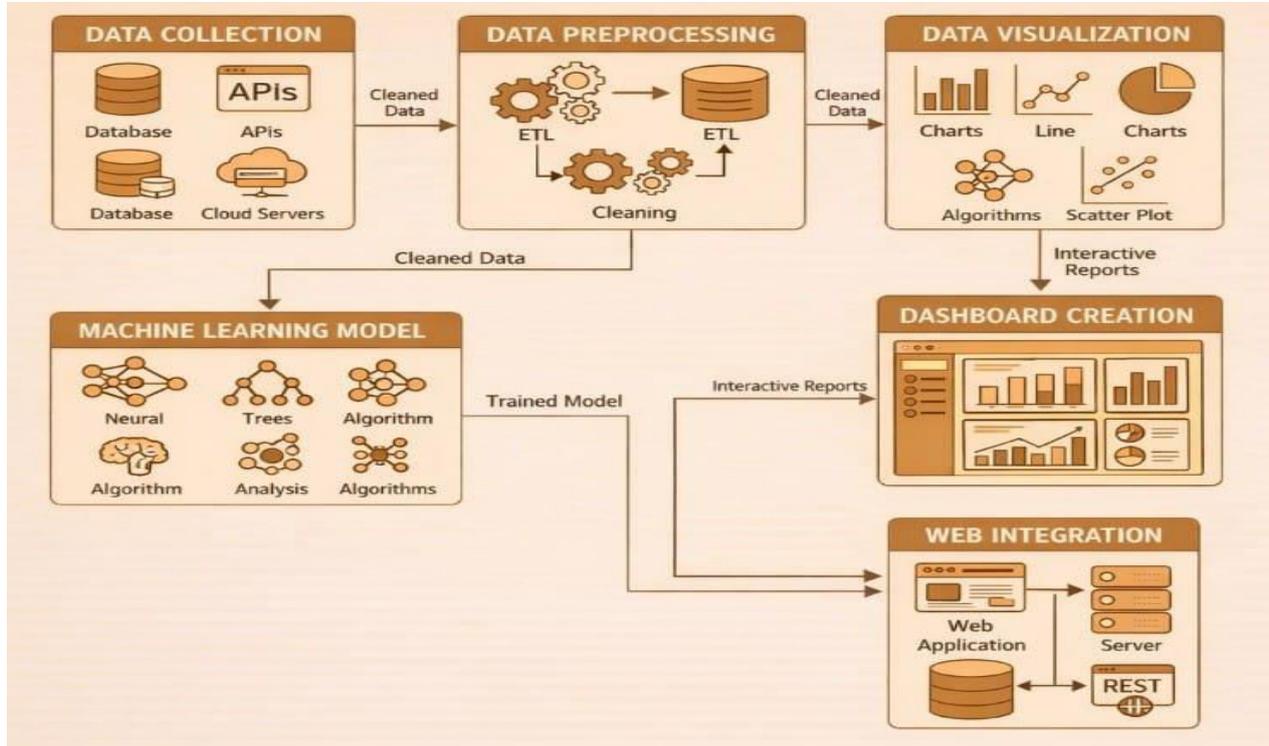
- **R<sup>2</sup> Score** – Measures prediction accuracy
- **Mean Absolute Error (MAE)** – Measures average prediction error.

### Web Integration

In the final stage, the Tableau dashboards are published to Tableau Public and integrated into a web application. The web interface is developed using HTML for page structure, CSS for styling, and Python with the Flask framework for backend functionality.

The dashboards are embedded into the web pages using embed links or iframe integration, allowing users to view and interact with the visualizations through a web browser.

This integration enables users to explore **housing market insights** through interactive dashboards and perform **house price predictions based on property features** within a single web application..



### Methodology Architecture

## IV. EXPERIMENTAL RESULTS

The proposed **housing market analytics system** was evaluated using a dataset containing property characteristics such as **number of bedrooms, number of bathrooms, house area (square feet), lot size, and year built**. These features were used to train a **Random Forest Regression model** to predict housing sale prices.

The Random Forest model was trained using the training dataset and evaluated on the testing dataset using standard regression performance metrics.

### Model Evaluation

Evaluation Metric	Description	Model Performance Value
<b>R<sup>2</sup> Score (R-Squared)</b>	Measures how well the independent variables explain the variation in housing prices..	<b>0.87 (87%)</b>
<b>MAE (Mean Absolute Error)</b>	Measures the average difference between actual house prices and predicted house prices.	<b>52,000</b>
<b>RMSE (Root Mean Squared Error)</b>	Measures the squared difference between predicted and actual prices, penalizing larger errors.	<b>81,000</b>

**Feature Importance Analysis**

This table presents the key housing attributes that have the greatest influence on **house price prediction**. Feature importance is calculated using the Random Forest algorithm, which determines how strongly each variable contributes to predicting housing prices.

Feature	Impact on EV Price
House Area (Square Feet)	High Impact – Larger houses generally have higher prices
Number of Bedrooms	Significant Impact – More bedrooms increase property value
Number of Bathrooms	Moderate Influence – Houses with more bathrooms tend to have higher prices
Year Built	Moderate Impact – Newer houses usually have higher market value
Lot Size	Low to Moderate Influence – Larger property land increases value

**Tableau Dashboard Insights**

Several key insights were obtained from the Tableau dashboards:

Analysis Area	Key Insight
Price Distribution	Houses with larger floor areas generally have higher sale prices
Bedroom Analysis	Properties with more bedrooms tend to have higher market value
Location Analysis	Urban areas and developed regions show higher housing prices
Property Size vs Price	House area has a strong positive relationship with sale price

These visual insights help users better understand housing market patterns and factors influencing property prices.

**User Interface Evaluation**

The usability and clarity of the web interface were evaluated in this project. The web application provides simple and intuitive navigation, enabling users to easily explore the housing market analysis results. It integrates Tableau dashboards with the machine learning prediction model to present both visual insights and predictive outcomes.

The interactive dashboards display housing market trends such as property prices, number of bedrooms, house size, and location-based comparisons. These visualizations help users understand patterns and relationships within the housing dataset. The interface also allows users to enter property features such as number of bedrooms, number of bathrooms, house area, and year built to predict the house price using the Random Forest regression model. Overall, the system is user-friendly and interactive, effectively supporting housing market data analysis and house price prediction.



Fig 1: Home Page

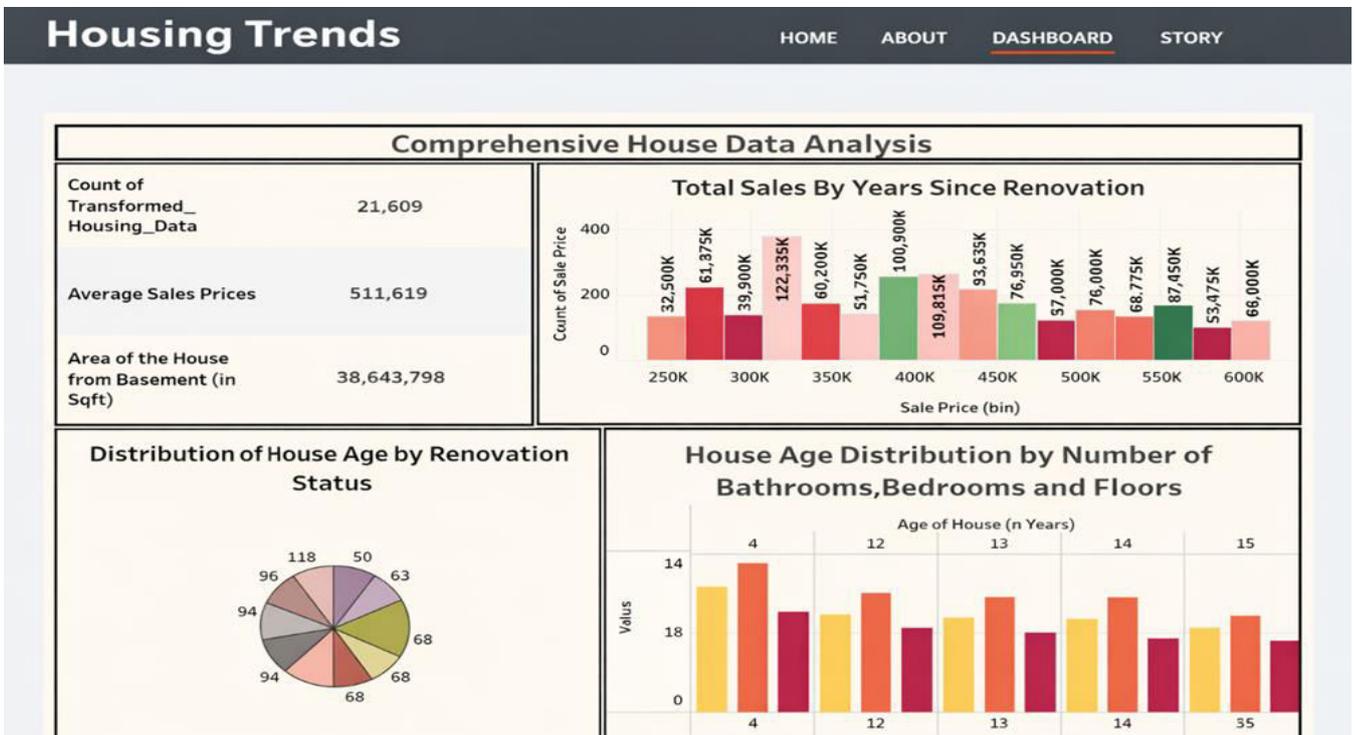


Fig 2: Dashboard Page

## Housing Trends

HOME ABOUT DASHBOARD STORY

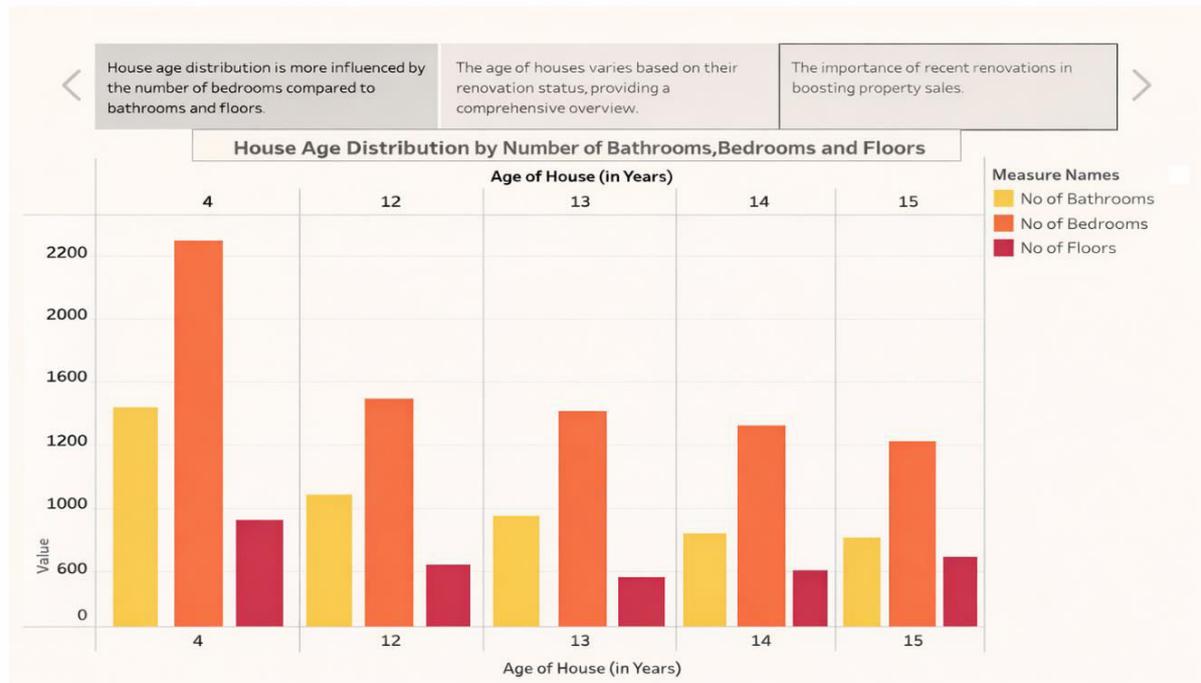


Fig 3: Story Page

### prediction for house prices

Enter the required property features to predict the estimated house price.

**Number of Bedrooms:**

**Number of Bathrooms:**

**House Area (sq ft):**

**Lot Size (sq ft):**

**Year Built:**

Fig 4: Prediction Page

## V. DISCUSSION

The results show that combining machine learning algorithms with visualization tools provides valuable insights into housing market trends. The Random Forest regression model demonstrated strong predictive performance because it

## Volume 15, Issue 3, March 2026

**|DOI: 10.15680/IJIRSET.2026.1503178|**

effectively captures complex relationships between property features such as house area, number of bedrooms, bathrooms, and construction year with housing prices.

The integration of Tableau dashboards significantly improves the interpretability of the results by presenting insights through interactive charts, graphs, and visual stories. These visualizations help users easily understand patterns such as price distribution, location-based variations, and the relationship between property size and sale price.

This combination of predictive analytics and visualization allows users, real estate analysts, and potential buyers to understand housing market behavior more clearly and make data-driven decisions regarding property investments and housing price evaluation.

### VI. CONCLUSION

This project presented a data-driven housing market analysis system that integrates machine learning prediction with interactive data visualization dashboards. The system analyzes housing market data and identifies important relationships between property features such as house area, number of bedrooms, bathrooms, and construction year with housing prices.

The Random Forest regression model demonstrated reliable prediction performance in estimating house prices based on property characteristics. At the same time, Tableau dashboards provided clear and interactive visualizations that help users understand housing market trends, price distributions, and the impact of different housing features.

The integration of machine learning with visualization tools enhances the overall analysis by combining predictive capabilities with graphical insights. The Flask-based web application enables users to explore dashboards and perform house price predictions through a user-friendly interface.

Overall, the project demonstrates that combining machine learning and visualization technologies can significantly improve the analysis of housing market data and support better decision-making in real estate applications.

### VII. FUTURE SCOPE

#### 1. Integration with Real-Time Data Sources:

In future work, the system can be connected with real-time housing market data obtained from real estate websites or public APIs. This will enable automatic updates of the dashboard and provide users with more accurate and current insights about housing price trends and market conditions.

#### 2. INCORPORATION OF PREDICTIVE ANALYTICS MODELS:

Advanced machine learning techniques such as Linear Regression, Decision Trees, or Random Forest algorithms can be incorporated to predict future housing prices. This enhancement would help investors, buyers, and analysts make informed decisions based on predicted market trends.

#### 3. Expansion of Dataset and Geographic Coverage:

The project can be extended by including larger datasets that cover multiple cities, regions, or countries. A broader dataset would allow comparative analysis of housing market trends across different geographical locations and improve the depth of the analysis.

#### 4. Development of Advanced Interactive Dashboards:

Future improvements can include more advanced visualization techniques such as dynamic filters, drill-down analysis, and user-controlled parameters in Tableau. These features would allow users to explore the data in greater detail and gain deeper insights into housing market patterns.

#### 5. Deployment as a Complete Web-Based Analytics System:

The project can be further developed into a fully functional web-based analytics platform using technologies such as

**Volume 15, Issue 3, March 2026****|DOI: 10.15680/IJIRSET.2026.1503178|**

Flask, HTML, CSS, and JavaScript. This would allow users to access the dashboards online and interact with the visualizations through a user-friendly interface.

**6. Cloud Deployment and Mobile Accessibility:**

The system can also be deployed on cloud platforms to improve accessibility and scalability. With cloud integration, users will be able to access the housing market dashboards through mobile devices and web browsers from any location, enhancing usability and convenience.

**VIII. ACKNOWLEDGMENT**

The authors would like to express their sincere gratitude to the faculty members and project supervisors of the Department of Computer Science and Engineering for their guidance and support throughout the development of this project.

**REFERENCES**

1. J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, 2011.
2. G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, Springer, 2013.
3. L. Breiman, "Random Forests," *Machine Learning Journal*, vol. 45, no. 1, pp. 5–32, 2001.
4. A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, O'Reilly Media, 2019.
5. Tableau Software, "Tableau Public Documentation." Available: <https://public.tableau.com/>
6. Flask Documentation, "Flask Web Framework." Available: <https://flask.palletsprojects.com/>
7. F. Chollet, *Deep Learning with Python*, Manning Publications, 2018.
8. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
9. S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed., Pearson, 2020.
10. Kaggle, "House Prices – Advanced Regression Techniques Dataset." Available: <https://www.kaggle.com/>
11. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer, 2009.
12. IBM, "Data Visualization and Business Intelligence Concepts."



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details