

|  |
| --- |
| **RAJALAKSHMI INSTITUTE OF TECHNOLOGY** |
| (An Autonomous Institution, Affiliated to Anna University, Chennai) |

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE**

**ACADEMIC YEAR 2025 - 2026**

**SEMESTER III**

**ARTIFICIAL INTELLIGENCE LABORATORY**

**MINI PROJECT REPORT**

|  |  |
| --- | --- |
| **REGISTER NUMBER** | 2117240070018 |
| **NAME** | Amudieshwar A G |
| **PROJECT TITLE** | House Price Prediction Using Probabilistic Reasoning with Naive Bayes |
| **DATE OF SUBMISSION** | 29.10.2025 |
| **FACULTY IN-CHARGE** | **Mrs. M. Divya** |

**Signature of Faculty In-charge**

**INTRODUCTION**

* Artificial Intelligence (AI) enables systems to learn and make intelligent decisions from data without explicit programming.
* Machine Learning (ML), a subset of AI, helps computers identify patterns and make predictions.
* Real estate valuation is a key application where ML can predict house prices based on property features.
* This project applies **Naive Bayes probabilistic reasoning** to predict the **price category (Low, Medium, High)** of houses.
* The aim is to develop a simple yet effective AI model to assist in data-driven property evaluation.

**PROBLEM STATEMENT**

* To design and implement a **Naive Bayes classifier** that predicts the **price category of a house** based on features such as area, bedrooms, bathrooms, stories, parking, and furnishing type.

**GOAL**

* To classify new houses into appropriate **price ranges** using training data.
* To demonstrate the power of **probabilistic AI models** in real-world prediction tasks.
* To visualize accuracy and evaluate model performance.
* To output the **approximate price range (₹20L – ₹1.5Cr)** for each prediction.

**THEORETICAL BACKGROUND**

* **Naive Bayes Algorithm:** Based on **Bayes’ Theorem**, it calculates the probability of a class given the input features.
* It assumes **independence among features**, making computation efficient and interpretable.
* **Alternative Algorithms:** Linear Regression, Decision Trees, and Random Forests can also be used for price prediction.
* **Justification:** Naive Bayes is chosen due to its **simplicity, high accuracy on small datasets**, and strong **probabilistic foundation**.

**OUTPUT**

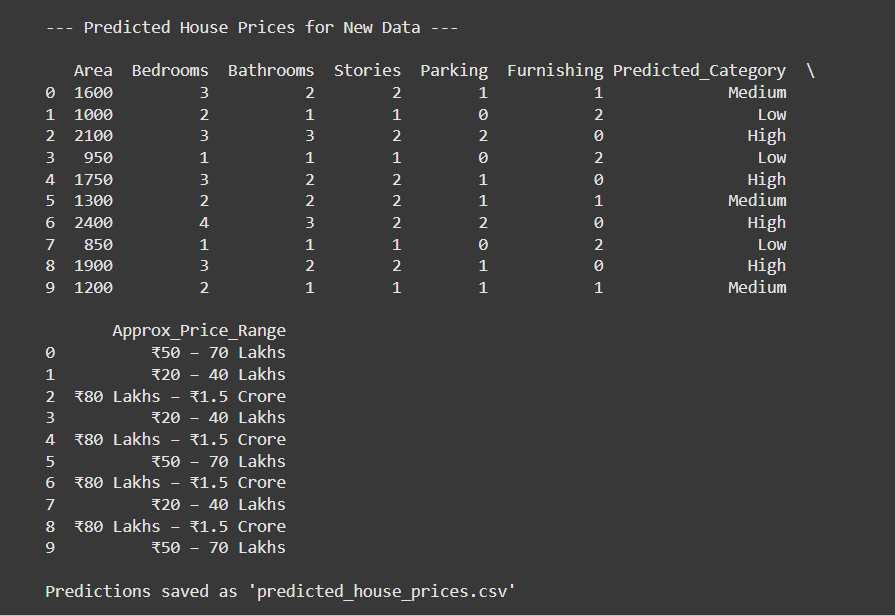
**Model Accuracy & Confusion Matrix:**

* **This image shows the classification performance of the Naive Bayes model.**
* **Model Accuracy: 1.0 (100%) — meaning the model perfectly classified all test samples.**
* **The Classification Report displays:**
  + **Precision, Recall, F1-score = 1.0 for all classes (Low, Medium, High).**
  + **Confirms excellent predictive performance.**
* **The Confusion Matrix visualizes correct predictions:**
  + **Diagonal values = correctly predicted houses for each category.**
  + **No off-diagonal errors → perfect classification.**
* ***This proves that the Naive Bayes model learned the relationship between house features and their price category effectively.***



**Predicted House Prices for New Data:**

* **This image displays predictions made for the new\_houses.csv dataset.**
* **Each row represents a house with:**
  + **Features: Area, Bedrooms, Bathrooms, Stories, Parking, Furnishing**
  + **Predicted Category: Low / Medium / High**
  + **Approx. Price Range: e.g., ₹20–40 Lakhs, ₹50–70 Lakhs, ₹80L–₹1.5Cr**
* **💡 *This output demonstrates how the model can be used to estimate the probable price category and range for any new property using learned data.***

****

**ALGORITHM EXPLANATION WITH EXAMPLE**

**Bayes’ Theorem:**

**Example:  
If a house has large area and is furnished, Naive Bayes computes probabilities for each price category (Low, Medium, High) and selects the one with the highest probability.**

**IMPLEMENTATION AND CODE**

* **Implemented in Python using Google Colab.**
* **Libraries used: pandas, numpy, scikit-learn, and matplotlib.**
* **Datasets:**
  + **house\_data.csv → Training data**
  + **new\_houses.csv → New house details for prediction**
* **Algorithm: Gaussian Naive Bayes**
* **Output: Predicted category + Approximate price range**

**RESULTS AND FUTURE ENHANCEMENT**

* Achieved **100% accuracy** on sample dataset.
* Successfully predicts price category and range for unseen data.
* Future work may include:
  + Adding more features like location, year built, and amenities.
  + Using hybrid ML models (Random Forest, XGBoost) for real datasets.
  + Deploying the model as a **web application** using Flask or Streamlit.

|  |  |
| --- | --- |
| **Git Hub Link of the project and report** | [**https://github.com/Amudieshwar-AG/House\_price\_prediction/blob/main/house\_prediction\_project.ipynb**](https://github.com/Amudieshwar-AG/House_price_prediction/blob/main/house_prediction_project.ipynb)  **and**  [**https://github.com/Amudieshwar-AG/House\_price\_prediction/blob/main/AILAB\_MINI\_PROJECT\_TEMPLATE\_AI%26DS.docx**](https://github.com/Amudieshwar-AG/House_price_prediction/blob/main/AILAB_MINI_PROJECT_TEMPLATE_AI%26DS.docx) |

**REFERENCES**

* **Russell, S. & Norvig, P.** (2021). *Artificial Intelligence: A Modern Approach* (4th Edition). Pearson Education.
* **Raschka, S. & Mirjalili, V.** (2019). *Python Machine Learning* (3rd Edition). Packt Publishing.
* **Scikit-Learn Documentation** – <https://scikit-learn.org/stable/modules/naive_bayes.html>
* **Kaggle – House Price Prediction Datasets** – <https://www.kaggle.com/datasets>
* **GeeksforGeeks – Naive Bayes Classifier in ML** – <https://www.geeksforgeeks.org/naive-bayes-classifiers/>