CAR PRICE PREDICTION USING MACHINE LEARNING ALGORITHM

Machine Learning Project Report

Submitted to the faculty of Engineering of
JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY KAKINADA, KAKINADA

In partial fulfillment of requirements for the award of Degree of

BACHELOR OF TECHNOLOGY
IN
ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

By

DASU HANU VENKATA NAGA SAI	**(22481A5415)**
GANJI MATHEWS HENRY	**(22481A5421)**
GUDURI MOHANA SAI AVINASH	**(22481A5429)**
KANTHETE NAGA DURGA YASWANTH	**(22481A5438)**
KRISHNASAI MALLADI	**(22481A5456)**

Under the guidance of

Dr.K.AshokReddy M Tech,Ph.D
Assistant Professor,Department of AI&DS

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE
SESHADRI RAO GUDLAVALLERU ENGINEERING COLLEGE
(An Autonomous institute with permanent Affiliation to JNTUK, Kakinada)
SESHADRI RAO KNOWLEDGE VILLAGE
GUDLAVALLERU-521356

2022-2026

SESHADRI RAO GUDLAVALLERU ENGINEERING COLLEGE

(An Autonomous institute with permanent Affiliation to JNTUK, Kakinada)

SESHADRI RAO KNOWLEDGE VILLAGE,

GUDLAVALLERU-521356

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

CERTIFICATE

This is to certify that the project report entitled **“Car price prediction Using Machine Learning Algorithm**” is a bonafide record of work carried out by **Dasu Hanu Venkata Naga Sai (22481A5415), Ganji Mathews Henry (22481A5421), Guduri Mohana Sai Avinash (22481A5429), Kanthete Naga Durga Yaswanth(22481A5438) and Krishnasai Malladi(22481A5456**) Under the guidance and supervision of **Dr. K. Ashok Reddy** in partial fulfillment of the requirements for the award of degree of Bachelor of Technology in Artificial Intelligence And Data Science of Jawaharlal Nehru Technological University Kakinada, Kakinada during the academic year 2024-2025.

Project Guide	**Head of the Department**
Dr. K. ASHOK REDDY	**Dr. S. NARAYANA**

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of any task would be incomplete without the mention of people who made it possible and whose constant guidance and encouragement crown all the efforts with success.

We would like to express our deep sense of gratitude and sincere thanks to
Dr. K. Ashok Reddy Assistant professor, Department of Artificial Intelligence and Data Science for his constant guidance, supervision and motivation in completing the project work.

We feel elated to express our floral gratitude and sincere thanks to **Dr. S. Narayana**, Head of the Department, Artificial Intelligence and Data Science for his encouragements all the way during analysis of the project. His annotations, insinuations and criticisms are the keys behind the successful completion of the project work.

We would like to take this opportunity to thank our beloved principal **Dr. B. Karuna Kumar** for providing a great support for us in completing our project and giving us the opportunity for doing project.

Our Special thanks to the faculty of our department and programmers of our computer lab. Finally, we thank our family members, non-teaching staff, attendants and our friends, who had directly or indirectly helped and supported us in completing our project in time.

By

DASU HANU VENKATA NAGA SAI (22481A5415)
GANJI MATHEWS HENRY (22481A5421)
GUDURI MOHANA SAI AVINASH (22481A5429)
KANTHETE NAGA DURGA YASWANTH (22481A5438)
KRISHNASAI MALLADI (22481A5456)

ABSTRACT

Cars are the most common means of transport used by most people often to travel on a regular basis for personal, professional and other leisure purposes too. Prediction of car prices is very crucial for individuals who are interested in cars such as Buyers, Sellers and Manufacturers of cars. This project presents a machine learning based approach for predicting the prices of the cars based on the features of the cars.

By analyzing features such as manufacturer or brand, year of production, interior finish, fuel type, engine volume, gear box type and airbags, a predictive model has been developed using the K Nearest Neighbors Algorithm. The model was rigorously evaluated using standard metrics like Adjusted R-squared value and mean-squared error.

The achieved performance shows the accuracy in predictions. The proposed system once trained with real-time data of car prices and features can efficiently predict prices of the cars. The K Nearest Neighbors algorithm was chosen for obtaining higher accurate results. After training and testing, the model displayed strong predictive power, demonstrating its suitability for integration into car markets.

Such a system can help predicting the price of the cars in an effective way without any human interference. The project includes comprehensive stages of data preprocessing, feature selection, model training, and evaluation. The final model demonstrated high accuracy and reliability in predicting the prices of the cars. Through the data preprocessing, feature engineering, and model evaluation, a robust predictive system was developed.

The model was assessed using various performance metrics, indicating high reliability and generalizability. This work contributes to ongoing research in car markets, transportation researches and finance researches and showcases the potential of machine learning in building an effective car price prediction system that can support car value analysis.

This system can be integrated with real-time cars buy and sale i.e., retail platforms helping the users in predicting the prices of the cars by themselves without any human interference. The results demonstrate the potential of machine learning in diagnosing the car asset value by predicting its price.

INDEX

CHAPTER 1
INTRODUCTION

1.1 INTRODUCTION

Cars are most essential parts of modern transportation and have become a symbol of convenience, status, and personal freedom. With a wide variety of makes, models, features, and performance levels, car prices can vary significantly.

Factors such as brand reputation, engine specifications, mileage, fuel type, transmission, and market trends all contribute to the overall pricing of a vehicle. For buyers and sellers alike, accurately estimating the price of a car is crucial whether for making informed purchasing decisions or setting competitive selling prices.

In recent years, the automotive industry has increasingly adopted data-driven approaches to tackle complex problems, and machine learning has emerged as a powerful tool in this context. Car price prediction is one such application where machine learning algorithms can analyze historical data and learn patterns to predict the price of a car based on various features. These models can significantly enhance the accuracy and efficiency of price estimation compared to traditional manual methods or rule-based systems.

This project focuses on building a machine learning model capable of predicting car prices based on key attributes. By training the model on a dataset of used cars, it learns to recognize patterns and relationships between the input features and the car's price. The goal is to develop a robust and reliable system that can assist both consumers and businesses in evaluating the fair market value of a vehicle.

1.2 PROBLEM STATEMENT

Predicting the price of a car is crucial for buyers, sellers, and manufacturers to make informed decisions. Car prices are influenced by various features like brand or the manufacturer, Year of Manufacture, fuel type, engine capacity or volume,etc. The goal is to build a machine learning model that accurately estimates car prices based on such features. This helps in automating price evaluation, reducing human error, and enhancing transparency in the automobile market.

1.3 EXISTING SYSTEM

1.Check value of your car for FREE – CarWale

- Platform for buying and selling cars.
- Uses Manufacturing Year, City, Car Type, owners value and kilometers driven
- Does not include other features like fuel type, engine capacity, airbags for safety standards,etc. Also shows right price not available for listed cars too.

2. Car-price-prediction-streamlit

- An app helping in predicting selling price of cars
- Uses ex-showroom price, driven distance, fuel type, seller type, transmission, owners and car purchased year.
- Does not focus on brand market value and price depreciation of brands, also the manufacture year of car is not considered.

3.Used Car Valuation – Calculate Value of Car for FREE at CarDekho

- Platform to calculate value of car and to sell car.
- Uses brand, Registration year, variant, fuel type, ownership, distance driven, location.
- Does not provide a direct interface for calculating price. It is shown only when registration number or exact car details are present i.e., helpful only for sellers but not buyers.

1.4 PROPOSED SYSTEM

The proposed system aims to build a predictive model for predicting car prices using better accurate machine learning algorithms on the available data. It addresses the limitations of the traditional and existing systems by including all the insights required of the data, providing an UI to calculate price by any interested individual say buyers and sellers too and focusing on the depreciation of the car price based on brand type present in historic data of dataset.

The proposed system uses car-price-prediction dataset consisting of information of several cars such as Price of the car based on different features of the cars.The objective is to calculate the price of the cars by understanding the underlying patterns of the cars.

Key Components of the Proposed System:

1. Data Collection:
    - Car Price and Feature related information from the dataset

2. Data Preparation or Preprocessing:
    - Handling missing values, Duplicate values, null values and outliers.
    - Label encoding for categorical variables.
    - Feature scaling using Standard Scaler for consistent input ranges.
    - Feature Selection for identifying significant factors using correlation and other graphical mechanisms.
3. Model Development:
    - Split the data into testing and training data for calculating accuracy of models.
    - Developed different Machine Learning models such as Linear Regression, Lasso Regression, K Nearest Neighbors Regression and Support Vector Machine.
4. Performance Evaluation:
    - Calculated Adjusted R-Squared Factor value and Mean Squared Error of each model.
    - Model with highest R-Squared Adjustment value and least Mean Squared Error is considered and choosen as the best model or fit for the data.(Here KNN).
5. Model Deployment:
    - Developed a Flask API for real-time prediction.
    - Deployed the K Nearest Neighbors Regressor model in local machine for checking the working and performance of system.

1.5 ADVANTAGES

1.Higher Accuracy:

Using K Nearest Neighbors Regressor produces the highest Adjusted R-Squared value means that it best suits or fits the data from the dataset.

2.Reduces Overfitting:

By using the larger and varying k values in K Nearest Neighbors Regressor, the model reduces overfitting on the data.

3.Robust to Noisy and Missing Data:

The algorithm is robust to noise and can handle missing values, making it suitable for real-world market analysis of car price data that may be incomplete or inconsistent.

4. Feature Selection:

Features are appropriately selected using correlation approach with the price value of the cars.

5.Scalability:

Model performs well on large datasets with many features too.

6.Disciplined Procedure:

System was developed by following all the steps in a sequential way as :

Data Handling → Feature Selection→ Model Development → Evaluation → Deployment.

1.6 DISADVANTAGES

There are some limitations to the developed system as follows:

1.No Spatial Feature Integration:

The dataset lacks geographical features like city name or location, which sometimes play key role in deciding the prices of the cars.

2. Low Scalability:

Although a Flask API and local machine deployment were made, the system hasn't been well designed for large-scale deployment or cloud-based processing, which may be necessary for operational use.

3.Computational Complexity:

As dataset size or features increases, Choosing KNN as model for prediction results in more computational complexity.

4.Choosing the Optimal Value of K:

Selecting the optimal value of k is always challenging as it depends on the data and influences the model's performance too.

CHAPTER 2
REQUIREMENT ANALYSIS

2.1 FUNCTIONAL REQUIREMENTS

The functional requirements define the core functionalities that the system must perform to achieve the goal of predicting prices of the cars. These requirements ensure that the system operates efficiently and accurately using machine learning techniques, particularly the
K Nearest Neighbor Regressor.

1. **Data Requirements:**

 Historical data of cars is required i.e., we need to import cars data in structured formats such as csv, xlsx, etc for model development.

2. **Model Requirements:**

 Data Preprocessing modules are needed to handle missing values, null values, outliers and categorical values(such as label encoders).

 Feature Engineering and Transformation modules are needed.

 Machine Learning Algorithms Module such as KNN,SVM, Linear Regression,etc needed.

Model Evaluation metrics such as Mean Squared Error and R-Squared Adjustment Values are needed.

3. **Deployment Requirements:**

A web Application or an API is needed to deploy the developed system.

A user friendly interface needed to be created to input car details and predict car prices.

Regular update to the developed model is needed for maintaining accuracy.

2.2 NON – FUNCTIONAL REQUIREMENTS

Non-functional requirements define the overall qualities and constraints of the system that influence user experience, performance, and maintainability. These requirements ensure that the system operates reliably and efficiently, even under various conditions.

1. **Performance Requirements:**

Ensure that the model provides better accurate predictions.

Model must be capable of responding within short time such as less than or equal to 2 seconds.

Maintaining minimum scalability of the API or model is required.

System must be able to maintain a good throughput i.e., good count of requests/minute.

2. **Security Requirements:**

Any important information related to model or data must be kept secure so that the model is not collapsed by any evil parties resulting wrong outputs.

Access to definite parties must be accessed based on previliges.

3. **Reliability Requirements:**

The system should perform consistently and produce accurate predictions when provided with valid input data. It should handle unexpected input gracefully and avoid crashing.

4. **Usability Requirements:**

UI must be simple, clear, easy and intuitive for users with different levels of expertise.

5. **Maintainability Requirements:**

The system should be modular, allowing for easy updates, such as re-training the model with new data or replacing the algorithm. Code should be well-documented for future enhancements.

2.3 SOFTWARE REQUIREMENTS

The following software components are necessary to develop, run, and evaluate the machine learning model for predicting car prices:

1. **Programming Language & Environment**

Language: Python 3.7 or above.

IDE: Jupyter Notebook for development and testing, Alternatively: Visual Studio Code or PyCharm.

2. **Libraries and Frameworks**

Data Analysis: pandas, numpy – Data Manipulation.

Visualization: matplotlib, seaborn – Plots and Graphs.

Machine Learning: scikit-learn – Model Building (KNN, Linear Regression, Lasso Regression, SVM, Label Encoder, Standard Scaler, etc).

Web Framework: Flask – For Creating Restful APIs.

pickle – A module to save the models i.e., storing sequence formats or patterns.

3. **Platforms and Systems**

Operating System: Cross-platform (Windows, Linux, macOS).

Browser: For accessing local Flask server endpoints.

CHAPTER 3

DESIGN

3.1 SYSTEM ARCHITECTURE

The system architecture for the car price prediction using Machine Learning Algorithms project consists of multiple interconnected components that handle data processing, model training, evaluation, and prediction. Each layer plays a crucial role in identifying underlying patterns associated with the variable price of the car.

1. **Data Source Layer-**

Inputs: Historical cars data (CSV, Excel, database, or API)

Details: This includes essential fields such as car price, manufacturer, production year, Interior Finish (Leather Interior), FuelType, Engine Volume, Gear box Type and Airbags

Purpose: Serves as the foundation for analysis and model training.

2. **Data Preprocessing Module-**

Tasks:

- Handles missing values, null values, duplicates in the data and outliers.
- Converts the large data into smaller ranges by scaling the data.

Purpose: Ensures the data is in a usable format for machine learning.

3. **Feature Engineering Module-**

Tasks:

- Selects most significant features based on correlation or importance

Purpose: Enhances model accuracy and performance.

4. **Data Splitting Module-**

Tasks:

- Split the entire data into training data and testing data in the ratio 7:3.

Purpose: To keep some values unseen by the model that could be used to assess the accuracy of the model.

5. **Machine Learning Model Module-**

Algorithms Used: Linear Regression, Lasso Regression, K Nearest Neighbors, Support Vector Machine

Tasks:

- Trains the model using processed and engineered data.

Purpose: Learns from historic training data to predict future prices.

6. **Model Evaluation Module-**

Tasks:

- Tests model performance using metrics like R^2 and MSE.

Purpose: Validates the model's accuracy before deployment.

7. **Prediction Module-**

Input: Test Data

Output: Predicted price of the car.

Task: Applies the trained model to unseen input features.

8. Output/User Interface-

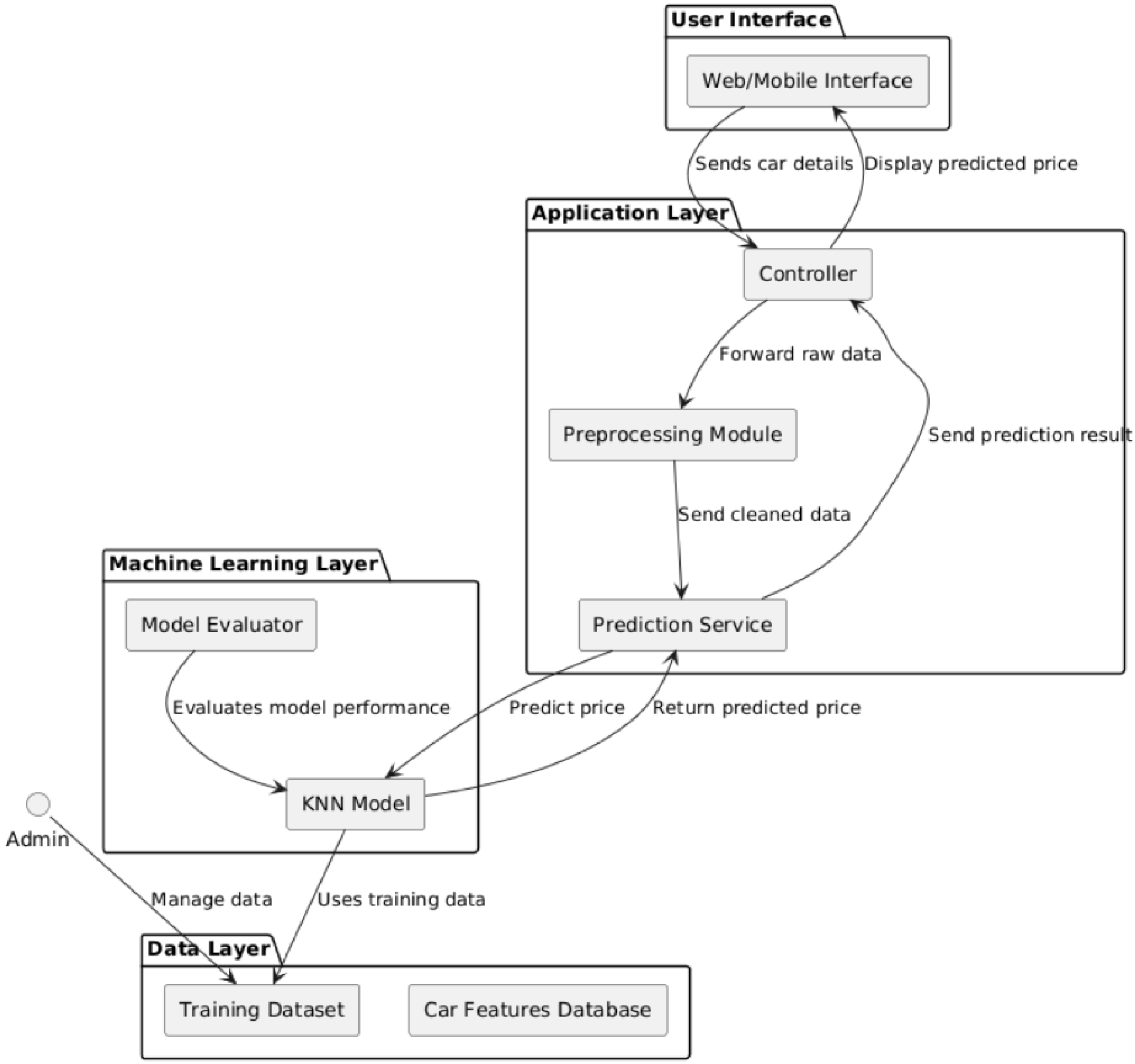
Tasks:

Allows users to input features of the car.

Displays predicted price of the car using the best fitted model.

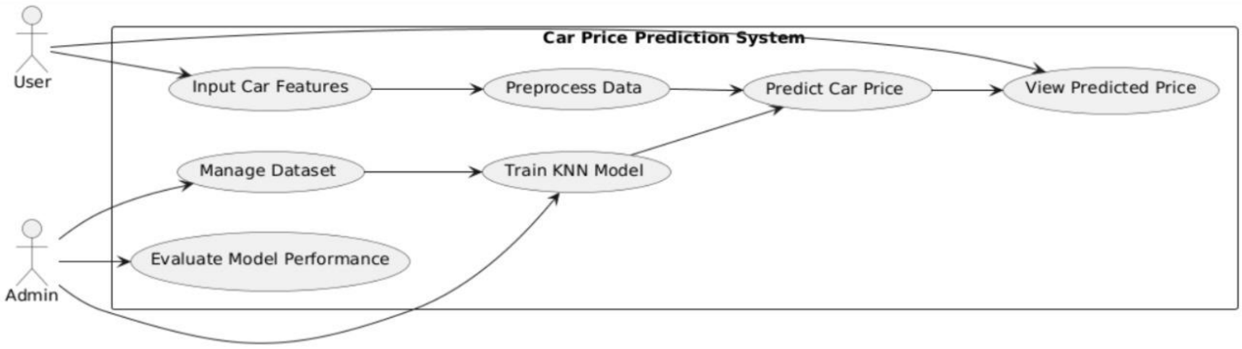
Tools: Flask(model connectivity) and HTML(Front-end UI Page)

Purpose: Enhances usability and user interaction with the system.



3.2 UML DIAGRAMS

3.2.1 USECASE DIAGRAM



The use case diagram represents how users and administrators interact with the Car Price Prediction System. It identifies the major functionalities (use cases) and the actors involved in each action, helping to understand the system's behaviour from an external perspective.

Actors:

1. User
   - A buyer, seller, or any individual who wants to predict the price of a car.
   - Interacts with the system by entering car details and viewing predicted prices.
2. Admin
   - A system administrator who manages the system operations.
   - Responsible for maintaining the dataset, training the KNN model, and evaluating model performance.

Use Cases:

- Input Car Features
  The user enters car information such as brand, year, engine volume, fuel type, gearbox, airbags, etc.
- Preprocess Data
  The system processes and cleans the input data to make it suitable for prediction.
- Train KNN Model
  The admin trains the KNN model using the latest dataset to improve prediction accuracy.
- Predict Car Price
  Based on user inputs and the trained model, the system predicts the car's price.
- View Predicted Price
  The predicted price is displayed to the user.
- Manage Dataset
  The admin can add, update, or delete records in the dataset used for training.
- Evaluate Model Performance
  The admin evaluates how well the KNN model performs using metrics such as Mean Squared Error and R^2.

3.2.2 CLASS DIAGRAM

Car Price Prediction System using KNN

This project implements a machine learning-based system to predict car prices using the K-Nearest Neighbors (KNN) algorithm.

The system analyzes car features such as brand, year, engine volume, fuel type, gearbox type, and airbags to estimate the price.

It targets buyers, sellers, and manufacturers seeking quick and accurate car valuations.

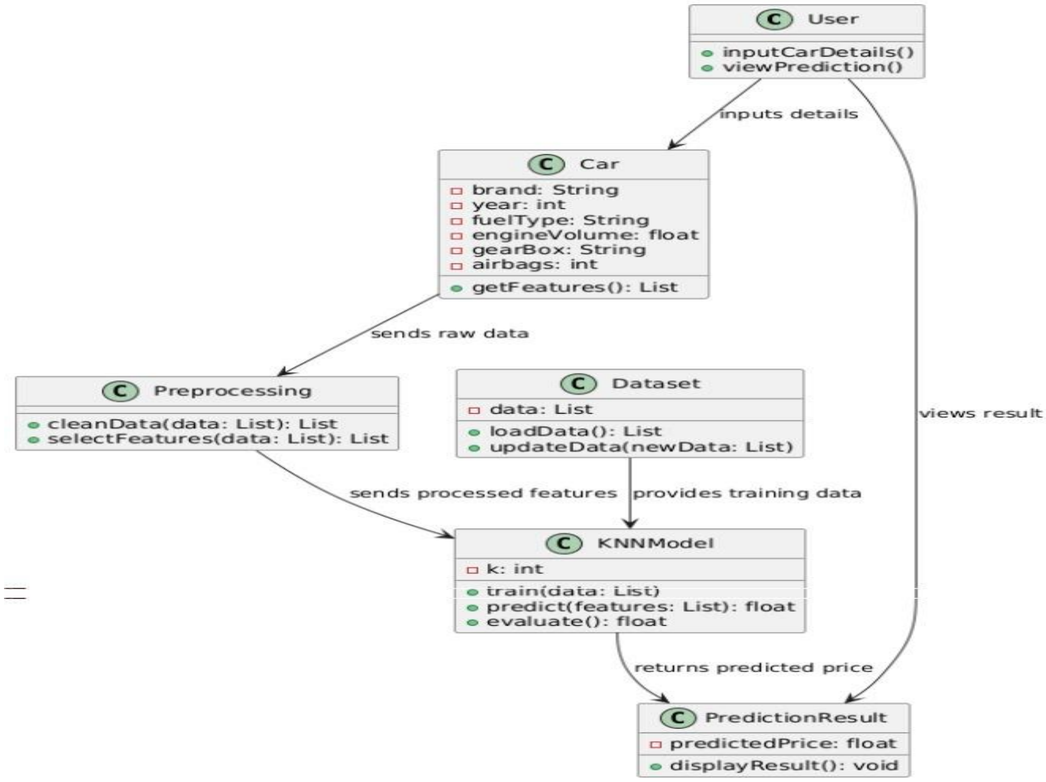
The system is composed of four main layers:

1. User Interface: Allows users to input car details and view results.
2. Application Layer: Handles preprocessing, prediction requests, and system control.
3. Machine Learning Layer: Contains the trained KNN model and evaluation module.
4. Data Layer: Stores the training dataset and car features.

Users can predict prices by entering car details, while admins manage the dataset, train the model, and evaluate performance using metrics like Mean Squared Error and Adjusted R^2.

The system is accurate, scalable, and suitable for integration with car retail platforms.

This work demonstrates the potential of machine learning in enhancing decision-making and price transparency in the automotive sector.

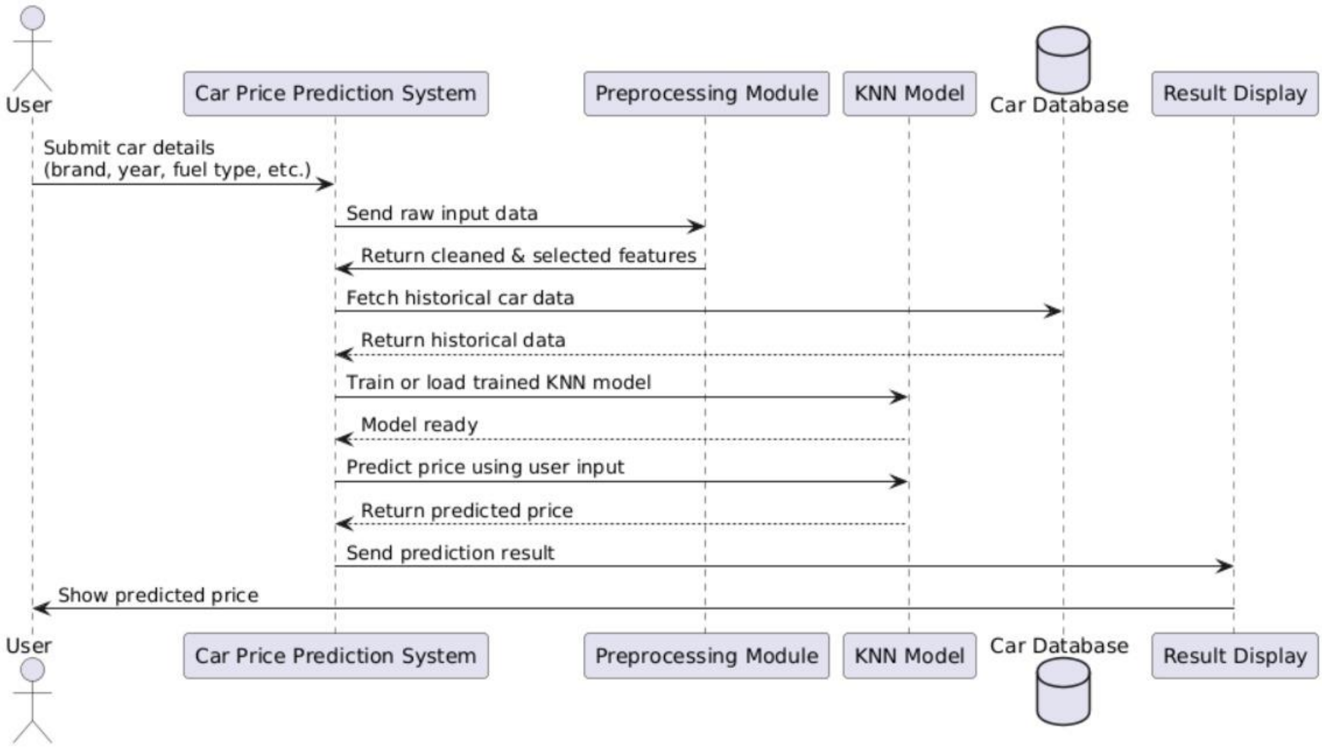


3.2.3 SEQUENCE DIAGRAM

The sequence diagram illustrates the interaction between the user and system components during car price prediction. The User inputs car details via the UI, which forwards them to the Prediction Service. The service calls the Data Preprocessor to clean and transform the data, then sends the processed data to the KNN Model for prediction. The predicted price is returned to the User through the UI.

In the backend, the Admin can trigger model training and evaluation using updated datasets. This ensures accurate predictions and continuous system improvement.

The diagram highlights the sequential flow of data and operations, supporting real-time, automated price estimation.



CHAPTER 4
IMPLEMENTATION

4.1 TECHNOLOGY DESCRIPTION AND TRAINING PROCESS

The Car Price Prediction System is a machine learning-based application that estimates the price of a car based on its features such as brand, year of manufacture, fuel type, engine capacity, gearbox type, and number of airbags. This project aims to provide an intelligent solution for predicting the market value of a car without manual intervention, helping buyers, sellers, and manufacturers make informed decisions.
The K-Nearest Neighbors (KNN) algorithm is used for building the predictive model due to its simplicity and effectiveness in regression tasks. The system is capable of learning from historical car data and producing accurate price predictions for new car entries.

Technology Description

The system has been developed using the following technologies:

- Programming Language: Python
- Libraries and Frameworks:
    - pandas: Data manipulation and analysis
    - numpy: Numerical computations
    - scikit-learn: Machine learning model building and evaluation
    - matplotlib / seaborn: Data visualization
- Machine Learning Algorithm: K-Nearest Neighbors (KNN)
- Model Evaluation Metrics:
    - Adjusted R^2 Score
    - Mean Squared Error (MSE)
- Development Environment: Jupyter Notebook
- Data Storage: CSV file
- Optional Tools: Flask for deployment

Dataset Collection

The dataset consists of historical car listings including attributes such as:

- Manufacturer/Brand
- Year of production
- Fuel type
- Engine volume
- Gearbox type

- Number of airbags
- Interior finish
- Car price (target variable)

Data Preprocessing

Before training the model, the dataset undergoes preprocessing:

- Handling of missing values using imputation methods (e.g., forward fill)
- Encoding categorical variables using techniques like Label Encoding .
- Normalization or standardization of numerical attributes for uniform scaling using Standard Scaler.

Data Splitting

The data is split into training and testing sets in an 70:30 ratio to ensure the model is evaluated on unseen data.

Model Training

The KNN regression model is trained using the training dataset. The value of k (number of neighbors) is selected based on accuracy during validation.

Evaluation

The model is evaluated using metrics such as:

- Adjusted R^2 Score: Measures how well the model explains the variability of the data.
- Mean Squared Error (MSE): Measures the average squared difference between actual and predicted values.

PSEUDO CODE:

```
Start
1. Import Libraries
   - Import pandas as pd
   - Import numpy as np
   - Import required modules from sklearn:
      - LabelEncoder
      - train_test_split
      - KNeighborsRegressor
      - Evaluation metrics: r2_score, mean_squared_error

2. Load and Inspect Dataset
   - Read CSV file: 'car_price_prediction.csv' into DataFrame (df)
```

- Display dataset information using df.info()
- Print number of unique values for selected features to understand data diversity

3. Data Cleaning and Preprocessing
   - Remove ' km' from 'Mileage' column and convert it to integer
   - Identify all columns with 'object' (categorical) data types
   - Apply LabelEncoder to convert categorical columns to numeric format

4. Feature Selection and Data Splitting
   - Define input features (X) and target variable (y)
   - Split the data into training and test sets using train_test_split

5. Model Training
   - Work with different models such as Linear Regression, Lasso, SVM, KNN.
   - Initialize KNeighborsRegressor model with desired parameters (e.g., n_neighbors=5)
   - Repeat the working for all the selected models.
   - Train the model using training data (X_train, y_train)

6. Model Prediction and Evaluation
   - Predict car prices using the trained model (y_pred = model.predict(X_test))
   - Calculate evaluation metrics:
     - R^2 Score
     - Mean Squared Error (MSE)
   - Print evaluation results

7. Conclusion
   - Summarize model performance and findings
   - Suggest possible improvements such as trying other algorithms or tuning hyper parameters

End

CHAPTER 5

RESULTS

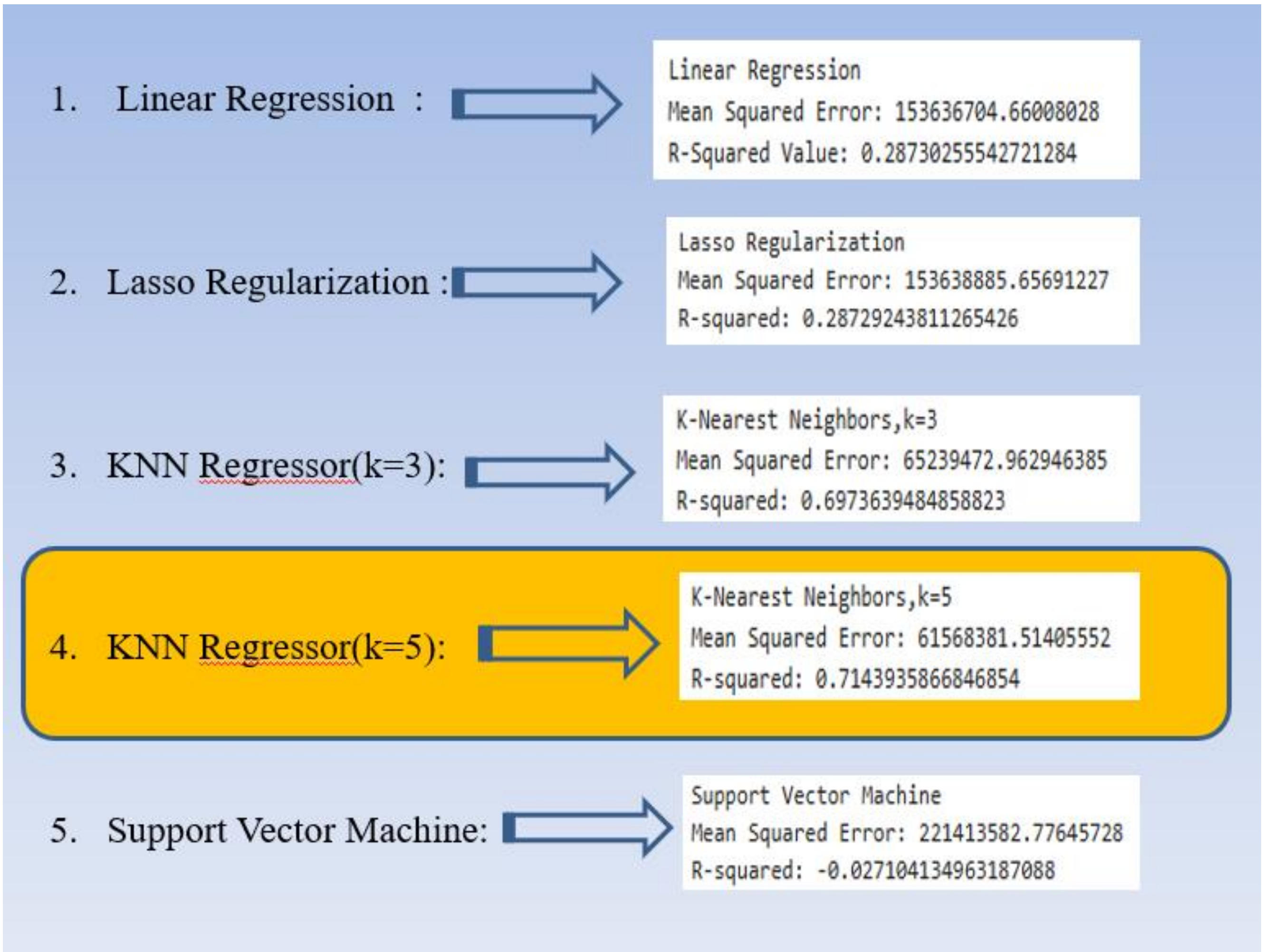


Fig: Performance Evaluation Metrics

By observing the performance evaluation metrics, we can clearly evaluate the working of different models on the data in the given dataset.

***** Evaluating using the Mean-Squared Error (Least MSE value implies a best fit model):

Sorted Values: **KNN (k=5) < KNN (k=3) < Linear < Lasso < SVM**

Order of Best Fit by MSE: **KNN (k=5) > KNN (k=3) > Linear > Lasso > SVM**

*****Evaluating using the R-Squared Adjustment Factor (Higher R-squared value implies a best fit model):

Sorted Values: **KNN (k=5) > KNN (k=3) > Linear > Lasso > SVM**

Order of Best Fit by R-squared Factor: **KNN (k=5) > KNN (k=3) > Linear > Lasso > SVM**

BEST FIT MODEL CHOSEN BY METRICS: **KNN (k=5)**

→Thus we developed the User interface for interaction with the user using the best fit model being integrated as the resource for prediction of the prices of the car. The User Interface looks as:

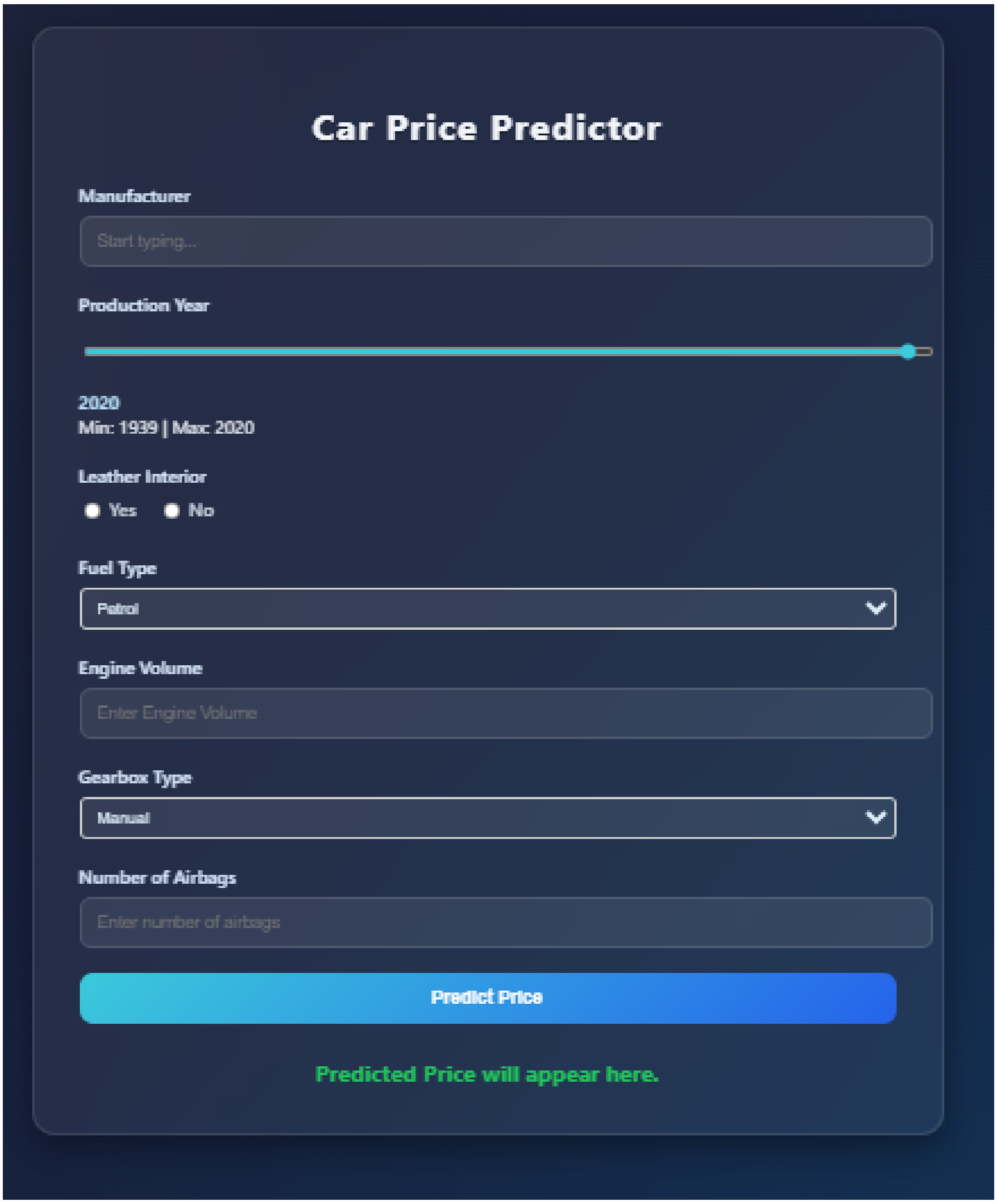


Fig: User – Interface for Car Price Prediction

When the user inputs the details of the manufacturer brand, Production Year, Leather Interior (Yes or No), fuel Type, Engine Volume, Gearbox Type, Number of Airbags and hits the predict button available, the page processes the input data and uses the trained model information to predict the price of the car and gives the value to the user.

CHAPTER 6
CONCLUSION

6.1 CONCLUSION

The Car Price Prediction System has significant potential for future improvements and could play an even more pivotal role in the automotive industry. One of the key developments would be its integration with online car sales platforms, such as car dealership websites, online marketplaces, and e-commerce platforms. By enabling real-time price estimates, users can instantly receive accurate car valuations when listing their cars for sale or while browsing potential vehicles for purchase. This would streamline the process for both buyers and sellers, making it more efficient and transparent. In addition to real-time integration, incorporating real-time data from APIs or external databases will ensure that the system stays continuously updated with the latest market trends. For example, incorporating live car sales data from platforms like Autotrader or Kelley Blue Book would help adjust the predictions based on fluctuations in car prices, regional demand, and economic factors. This would further increase the relevance of the predictions, making them more aligned with the current market scenario. To enhance the system's prediction accuracy, exploring advanced machine learning algorithms such as Random Forest, Gradient Boosting Machines (GBM), or Deep Neural Networks would be beneficial. These models could capture more complex relationships between the car's features and its market value, improving accuracy beyond the capabilities of KNN. Additionally, implementing ensemble methods that combine multiple algorithms could lead to even more robust results by reducing overfitting and increasing model generalization.

Another exciting direction for development is the inclusion of image-based predictions. By leveraging deep learning models such as Convolutional Neural Networks (CNNs), the system could analyze car images to assess its condition, detect damages, or even evaluate the external aesthetics, such as paint quality or tire wear. This would provide a more comprehensive valuation by considering both the car's physical condition and its technical specifications, leading to more accurate price estimations.

Furthermore, the development of a mobile application would make the system accessible on-the-go. Users could snap pictures of cars and input details through their smartphones, receiving instant price predictions. This would also be a valuable tool for car dealerships and buyers at auctions or in person, enabling them to make informed decisions in real-time.The system could also be improved by incorporating additional features such as mileage, service history, and accident reports into the model. By including these elements, the model would be able to give more accurate predictions based on factors that significantly impact a car's price. Similarly, the system could be adapted to account for regional pricing trends by integrating location-based data, enabling predictions tailored to specific geographical markets, and addressing variations in pricing due to factors like local demand and supply.

6.2 FUTURE SCOPE

The Car Price Prediction System has a very great significant potential for future enhancements and broader applications. One of the primary areas for improvement is the integration with online platforms, such as car resale websites and dealership portals, where users can obtain real-time price estimates while listing or browsing vehicles. By incorporating real-time data sources, such as APIs from automotive marketplaces and car listing platforms, the system can remain continuously updated, ensuring that predictions reflect the latest market trends and fluctuations. Additionally, the use of advanced machine learning algorithms such as Random Forest, Gradient Boosting, or deep learning techniques (like neural networks) can be explored to further improve accuracy, especially when dealing with complex, non-linear relationships between car features and pricing.

Another promising extension is the inclusion of image-based predictions using computer vision techniques, where the system could analyze car images to assess external condition and factor it into the price estimation, providing a more comprehensive evaluation of the vehicle's value. Furthermore, a mobile application version of the system could make it more accessible to users, allowing them to check car prices on-the-go, whether they're at a dealership or inspecting a car in person.

To further enhance the system's performance and reliability, additional features such as mileage, accident history, service records, and regional pricing trends could be incorporated into the model, creating a richer dataset for more accurate predictions. Moreover, the implementation of automated retraining will allow the system to adapt to changes in the market by continuously learning from new data, ensuring that the predictions remain relevant and up-to-date over time. The system can also be expanded for multilingual and regional adaptability, providing localized pricing predictions tailored to different countries, languages, and regional market conditions, thus broadening its user base and applicability worldwide.

In the long term, this system could evolve into a fully autonomous tool that aids buyers, sellers, and manufacturers in making data-driven decisions, thus transforming how car pricing is determined and assessed in the automotive industry

6.3 REFERENCES

[1] Pudaruth, S. (2014). Predicting the Price of Second-hand Cars using Artificial Neural Networks. *Proceedings of the Second International Conference on Data Mining, Internet Computing, and Big Data*, SDIWC.

[2] Selvaraj, V. (2022). Predicting Resale Car Prices Using Machine Learning Regression Models with Ensemble Techniques. *International Journal of Computer Science and Mobile Computing*, 11(2), 45-53.

[3] Longani, C., Rajesh, M., & Pudaruth, S. (2021). Price Prediction for Pre-Owned Cars Using Ensemble Machine Learning Techniques. *Journal of Intelligent & Fuzzy Systems*, 41(1), 2305–2312.

[4] Gegic, E., & Pudaruth, S. (2015). Car Price Prediction Using Machine Learning Techniques. *International Journal of Data Mining & Knowledge Management Process*, 5(1), 1–10.

[5] Jain, D., & Kataria, A. (2020). Machine Learning Approaches to Predict Car Prices. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 9(3), 289–292.

[6] Jayanthi, A., & Rajalakshmi, P. (2023). Analysis and Prediction of Car Prices Using Supervised Learning Algorithms. *Journal of Computer Applications*, 46(1), 33–39.

[7] Gaur, S., & Gupta, R. (2021). Second-hand Car Price Prediction Using Data Analytics. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(2), 231–237.

[8] Khan, M. S., & Singh, R. (2022). Machine Learning Based Used Car Price Prediction Using Linear and Non-Linear Models. *Journal of Emerging Technologies and Innovative Research*, 9(4), 874–880.

[9] Sharma, A., & Verma, N. (2020). Used Car Price Prediction Using K-Nearest Neighbors. *International Journal of Advanced Research in Computer Science*, 11(6), 110–115.

[10] Thakur, A., & Bansal, R. (2023). Predictive Modeling for Car Pricing Using Data Science Techniques. *International Journal of Artificial Intelligence and Applications*, 14(1), 19–26.

Program Outcomes (POs):

Engineering Graduates will be able to:

1. **Engineering knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals and an engineering specialization to the solution of complex engineering problems

2. **Problem analysis:** identify, formulate, review research literature, and analyse complex engineering problem reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

3. **Design/development of solutions**: Design solutions for complex engineering problems and designsystemcomponentsorprocessesthatmeetthespecifiedneedswithappropriateconsideration for the public health and safety, and the cultural, societal, and environmental considerations.

4 Conduct investigations of complex problems: Use research-hased knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

5. **Modern tool usage:** Create, select, and apply appropriate techniques, resources and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

6. **The engineer and society:** Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

7. **Environment and sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

8. **Ethics:**Applyethicalprinciplesandcommittoprofessionalethicsandresponsibilitiesandnorms of the engineering practice

9. **Individual and team work:** Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

10. **Communication:** Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

11. **Project management and finance:** Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

12. **Life-long learning:** Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change

Program Specific Outcomes (PSOs):

Engineering students will also be able to

1. Process, interpret the real-world data to formulate the model for predicting and forecasting.

2. Apply machine learning techniques to design and develop automated systems to solve real world problems.

PROJECT PROFORMA

Classification of Project	Application	Product	Research	Review

Note: Tick the appropriate box.

Project Outcomes	
Course Outcome (CO1)	Acquire technical competence in the specific domain during the training.
Course Outcome (CO2)	Identify the problem statement based on the requirements of the industry.
Course Outcome (CO3)	Adapt project management skills on par with industrial standards.
Course Outcome (CO4)	Develop a system model to obtain solution and generate a report.