A Major Project Synopsis on

**Car Price Prediction**

Submitted to Manipal University, Jaipur

Towards the partial fulfillment for the Award of the Degree of

**MASTER OF COMPUTER APPLICATIONS**

2023-2025

by

HIMANSHU TANWAR

23FS20MCA00054

A close up of a sign

AI-generated content may be incorrect.

Under the guidance of

Dr. Geeta Chhabra Gandhi, Assistant Professor

**Department of Computer Applications**

**School of AIML, IoT & IS, CCE, DS and Computer Applications**

**Faculty of Science, Technology and Architecture**

**Manipal University Jaipur**

**Jaipur, Rajasthan**

**2025**

**I. Introduction**

The used car market in India is highly dynamic, with prices fluctuating based on various factors such as make, model, mileage, condition, and market trends. Sellers often struggle to determine the right price for their vehicles, leading to inefficiencies in the market. To address this challenge, we propose developing a Machine Learning (ML) model that predicts the price of a used car based on its features. Car price prediction is crucial for manufacturers, dealers, and individual buyers or sellers. Traditional pricing methods often rely on manual comparisons or limited rule-based systems, which may not capture the full complexity of market dynamics

**II. Motivation**

Our proposed solution aims to assist both buyers and sellers by providing a systematic approach to price estimation. The lack of transparent pricing mechanisms often results in extended sales cycles and financial losses. By leveraging machine learning techniques, we intend to bring efficiency and fairness to the used car market.

**1. For Sellers:**

* Enable accurate pricing for faster sales and better returns.
* Reduce the chances of undervaluation or overpricing.

**2. For Buyers:**

* Assist in identifying fair-priced vehicles.
* Enhance the decision-making process with data-driven insights.

**3. For the Market:**

* Improve transparency and efficiency.
* Provide structured and automated pricing recommendations.

**III. Problem Statement**

* The used car market in India is a constantly changing landscape where pricing can be unpredictable due to various influencing factors.
* Sellers often face challenges in accurately pricing their cars, leading to longer sales cycles and potential financial losses.
* Buyers, on the other hand, may struggle to identify fair-priced vehicles.
* The lack of a systematic approach to price estimation results in market inefficiencies.

**IV. Approach**

* To tackle this issue, we propose to build a Machine Learning model that utilizes historical data of used car sales from the dataset.
* The model will analyze key attributes such as brand, model, year of manufacture, mileage, fuel type, and location to predict an optimal selling price. By leveraging data analytics and predictive modeling, we aim to create a system that benefits both buyers and sellers.

**IV. Methodology / Planning of Work**

We propose to build a Machine Learning model that utilizes historical data of used car sales from the dataset.

By leveraging data analytics and predictive modeling, we aim to create a system that benefits both buyers and sellers.

**1. Project Workflow:**

1. **Understanding the Problem Statement** – Defining the scope and objectives.
2. **Getting System Ready** – Setting up the required tools and libraries.
3. **Data Collection** – Gathering relevant data from reliable sources.
4. **Data Understanding** – Data description and preliminary observations.
5. **Data Cleaning & Preprocessing (Phase I)** – Handling missing values, outliers, and inconsistencies.
6. **Exploratory Data Analysis (EDA):**
   * Univariate Analysis
   * Multivariate Analysis
7. **Data Cleaning & Preprocessing (Phase II)** – Feature engineering and selection.
8. **Insights from Data Visualization** – Understanding key patterns and relationships.
9. **Feature Engineering** – Creating and transforming features for better model performance.
10. **Model Building & Evaluation** – Training various ML models and assessing their performance.
11. **Selection of Best Model & Hyperparameter Tuning** – Optimizing the best-performing model.
12. **Generating Pickle File** – Saving the trained model for future predictions.

**2. Working with Backend:**

* **Programming Language:** Python
* **Machine Learning Libraries:** Scikit-Learn, TensorFlow, Pandas, NumPy
* **Database:** MySQL (for storing training data)
* **Frontend:** ReactJS (for building an interactive interface for users)

**V. Requirements for Proposed Work**

**1. Software Requirement:**

* **Operating System:** Windows, Linux
* **User Interface:** ReactJS
* **Database:** MySQL
* **Backend:** Python with Flask or Django

**2. Hardware Requirement:**

* **Processor:** Pentium-based systems with a minimum of P4
* **RAM:** 4GB (minimum recommended)
* **Hard Disk:** 20 GB Hard Disk Space

**VI. Benefits of This Model**

1. **Accurate Price Estimation**
   * Predicts fair and data-driven car prices, reducing chances of overpricing or under-pricing.
2. **Time-Saving**
   * Automates the valuation process, eliminating the need for manual market comparisons.
3. **Decision Support**
   * Assists both buyers and sellers in making informed and confident decisions.
4. **Scalability**
   * Can be easily scaled to include new data, brands, or markets for improved predictions.
5. **Customizable**
   * The model can be tailored for specific regions, brands, or market trends.
6. **Cost-Efficient**
   * Reduces dependency on paid car evaluation services or human experts.
7. **Business Insights**
   * Helps dealerships, online platforms, and insurance companies understand pricing patterns and customer behaviour.

**VII. Overall Requirements for Car Price Prediction using ML**

To build a car price prediction system using machine learning, the following components are essential:

1. **Dataset**
   * A dataset containing features such as car brand, model, year, mileage, fuel type, transmission, engine size, and selling price.
   * Example: Kaggle Car Dataset
2. **Data Preprocessing**
   * Handle missing values, encode categorical variables, normalize or scale numerical data, and remove outliers if necessary.
3. **Feature Selection/Engineering**
   * Choose or create features that most impact the price prediction, such as car age or mileage per year.
4. **Machine Learning Model**
   * Train regression models like Linear Regression, Random Forest, or XGBoost to predict car prices.
5. **Model Evaluation**
   * Use metrics like R² Score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) to evaluate model performance.
6. **Tools & Libraries**
   * Python, Pandas, NumPy, Scikit-learn, Matplotlib/Seaborn for analysis and visualization.

**VIII. Conclusion**

This project aims to provide an effective solution for used car price prediction using Machine Learning. By leveraging data-driven insights, it will benefit sellers, buyers, and the overall market, leading to fair pricing and a more efficient buying and selling process. The project has the potential to revolutionize the used car industry in India by improving transparency and decision-making.

**XI. References**

1.  **Patel, D. R., & Desai, P. S.** (2020). *Used car price prediction using Machine Learning techniques*. International Journal of Engineering Research & Technology (IJERT), 9(06).  
   <https://www.ijert.org/used-car-price-prediction-using-machine-learning-techniques>
2.  **Aggarwal, A., & Soni, R.** (2021). *Car price prediction using machine learning algorithms*. International Journal of Engineering Research & Technology (IJERT), 10(03).  
   <https://www.ijert.org/car-price-prediction-using-machine-learning-algorithms>
3.  Hossain, M. A., Ahmed, M., & Kabir, M. (2019). Prediction of used car price using machine learning techniques. In 2019 International Conference on Bangla Speech and Language Processing (ICBSLP).IEEE.  
   <https://ieeexplore.ieee.org/document/9034970>
4.  UCI **Machine Learning Repository.** (n.d.). Automobile Data Set.  
   <https://archive.ics.uci.edu/ml/datasets/Automobile>
5.  Reese**, A.** (n.d.). Craigslist Cars and Trucks Data [Dataset]. Kaggle.  
   <https://www.kaggle.com/datasets/austinreese/craigslist-carstrucks-data>
6.  Birla**, N.** (n.d.). Vehicle dataset from Car Dekho [Dataset]. Kaggle.  
   <https://www.kaggle.com/datasets/nehalbirla/vehicle-dataset-from-cardekho>
7.  Balcıoğlu**, Y. S., & Sezen, B.** (2024). Car Price Prediction Using Machine Learning Techniques. 6th International Artemis Congress on Health and Sport Sciences.  
   [https://www.researchgate.net/publication/379443793\_](https://www.researchgate.net/publication/379443793_CAR_PRICE_PREDICTION_USING_MACHINE_LEARNING_TECHNIQUES)
8.  Amshi**, A. T.** (2023). Vehicle Price Prediction By Aggregating Decision Tree Model With Boosting Model. arXiv preprint arXiv:2307.15982.  
   <https://arxiv.org/abs/2307.15982>
9.  Gegic**, E., Kadic, N., & Music, J.** (2020). Car Price Prediction using Machine Learning Techniques. 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO).  
   <https://ieeexplore.ieee.org/document/9245195>
10.  Pudaruth**, S.** (2014). Predicting the Price of Used Cars using Machine Learning Techniques. International Journal of Information & Computation Technology, 4(7), 753-764.  
    https://www.ripublication.com/irph/ijict\_spl/ijictv4n7spl\_14.pdf
11.  **Monburinon, N., & Chaiyaratana, N.** (2017). Prediction of Prices for Used Car by Using Regression Models. 2017 9th International Conference on Knowledge and Smart Technology (KST).  
    <https://ieeexplore.ieee.og/document/7886102>
12.  Noor**, K., & Jan, S.** (2019). Vehicle Price Prediction System using Machine Learning Techniques. 2019 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET).  
    <https://ieeexplore.ieee.org/document/8673454>
13.  Gajera**, H.** (2022). Car Price Prediction Using Machine Learning. International Journal for Research in Applied Science & Engineering Technology (IJRASET), 10(5).  
    <https://www.ijraset.com/research-paper/car-price-prediction-using-machine-learning>