# Machine Learning Based Cars Price Prediction & Recommendation Project

**\* EL KAHLAOUI Youssef**

**\*\* KHAMJANE Aziz**

\*\*\*Digital Transformation and Artificial Intelligence

National School of Applied Sciences of Al Hoceima

23/12/2024

***Abstract*-** The Car Price Prediction Project aims to develop a machine learning application that accurately predicts the prices of used cars based on various features such as manufacturer, model name, engine type, transmission, mileage, price, and age.

This project involves data scraping from the AA Cars website [1], comprehensive data cleaning and preprocessing, exploratory data analysis, and the implementation of multiple machine learning models. The final model, trained on a combined dataset, demonstrates a slight performance improvement over individual datasets. The project also includes an interactive web interface built with Flask, allowing users to input car details and receive price predictions and recommendations for similar cars. This paper presents the methodology, model evaluation, and results, highlighting the effectiveness of the chosen approach in predicting car prices and providing recommendations.

***Index Terms***- Car Price Prediction, Machine Learning, Recommendation System, Web Application, Data Scraping

T

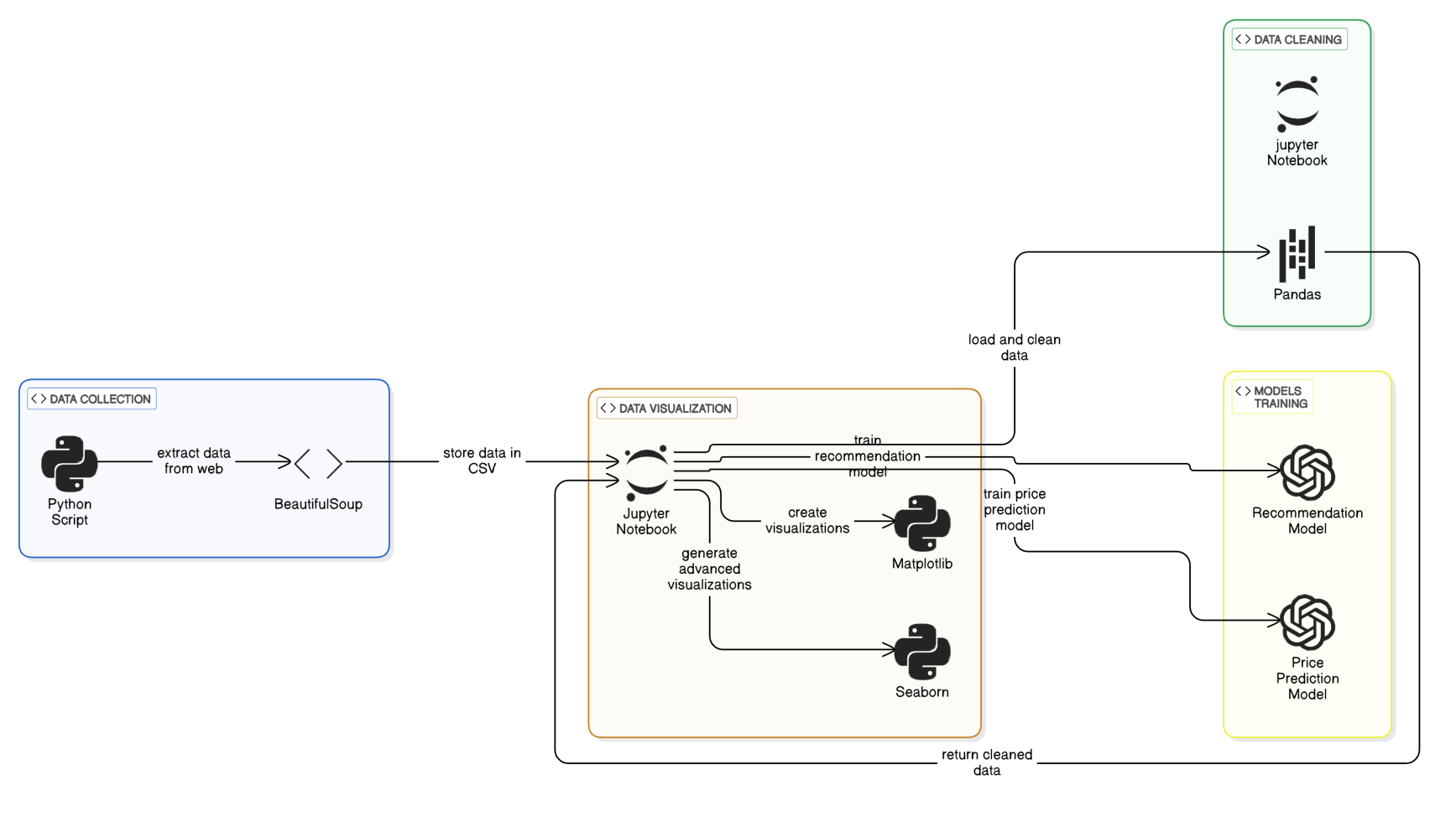
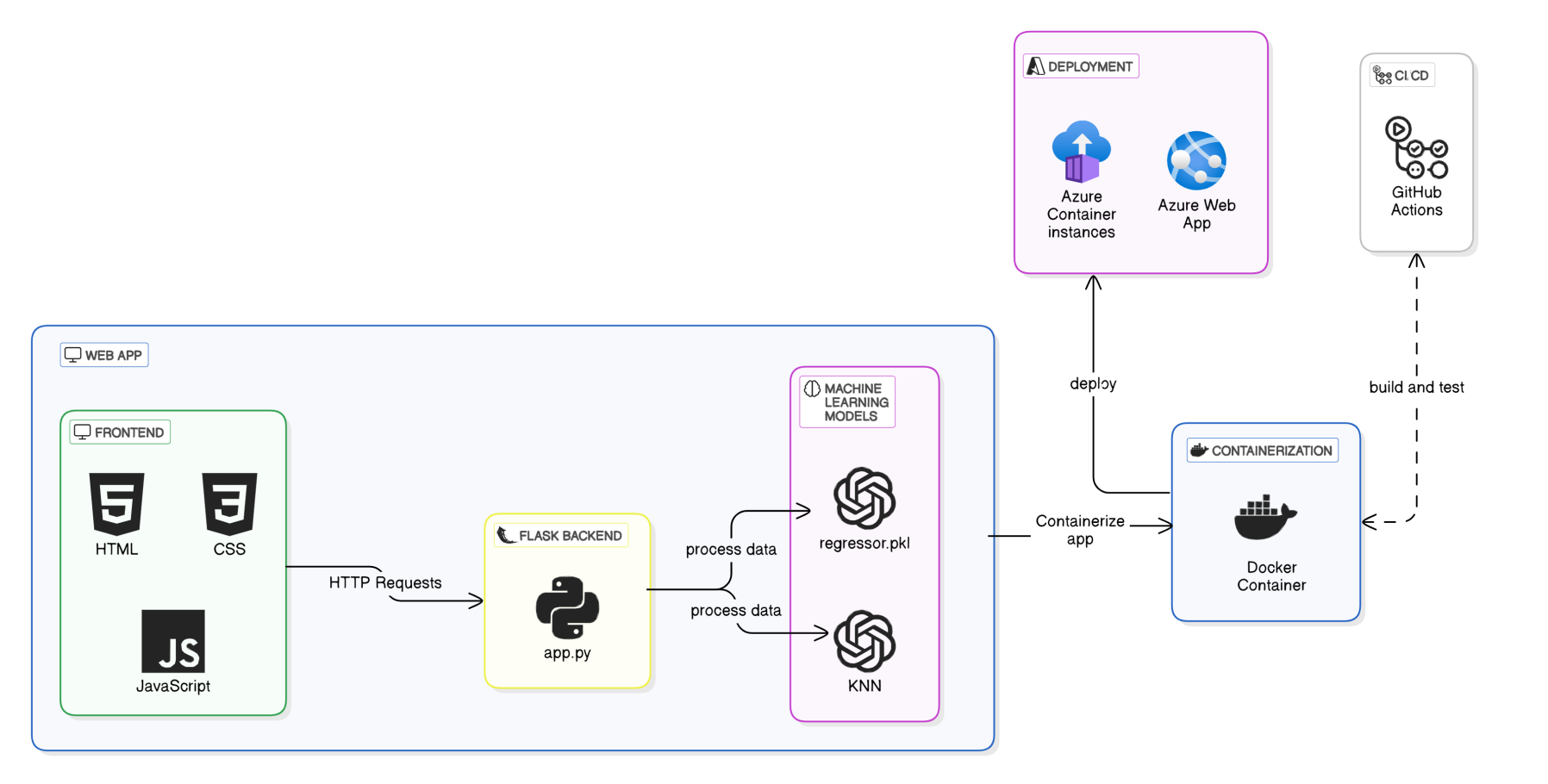
1. **Introduction**

he used car market is a dynamic and complex industry where accurate pricing is crucial for both buyers and sellers. Predicting the price of a used car involves considering various factors such as the car's manufacturer, engine type, transmission, mileage, price, and age. Traditional methods of car price estimation often rely on expert knowledge and manual assessments, which can be subjective and inconsistent. With the advent of machine learning, it is possible to develop models that can predict car prices more accurately and consistently.

This research paper presents the Car Price Prediction Project, which aims to leverage machine learning techniques to predict the prices of used cars and recommend similar cars based on input features. The project involves several key steps: data scraping from the AA Cars website, data cleaning and preprocessing, exploratory data analysis, and the implementation and evaluation of multiple machine learning models. The final model, trained on a combined dataset, shows a slight performance improvement, indicating the effectiveness of the chosen approach.

The paper is structured as follows: Section 2 describes the architecture of both model and the web application .Section 3 reviews related work in the field of car price prediction and machine learning models. Section 4 describes the methodology, including data collection, preprocessing, feature engineering, model training, and the recommendation system. Section 5 presents the discussion, including model evaluation and comparison as part A. Section 5 part B discusses the limitations of the approach and potential future work. Finally, Section 7 concludes the paper by summarizing the key findings and contributions.

By developing an accurate and reliable car price prediction model and a recommendation system, this project aims to provide valuable insights and tools for both buyers and sellers in the used car market, ultimately contributing to more informed decision-making and fairer pricing.

1. **Architecture**
   1. **Models Buildeing :**
   2. **Web Application :**
2. **Related Works**

R

esearch on estimating the price of used cars is relatively recent and not extensively covered. In her MSc thesis, Listiani [2] demonstrated that a regression model using support vector machines (SVM) can predict the residual price of leased cars more accurately than simple multiple regression or multivariate regression. SVMs are particularly effective in handling high-dimensional data (numerous features used to predict the price) and can avoid both overfitting and underfitting. She employed a genetic algorithm to optimize the SVM parameters efficiently. However, the study did not express the improvement of SVM regression over simple regression in straightforward measures like mean deviation or variance.

In another university thesis, Richardson [4] explored the hypothesis that car manufacturers aim to produce vehicles that do not depreciate quickly. Using multiple regression analysis, he found that hybrid cars (vehicles with both an internal combustion engine and an electric motor) retain their value better than traditional vehicles. This is likely due to increased environmental concerns and higher fuel efficiency. The study also considered other factors such as age, mileage, make, and MPG (miles per gallon). Data for this study was collected from various websites.

Wu et al. [5] utilized a neuro-fuzzy knowledge-based system to predict the price of used cars, considering only three factors: the make of the car, the year of manufacture, and the engine style. The proposed system produced results comparable to simple regression methods. In the USA, car dealers sell hundreds of thousands of cars annually through leasing. Most of these cars are returned at the end of the leasing period and must be resold. Accurately pricing these cars is crucial for economic success. To address this, Du et al. [6] developed the ODAV (Optimal Distribution of Auction Vehicles) system, which not only estimates the best resale price but also advises on the optimal location to sell the car. Given the vast size of the United States, the selling location significantly impacts the price of used cars. A k-nearest neighbor regression model was used for price forecasting. Since its inception in 2003, the system has distributed over two million vehicles.

Gonggi [7] proposed a model based on artificial neural networks to forecast the residual value of private used cars. The study focused on features such as mileage, manufacturer, and estimated useful life. The model was optimized to handle nonlinear relationships, which simple linear regression methods cannot manage. The model proved to be reasonably accurate in predicting the residual value of used cars.

In a study by Sameerchand Pudaruth [8], supervised machine learning techniques were applied to predict the price of used cars in Mauritius. The predictions were based on historical data collected from daily newspapers. Various techniques, including multiple linear regression analysis, k-nearest neighbors, naïve Bayes, and decision trees, were used to make the predictions. The predictions were evaluated and compared to identify the best-performing methods. The study concluded that predicting the price of used cars is a challenging problem that requires sophisticated algorithms for high accuracy. All four methods provided comparable performance.

Inspired by the study conducted by Sameerchand Pudaruth [8], this research explores the application of machine learning techniques to predict the price of used cars. These studies illustrate the diverse approaches and techniques used in predicting the price of used cars, such as support vector machines, multiple regression analysis, neuro-fuzzy systems, k-nearest neighbor regression, and artificial neural networks. Each method has its advantages and limitations, and the choice of method depends on the specific requirements and constraints of the problem.

1. **Methodology**

M

guides a stepwise walkthrough by Experts for writing a successful journal or a research paper starting from inception of ideas till their publications. Research papers are highly recognized in scholar fraternity and form a core part of PhD curriculum. Research scholars publish their research work in leading journals to complete their grades. In addition, the published research work also provides a big weight-age to get admissions in reputed varsity. Now, here we enlist the proven steps to publish the research paper in a journal.

* 1. **Data Collection and Training**

M

guides a stepwise walkthrough by Experts for writing a successful journal or a research paper starting from inception of ideas till their publications. Research papers are highly recognized in scholar fraternity and form a core part of PhD curriculum. Research scholars publish their research work in leading journals to complete their grades. In addition, the published research work also provides a big weight-age to get admissions in reputed varsity. Now, here we enlist the proven steps to publish the research paper in a journal.

* 1. **Model Description**

M

guides a stepwise walkthrough by Experts for writing a successful journal or a research paper starting from inception of ideas till their publications. Research papers are highly recognized in scholar fraternity and form a core part of PhD curriculum. Research scholars publish their research work in leading journals to complete their grades. In addition, the published research work also provides a big weight-age to get admissions in reputed varsity. Now, here we enlist the proven steps to publish the research paper in a journal.

* 1. **Feature Engeneering**

M

guides a stepwise walkthrough by Experts for writing a successful journal or a research paper starting from inception of ideas till their publications. Research papers are highly recognized in scholar fraternity and form a core part of PhD curriculum. Research scholars publish their research work in leading journals to complete their grades. In addition, the published research work also provides a big weight-age to get admissions in reputed varsity. Now, here we enlist the proven steps to publish the research paper in a journal.

* 1. **Model Training and Evaluation**

M

guides a stepwise walkthrough by Experts for writing a successful journal or a research paper starting from inception of ideas till their publications. Research papers are highly recognized in scholar fraternity and form a core part of PhD curriculum. Research scholars publish their research work in leading journals to complete their grades. In addition, the published research work also provides a big weight-age to get admissions in reputed varsity. Now, here we enlist the proven steps to publish the research paper in a journal.

1. **Results**

M

guides a stepwise walkthrough by Experts for writing a successful journal or a research paper starting from inception of ideas till their publications. Research papers are highly recognized in scholar fraternity and form a core part of PhD curriculum. Research scholars publish their research work in leading journals to complete their grades. In addition, the published research work also provides a big weight-age to get admissions in reputed varsity. Now, here we enlist the proven steps to publish the research paper in a journal.

1. **Discussion**
   1. **Analysis of Results**

guides a stepwise walkthrough by Experts for writing a successful journal or a research paper starting from inception of ideas till their publications. Research papers are highly recognized in scholar fraternity and form a core part of PhD curriculum. Research scholars publish their research work in leading journals to complete their grades. In addition, the published research work also provides a big weight-age to get admissions in reputed varsity. Now, here we enlist the proven steps to publish the research paper in a journal.

* 1. **Limitations**

guides a stepwise walkthrough by Experts for writing a successful journal or a research paper starting from inception of ideas till their publications. Research papers are highly recognized in scholar fraternity and form a core part of PhD curriculum. Research scholars publish their research work in leading journals to complete their grades. In addition, the published research work also provides a big weight-age to get admissions in reputed varsity. Now, here we enlist the proven steps to publish the research paper in a journal.

1. **Conclusion**

**References**

1. G. O. Young, “Synthetic structure of industrial plastics (Book style with paper title and editor),” in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.
2. W.-K. Chen, *Linear Networks and Systems* (Book style)*.* Belmont, CA: Wadsworth, 1993, pp. 123–135.
3. H. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1985, ch. 4.
4. B. Smith, “An approach to graphs of linear forms (Unpublished work style),” unpublished.
5. E. H. Miller, “A note on reflector arrays (Periodical style—Accepted for publication),” *IEEE Trans. Antennas Propagat.*, to be published.
6. J. Wang, “Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication),” *IEEE J. Quantum Electron.*, submitted for publication.

**Elements**

**\*Youssef El kahlaoui** – Author name, qualifications, associated institute (if any) and email address.

**\*\*Second Author** – Author name, qualifications, associated institute (if any) and email address.

**\*\*\*Third Author** – Author name, qualifications, associated institute (if any) and email address.

**Correspondence Author** – Author name, email address, alternate email address (if any), contact number.

**Appendices**