A

PROJECT REPORT

ON

# Used Car Price Predictor

Submitted in partial fulfilment of the requirements of the degree

## Bachelor of Engineering In Information Technology

By

Aditya Nar 34

Dattatray Narhe 35

Harsh Gawas 38

Harsh Pandey 10

Supervisor: Prof. Amarja Adgaonkar



Department of Information Technology

K.C. College of Engineering and Management Studies And Research,

Thane (E)

University of Mumbai

2023-24

## CERTIFICATE

This is to certify that the project entitled “Car Price Predictor” is a bonafide work of Aditya Nar (34 ) , Dattatray Narhe (35) , Harsh Pandey (38) , Harsh Gawas (10) submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of “Bachelor of Engineering” in “Information Technology”.

Name and sign

Co Supervisor/Guide Supervisor/Guide



Head of Department Principal

Prof.Amarja Adgaonkar Dr.Vilas Nitnaware

### Project Report Approval for T.E

This project report entitled Car Price Predictor by Aditya Nar, Dattatray

Narhe, Harsh Pandey, Harsh Gawas is approved for the degree of Bachelor of Engineering in Information Technology

Examiners

1.---------------------------------------------

2.---------------------------------------------

Date:

Place:

## DECLARATION

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Signature)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Aditya Nar - 34)

(Dattatray Narhe-35)

(Harsh Pandey – 38)

(Harsh Gawas - 10)

Date:

## DECLARATION

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Signature)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Aditya Nar - 34)

(Dattatray Narhe-35)

(Harsh Pandey – 38) (Harsh Gawas - 10)

Date:

## ACKNOWLEDGEMENT

We would like to express special thanks of gratitude to our guide Mrs.Punam Bagul as well as our Project Coordinator who gave us the golden opportunity to do this wonderful project on the topic of Car Price Predictor which also helped us in doing a lot of research and we came to know about so many new things. We are very grateful to our Head of the Department Mrs.Amarja Adgaonkar for extending her help directly and indirectly through various channels in our project work. We would also like to thank Principal Dr. Vilas Nitnaware for providing us the opportunity to implement our project. We are really thankful to them. Finally we would also like to thank our parents and friends who helped us a lot in finalizing this project within the limited time frame.

Thanking You.

### TABLE OF CONTENT

Certificate ..............................................................................................i

Approval Sheet .....................................................................................ii

Declaration …….......................……………………………………….iii

Acknowledgement ……………………………………..…..................iv

List of Figures ...………………………………………………………vi

List of Table ...………………………………………………………...vii

Abstract ....…………………………………………………………….viii

1. Introduction ………………………………………………………...9
2. Literature Survey…………………………………………………..10
3. Proposed Work
   1. Requirement Analysis…………………………………………
      1. Scope ……………………………………………………..11
      2. Feasibility Study ………………………………………….12
      3. Hardware & Software Requirement ………………………13
   2. Problem Statement ………………………………………….14
   3. Project Design……………………………………………….15
   4. Methodology…………………………………………………31

1. Conclusion and Future Scope………………………………………..32

1. References ………………………………………………………….. 33

### LIST OF FIGURES

Sr. No Topic Page No

1. Flow Chart 15

1. Block Diagram 15
2. Use Case Diagram 16
3. DFD Level 0 16

1. DFD Level 1 17

1. Output User Interface 17

## ABSTRACT

Abstract: The movie recommendation system using machine learning (ML) is a pivotal application in the realm of personalized content delivery. This paper introduces an innovative approach to enhance user experience by leveraging ML algorithms to offer tailored movie suggestions. The system employs collaborative filtering and content-based filtering techniques to analyze user preferences and movie attributes. Collaborative filtering identifies similarities between users and recommends movies based on their past interactions, while content-based filtering suggests movies with similar characteristics to those previously enjoyed by the user. The methodology incorporates matrix factorization to extract latent features from the user-item interaction matrix, facilitating accurate predictions of user preferences. Additionally, similarity-based methods enable the system to recommend movies based on their genre, cast, directors, and plot summaries. The effectiveness of the recommendation system is evaluated through metrics such as accuracy, coverage, and user satisfaction. Results demonstrate the system's ability to provide personalized recommendations, enhancing the movie-watching experience for users across various platforms. The proposed system not only addresses the challenge of information overload but also promotes content discovery and engagement. Future enhancements may include integrating additional features such as user demographics and contextual information to further refine recommendations. Overall, the movie recommendation system using ML represents a promising approach to deliver tailored movie suggestions, fostering user satisfaction and retention in the ever-expanding digital entertainment landscape.

# 1. Introduction

In today's era of digital streaming platforms, the vast array of movie choices can often overwhelm users, leaving them struggling to find films that match their interests. To address this challenge, we propose the development of a small-scale movie recommendation system powered by machine learning (ML). By harnessing ML algorithms, this system aims to provide personalized movie suggestions tailored to individual user preferences.

Traditional methods of browsing through extensive movie catalogs can be time-consuming and inefficient. Users may find themselves sifting through numerous titles without finding anything that truly captures their interest. The proposed recommendation system seeks to alleviate this frustration by analyzing user behavior and movie attributes to offer tailored recommendations.

The core concept of the system revolves around collaborative filtering and content-based filtering techniques. Collaborative filtering analyzes patterns in user interactions with movies, such as ratings and viewing history, to identify similarities between users and recommend titles based on the preferences of similar users. Content-based filtering, on the other hand, focuses on the intrinsic characteristics of movies, such as genre, cast, and plot, to recommend films that share similarities with those previously enjoyed by the user.

By leveraging these techniques, the recommendation system can provide users with personalized movie suggestions that align with their tastes and preferences. Through the integration of ML models, such as matrix factorization and similarity-based approaches, the system aims to improve the accuracy and relevance of its recommendations.

In summary, the proposed small-scale movie recommendation system offers a streamlined solution to the challenge of finding relevant movie content in a sea of options. By harnessing the power of ML algorithms, this system endeavors to enhance the movie-watching experience for users by delivering tailored recommendations that cater to their individual tastes and preferences.

# 2. Literature Survey

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | Title | Methodology | Advantages | Limitations |
| 2021 | "Predicting Used  Car Prices Using  Machine  Learning" | - Data collection from online listings - Preprocessing  (imputation, encoding) -  Regression models | - Accurate price predictions based on historical data - Scalable for large datasets - Simple model interpretation | - Limited to regression models, may not capture complex  patterns - Relies heavily on data quality and availability |
| 2020 | "Ensemble  Learning for Used  Car Price  Prediction" | - Ensemble methods  (random forests, gradient boosting) - Feature engineering | - Improved predictive accuracy through ensemble techniques - Handles  nonlinear relationships in  data | - Increased computational complexity compared to  individual models - Requires careful tuning of  hyperparameters for optimal  performance |
| 2019 | "Deep Learning  Approaches for  Predicting Used Car Prices" | - Deep learning models  (neural networks,  LSTM) - Feature scaling and normalization | - Captures complex patterns and nonlinear relationships in data - Potential for  automatic feature extraction | - Requires large amounts of data for training deep learning models effectively - Prone to overfitting  if not properly regularized and  validated |
| 2018 | "Hybrid  Approach for  Used Car Price Prediction" | - Hybrid of regression and clustering  techniques - Feature selection and extraction | - Integrates multiple modeling techniques for robust predictions -  Addresses heterogeneity in data | - Complexity in combining different models and  methodologies - Interpretability may be compromised with hybrid approaches |
| 2017 | "Sentiment  Analysis in Used  Car Price  Prediction" | - Sentiment analysis of customer reviews -  Feature engineering for sentiment features | - Incorporates qualitative data (customer sentiments) into pricing predictions -  Provides additional context for pricing decisions | - Relies on accurate sentiment analysis tools and datasets -  Limited to availability of customer reviews and sentiment  data |

## 3.1 Proposed Work

### 3.1 Requirement Analysis

#### 3.1.1 Scope

Scope:

The scope of the proposed small-scale movie recommendation system encompasses several key aspects, including:

1. Data Collection and Preprocessing: Gathering relevant data sources such as user ratings, movie metadata, and viewing history from available APIs or databases. Preprocessing this data to ensure consistency, accuracy, and compatibility with the recommendation algorithms.

2. Algorithm Selection and Implementation: Choosing appropriate ML algorithms, such as collaborative filtering, content-based filtering, or hybrid approaches, based on the available data and system requirements. Implementing these algorithms to generate personalized movie recommendations for users.

3. User Interface Design: Designing an intuitive and user-friendly interface through which users can interact with the recommendation system. This may include features such as search functionality, personalized recommendation lists, and user feedback mechanisms.

4. Evaluation and Testing: Assessing the performance of the recommendation system through metrics such as accuracy, coverage, and user satisfaction. Conducting thorough testing to identify and address any issues or limitations in the system's functionality.

5. Scalability and Performance Optimization: Considering the scalability of the system to handle increasing volumes of users and movie data. Implementing optimizations to enhance the system's performance, such as parallel processing, caching, or distributed computing.

6. Integration with Existing Platforms: Exploring opportunities to integrate the recommendation system with existing movie streaming platforms or applications. This integration may involve API integrations, data synchronization, and compatibility testing.

7. Deployment and Maintenance: Deploying the recommendation system on a suitable hosting platform or infrastructure. Implementing monitoring and maintenance procedures to ensure the system remains operational, secure, and up-to-date with evolving user preferences and movie trends.

8. User Feedback and Iterative Improvement: Establishing mechanisms for collecting user feedback and incorporating it into the recommendation algorithms. Iteratively improving the system based on user input and ongoing data analysis.

Overall, the scope of the small-scale movie recommendation system encompasses the entire lifecycle of development, from data collection and algorithm implementation to user interface design, evaluation, deployment, and ongoing maintenance. By focusing on these key aspects, the system aims to provide users with personalized movie recommendations that enhance their viewing experience and satisfaction.

#### 3.1.2 Feasibility Study

Before proceeding with the development of the small-scale movie recommendation system, it is essential to conduct a feasibility study to assess its viability and determine the potential challenges and opportunities associated with its implementation. The feasibility study encompasses several key areas:

1. Technical Feasibility:

- Availability of Data: Assess the availability and accessibility of relevant movie data sources, including user ratings, movie metadata, and viewing history. Determine if adequate data can be obtained to train and test the recommendation algorithms.

2. Financial Feasibility:

- Cost Analysis: Estimate the costs associated with developing, deploying, and maintaining the recommendation system. Consider expenses related to data acquisition, infrastructure, software development, testing, deployment, and ongoing maintenance.

3. Market Feasibility:

- Market Research: Conduct market research to understand the demand for personalized movie recommendation services. Identify target demographics, user preferences, and competitors in the movie streaming and recommendation space.

4. Legal and Regulatory Feasibility:

- Data Privacy and Compliance: Ensure compliance with data privacy regulations, such as GDPR or CCPA, when collecting and processing user data. Assess legal implications related to data storage, security, and user consent.

5. Operational Feasibility:

- Resource Availability: Assess the availability of human resources, expertise, and skill sets required for developing, deploying, and maintaining the recommendation system. Determine if additional training or hiring is necessary to fill any skill gaps.

Based on the findings of the feasibility study, stakeholders can make informed decisions regarding the viability and feasibility of the small-scale movie recommendation system. If the study indicates favorable conditions and opportunities outweigh potential challenges, stakeholders can proceed with the development and implementation of the system, taking into account the identified considerations and recommendations.

#### 3.1.3 Hardware and Software Requirement

Hardware Requirements:

1. Server Infrastructure:

* High-performance server or cloud-based computing resources for model training and inference.
* Adequate storage capacity to handle large datasets and model artifacts.
* Recommended specifications: multicore CPU (e.g., Intel Core i7 or AMD Ryzen 7), sufficient RAM (at least 16GB), and SSD storage for faster data access.

2. Networking Equipment:

* Reliable internet connectivity with sufficient bandwidth to support concurrent user requests.
* Networking hardware such as routers, switches, and firewalls for secure data transmission.

3. End-user Devices:

* Users can access the system via desktop computers, laptops, tablets, or smartphones.
* Ensure compatibility with various operating systems (Windows, macOS, Linux) and web browsers (Chrome, Firefox, Safari, Edge).

Software Requirements:

1. Operating System:

* Server-side: Linux distributions (e.g., Ubuntu Server, CentOS) or Windows Server for hosting the application and ML models.
* Development environments: Windows, macOS, or Linux for software development and testing.

2. Programming Languages and Libraries:

- Machine learning frameworks such as TensorFlow, PyTorch, or scikit-learn for implementing recommendation algorithms and training machine learning models.

- Web frameworks such as Flask or Django for building the backend server and API endpoints.

- Database management systems such as PostgreSQL or MySQL for storing movie data, user profiles, and system logs.

3. Version Control:

* Version control systems like Git for managing codebase changes and collaboration among developers..

4. Development and Deployment Tools:

* Integrated Development Environments (IDEs): Visual Studio Code, PyCharm, Jupyter Notebooks for coding and testing ML models.
* Version control systems (e.g., Git) for collaboration and code management.

3.2 Problem Statement

The proliferation of digital streaming platforms has led to an abundance of movie content available to consumers, resulting in a paradox of choice. With an overwhelming number of options to choose from, users often struggle to find movies that align with their individual preferences and tastes. Traditional methods of browsing through extensive catalogs of movies are time-consuming and inefficient, leading to frustration and dissatisfaction among users.

To address this challenge, this project aims to develop a small-scale movie recommendation system powered by machine learning (ML) algorithms. The primary objective is to provide users with personalized movie recommendations tailored to their unique preferences, thereby enhancing their movie-watching experience and increasing user satisfaction.

The problem statement can be summarized as follows:

1. Information Overload: Users are inundated with a vast array of movie choices across multiple streaming platforms, making it difficult to identify movies that match their interests and preferences.

2. Inefficient Browsing Experience: Traditional methods of browsing through movie catalogs are inefficient and time-consuming, resulting in decision fatigue and dissatisfaction among users.

3. Lack of Personalization: Existing recommendation systems often fail to deliver personalized recommendations, relying on generic algorithms that do not take into account individual user preferences and viewing habits.

4. Opportunity for Improvement: There is a clear opportunity to leverage machine learning algorithms to analyze user behavior and movie attributes, thereby offering personalized movie recommendations tailored to the unique preferences of each user.

By addressing these challenges and developing a small-scale movie recommendation system powered by ML algorithms, this project aims to provide users with a more efficient, personalized, and satisfying movie-watching experience..

3.3 Project Design

Figure 1 : Flowchart

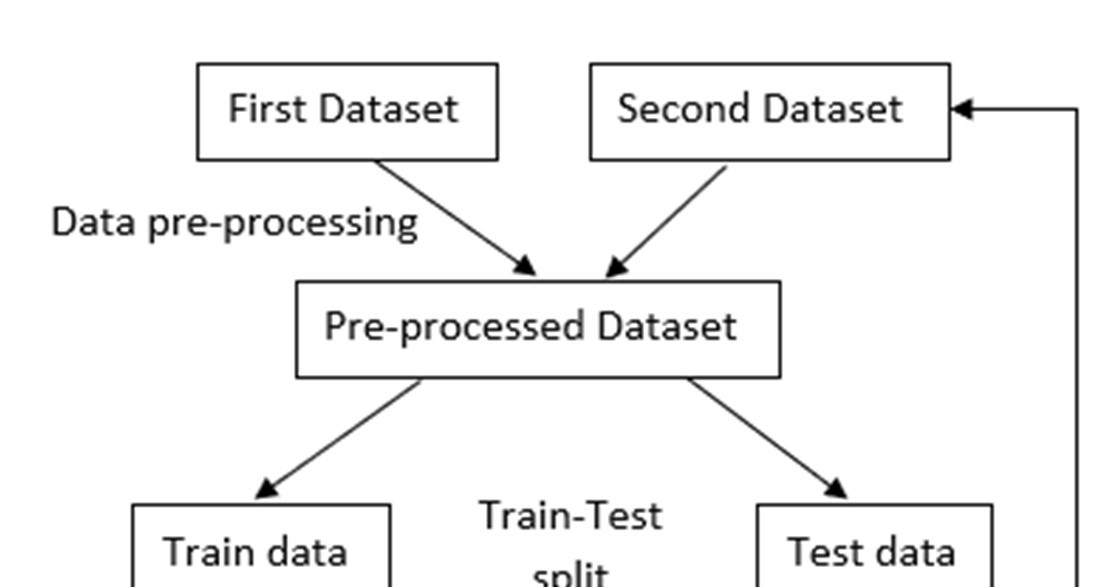




Figure 2 : Block Diagram

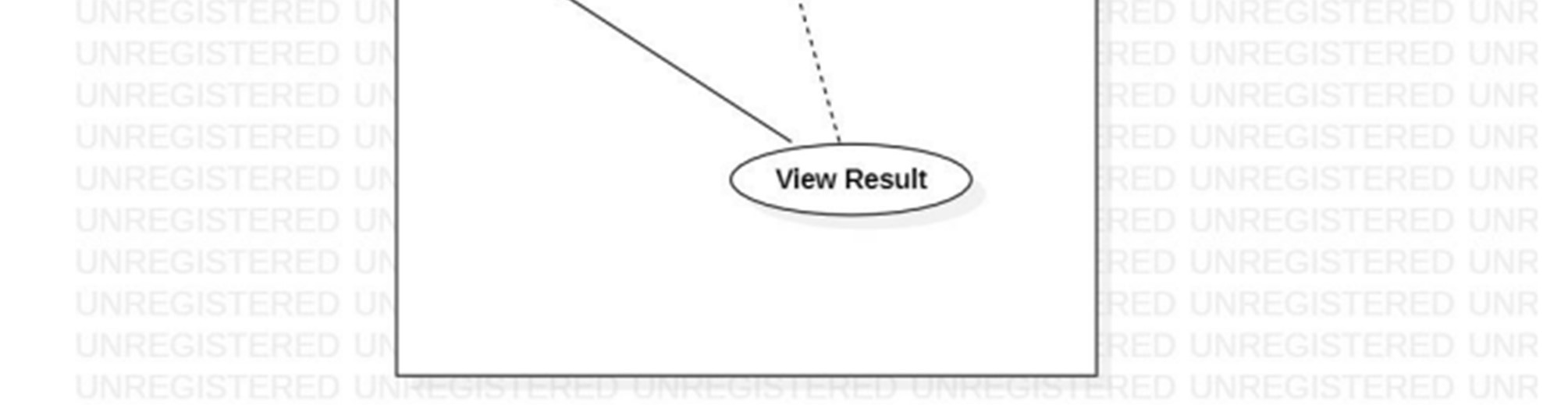
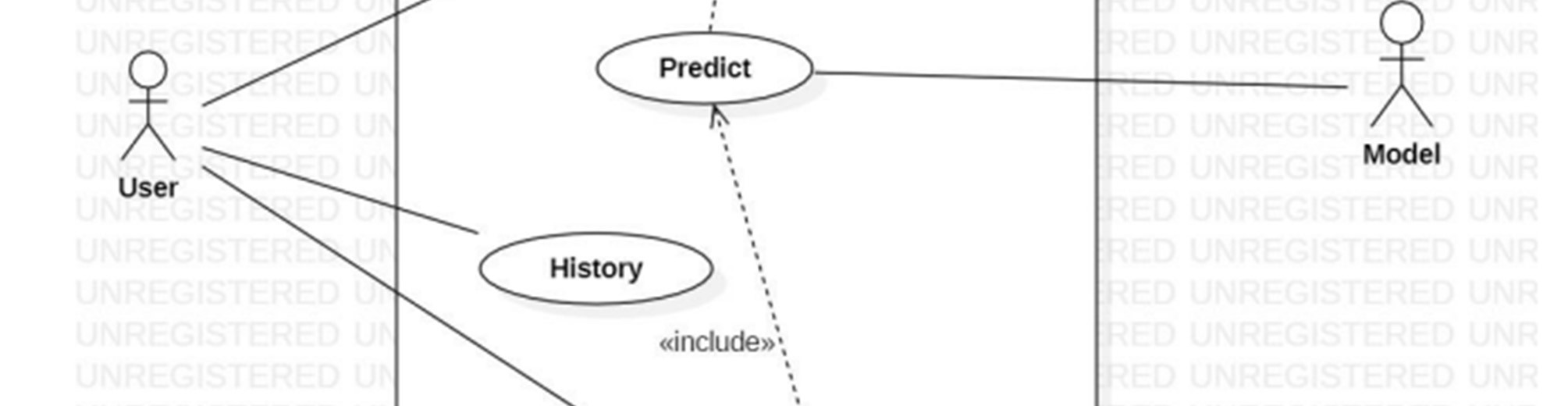
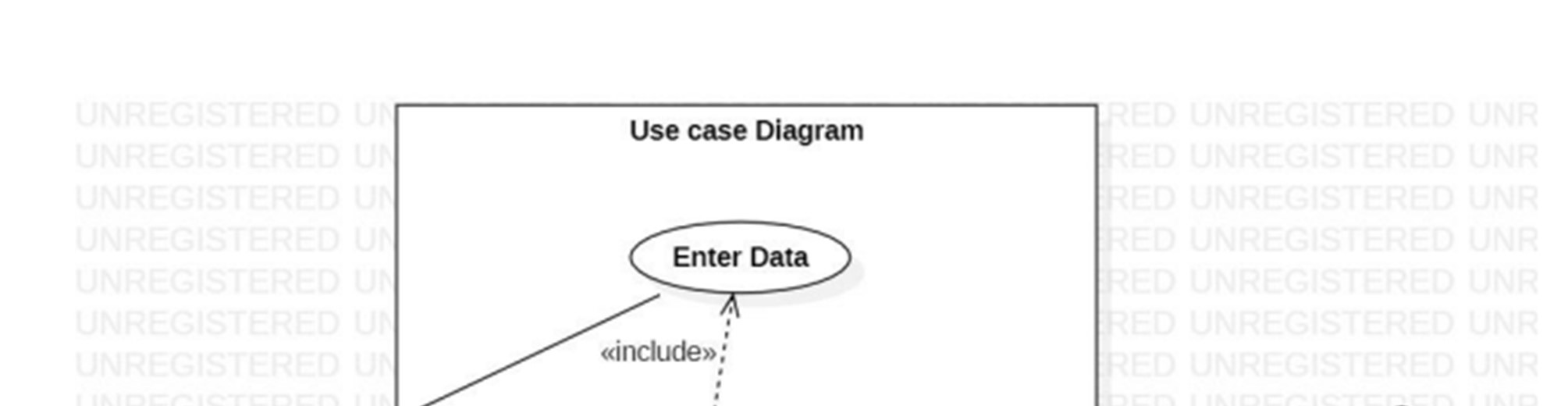


Figure 3 : Use Case Diagram

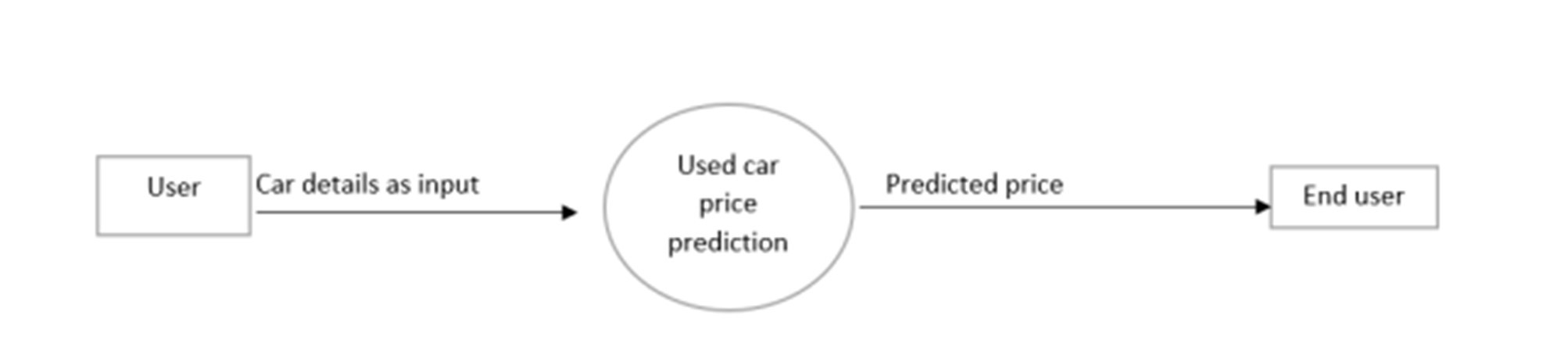


Figure 4 : DFD Level 0

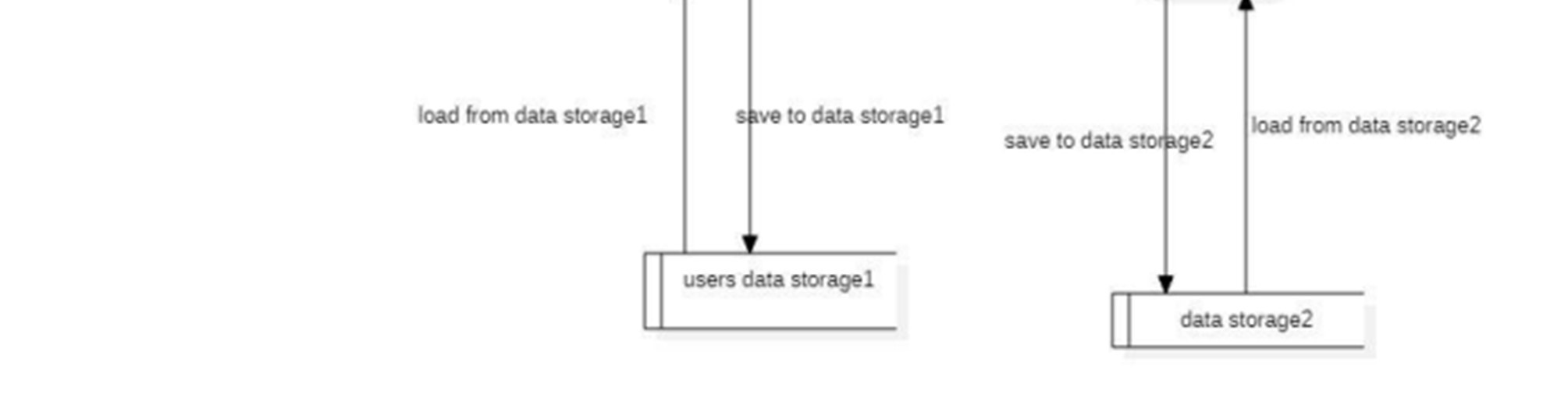
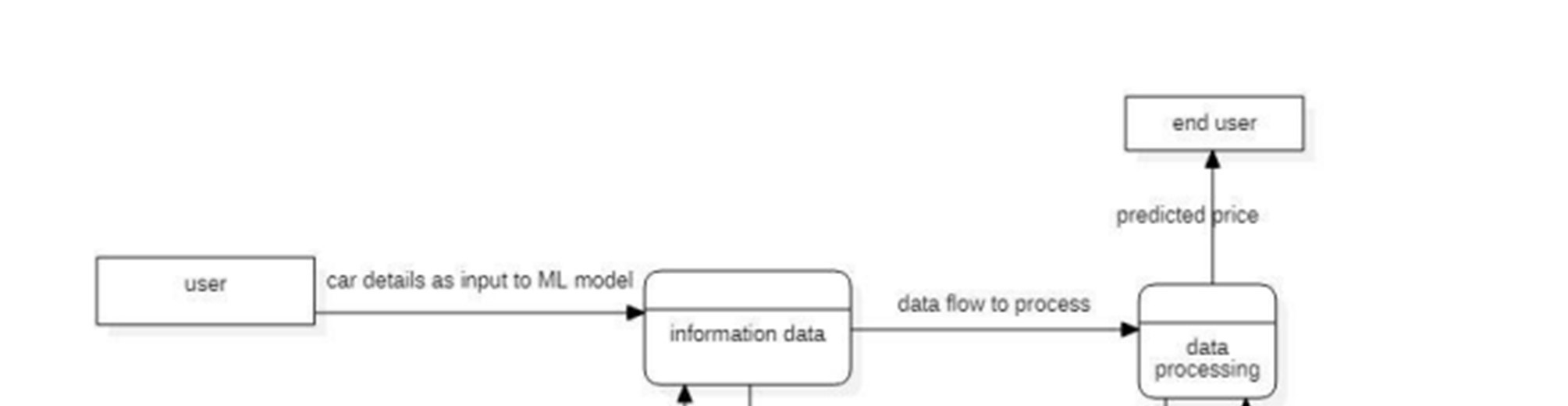
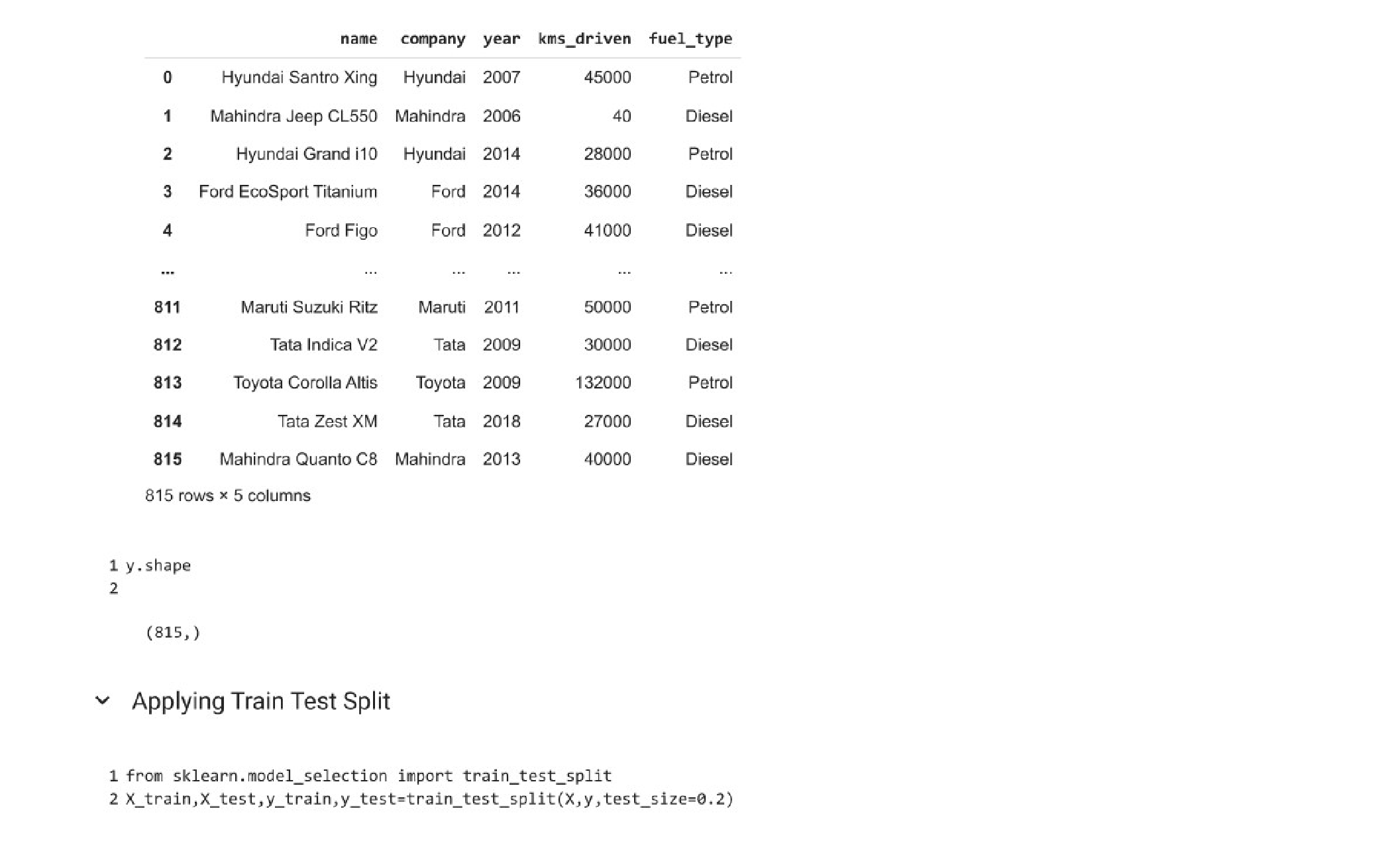
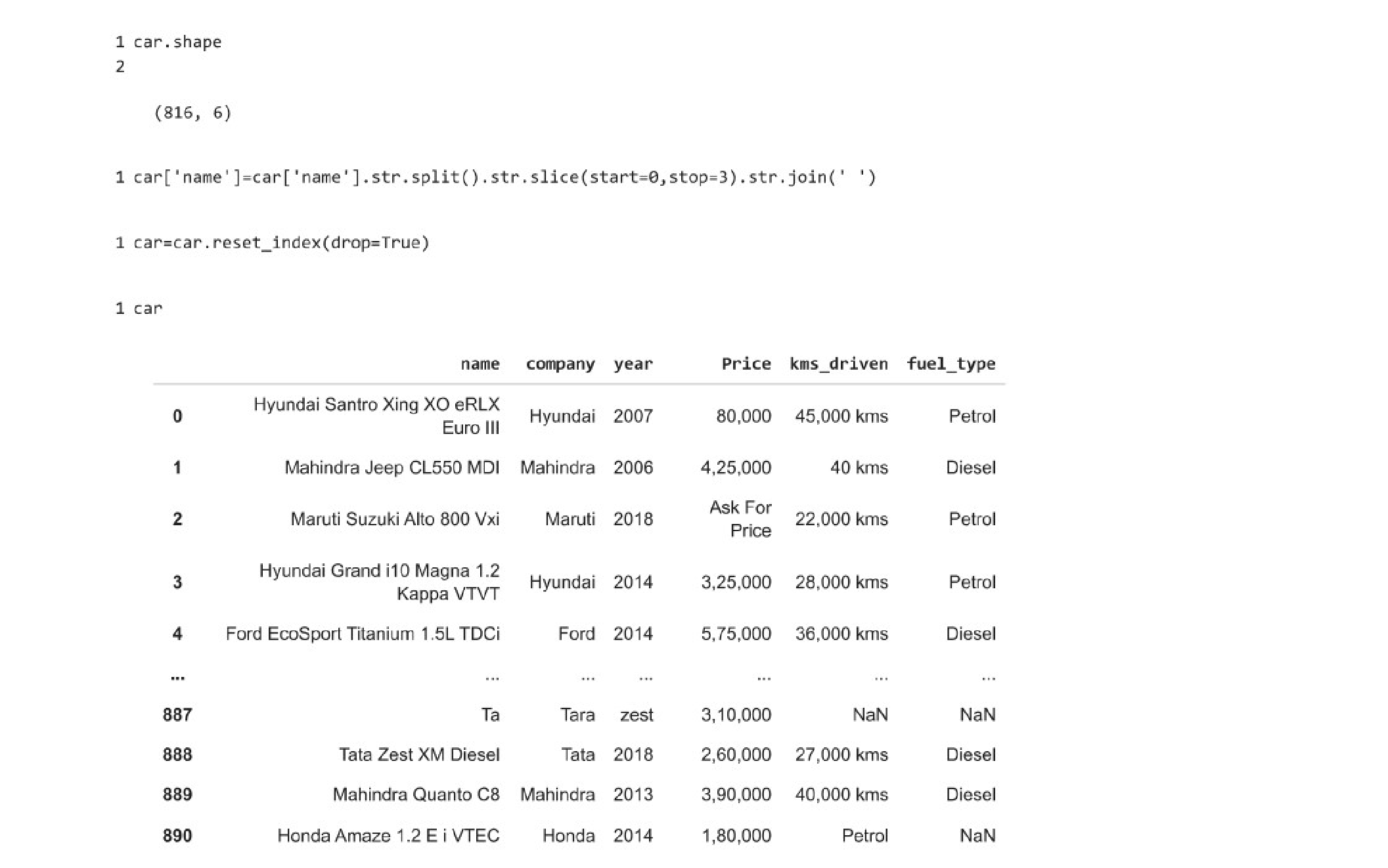
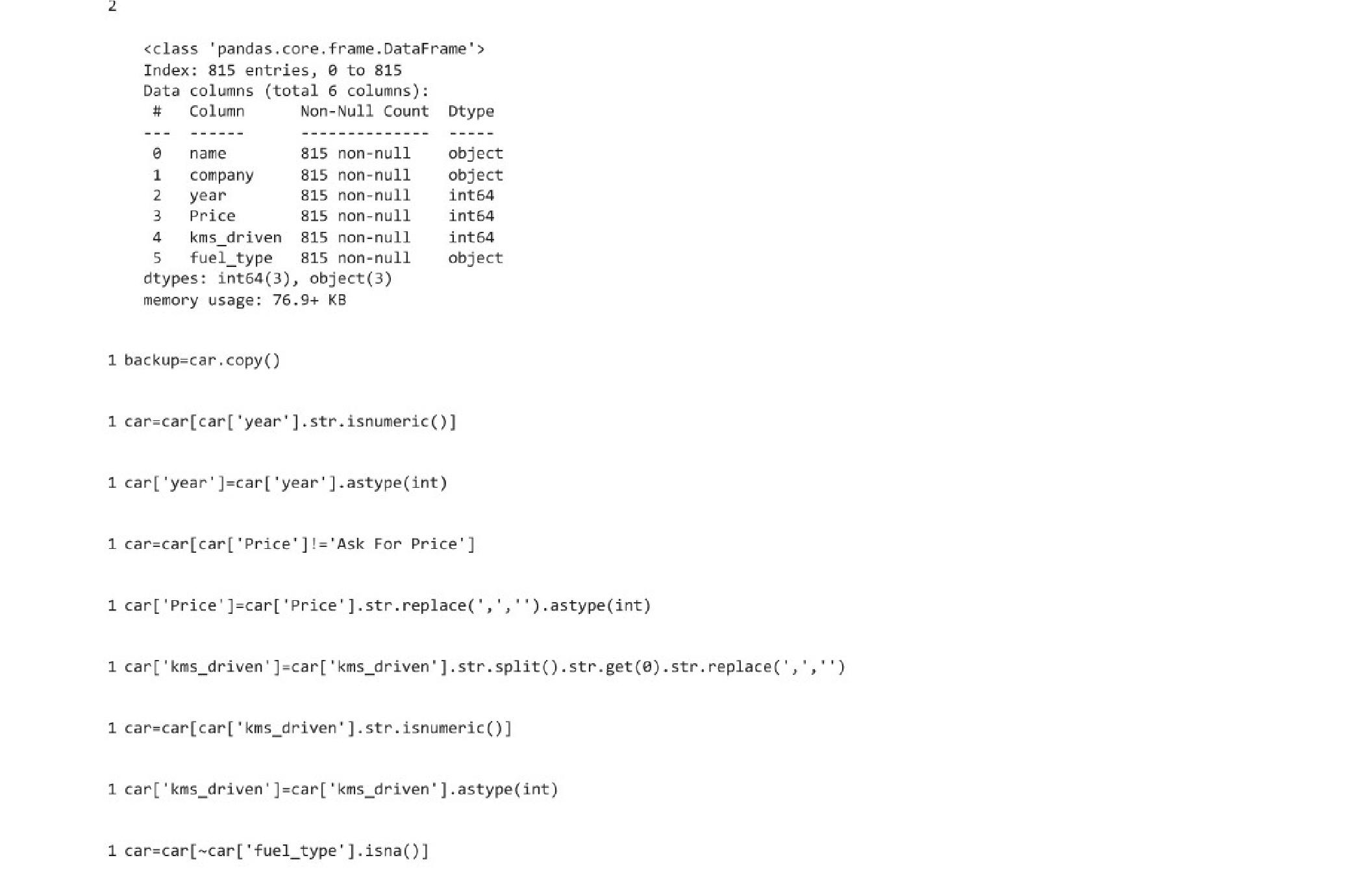
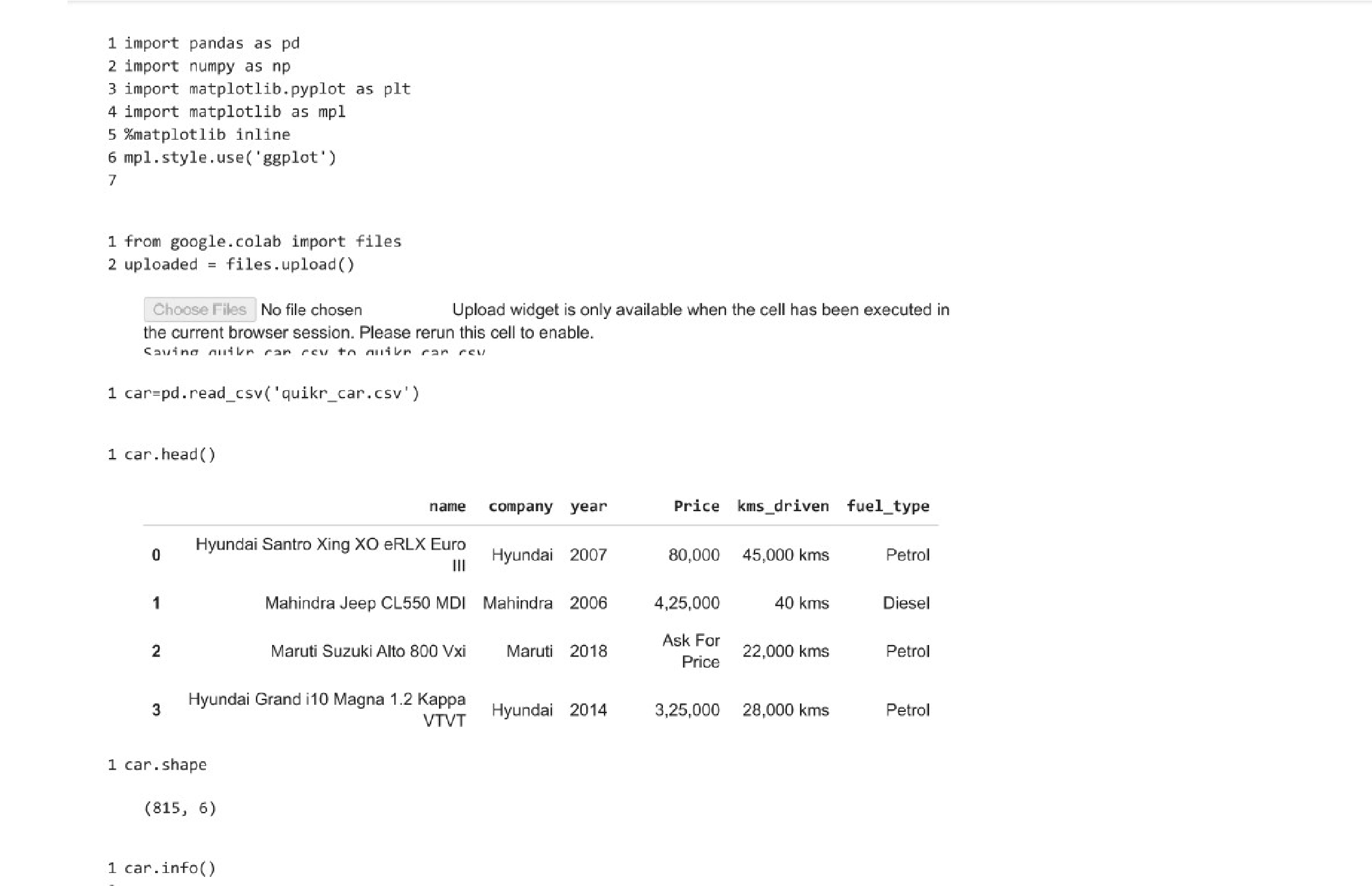


Figure 5 : DFD Level 1

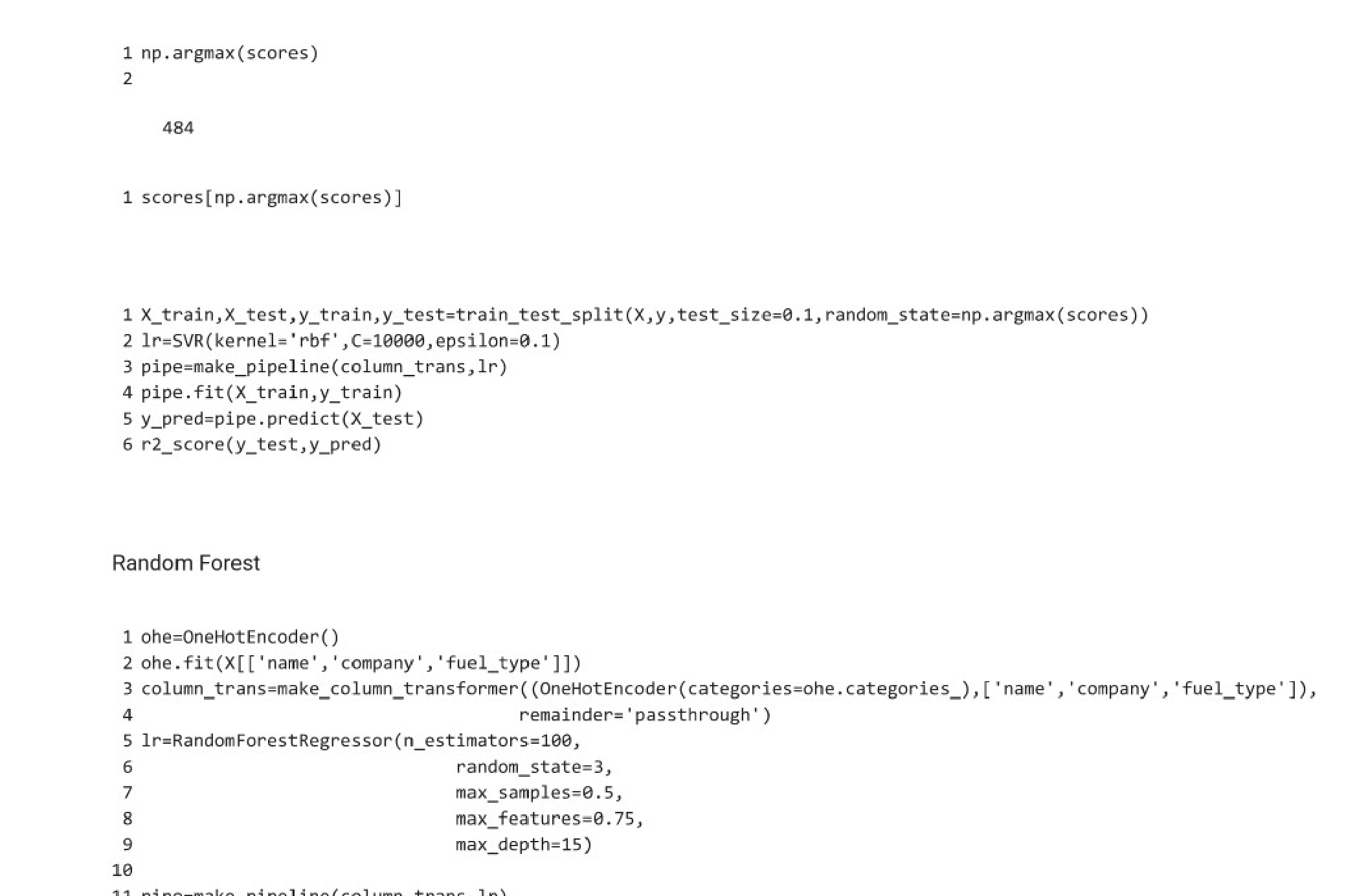
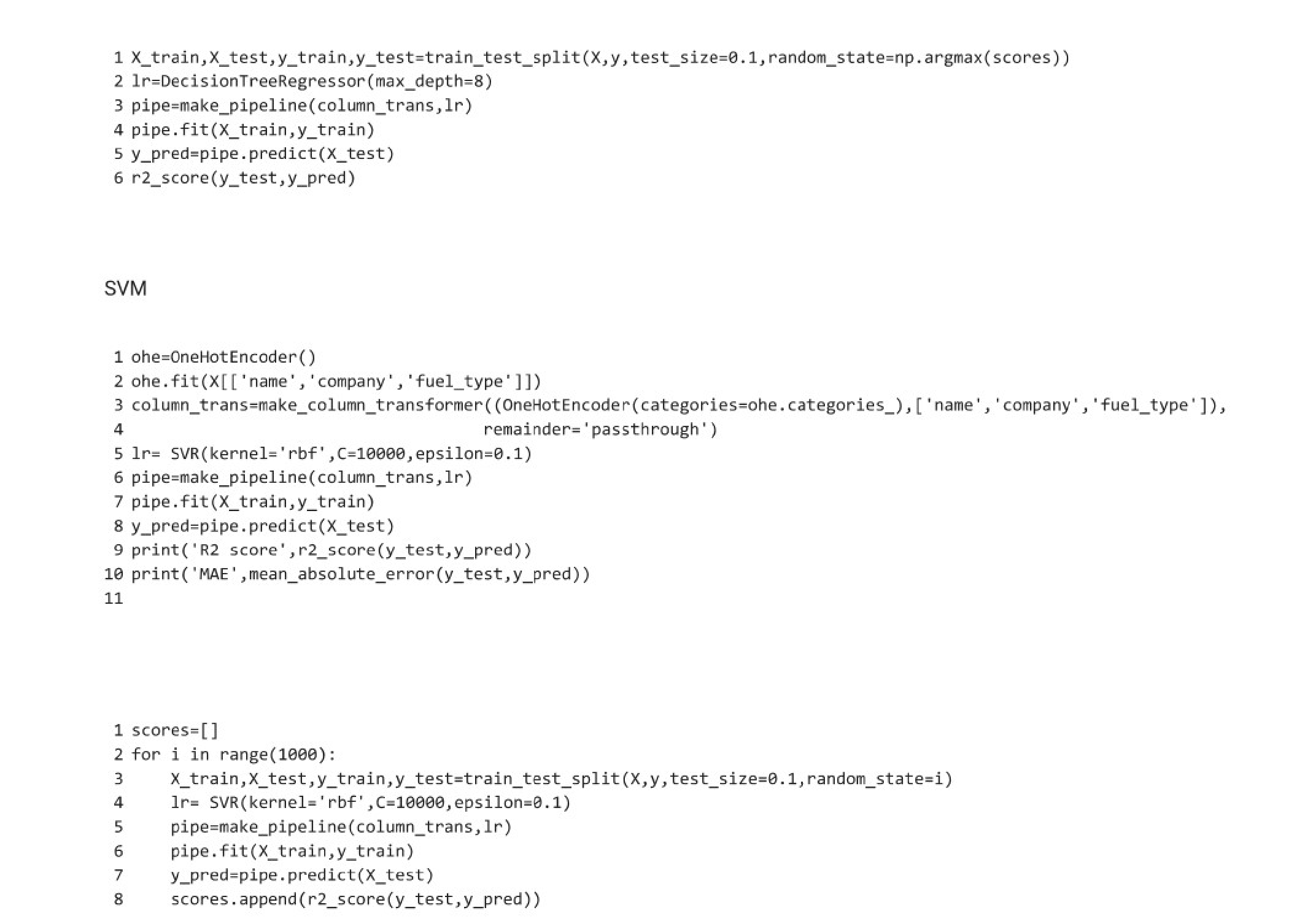
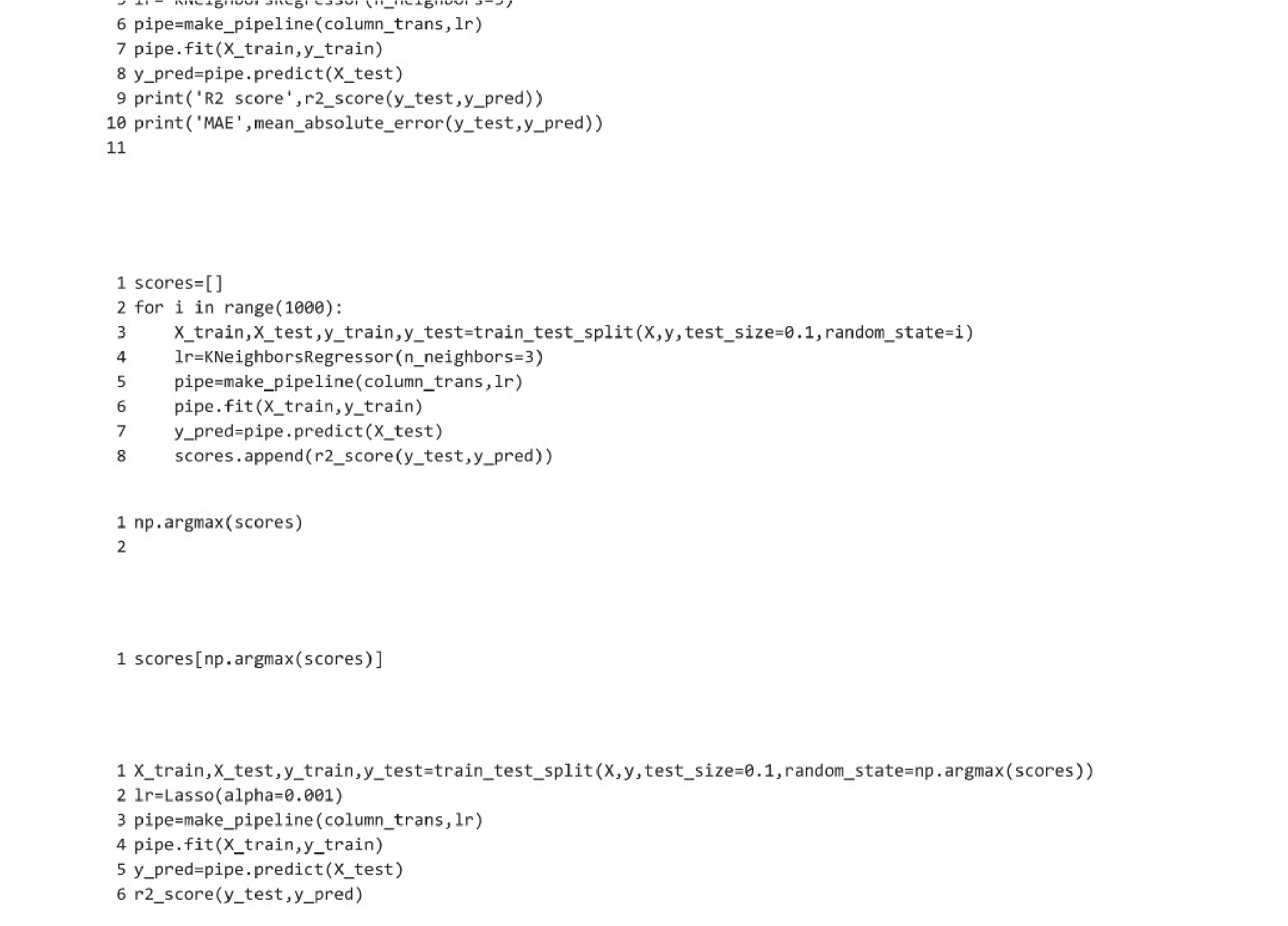


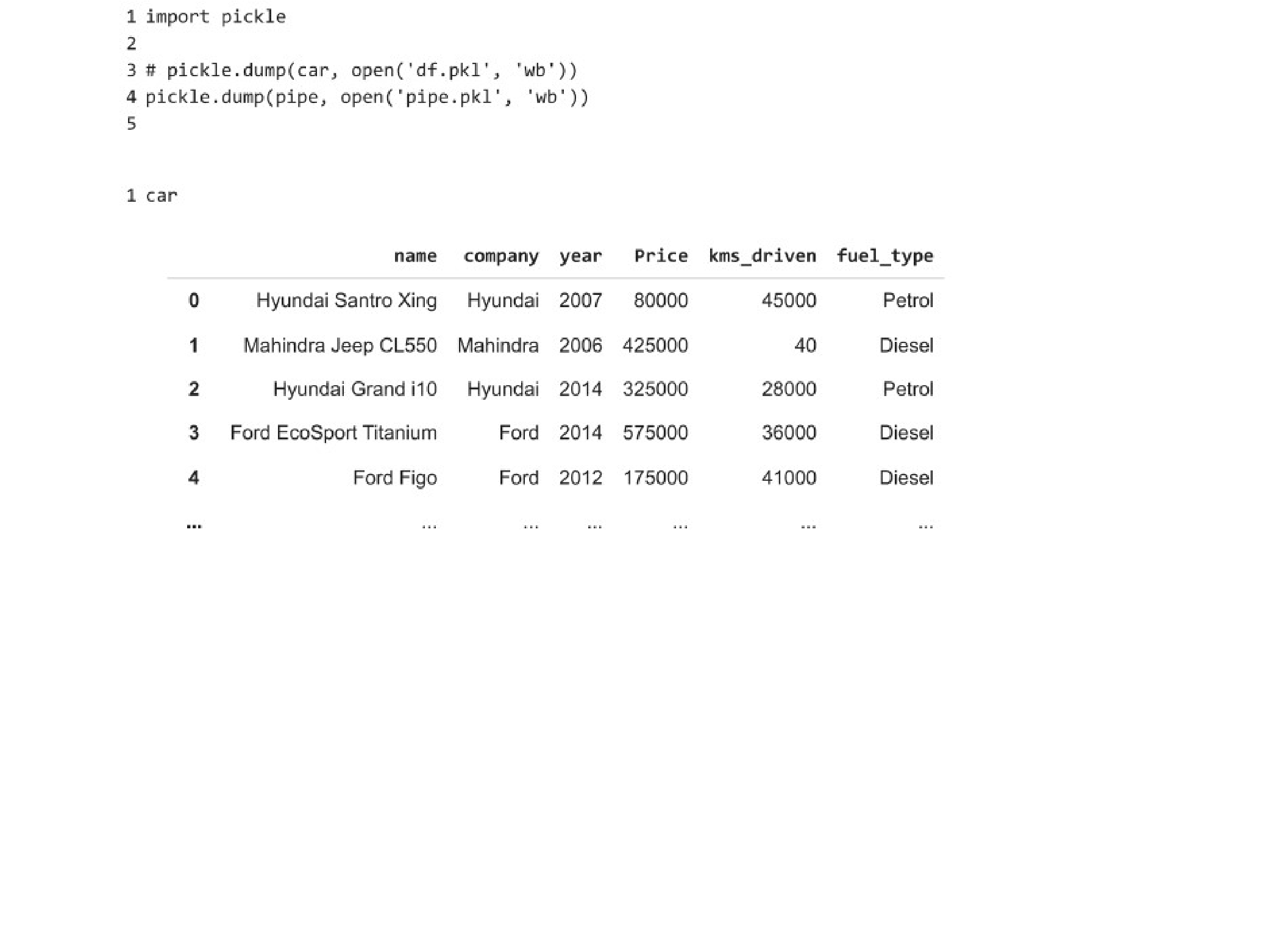
Figure 6 : Output Screenshot

### Code









# Methodology

The methodology for developing the Used Car Price Predictor project involves several key steps, including data collection, pre-processing, model development, evaluation, deployment, and ongoing maintenance. Here's a detailed breakdown of the methodology:

1. Data Collection:

* Gather comprehensive data on used cars from reliable sources, including attributes such as make, model, year, mileage, condition, region, and historical sales prices.
* Consider multiple data sources, including public datasets, APIs, web scraping from online listings, and collaboration with automotive industry partners for access to proprietary data.

2.Data Pre-processing:

* Cleanse the collected data to handle missing values, outliers, and inconsistencies using techniques such as imputation, outlier detection, and data validation.
* Perform data transformation and feature engineering to extract relevant features, encode categorical variables, and normalize numerical data for ML model compatibility.

3. Machine Learning Model Development:

* Select suitable ML algorithms for price prediction, such as decision trees, random forests, gradient boosting, or regression models, based on the nature of the data and prediction task.
* Split the pre-processed data into training and testing sets for model development and evaluation. - Train the ML models using the training data, optimize hyperparameters through techniques like grid search or random search, and validate the models using cross-validation methods

4. Model Evaluation:

* Evaluate the trained ML models using performance metrics such as mean absolute error (MAE), root mean squared error (RMSE), and R-squared (R2) to assess predictive accuracy and generalization.
* Conduct sensitivity analysis and feature importance ranking to identify the most impactful variables influencing price predictions.

5. User Interface Development:

* Design and develop a user-friendly web interface or mobile application for the Used Car Price Predictor system. - Implement input forms for users to enter car details (make, model, year, mileage, condition) and receive predicted price ranges based on the trained ML models.
* Incorporate visualizations or insights to enhance user experience and decision-making.

6. Deployment:

* Deploy the trained ML models and user interface components on a suitable hosting platform or server infrastructure.
* Ensure scalability, security, and performance optimizations for handling concurrent user requests and large datasets.
* Conduct thorough testing (unit testing, integration testing, performance testing) to validate system functionality and reliability.

# 4. Conclusion and Future Scope

Conclusion:

The development of the Used Car Price Predictor project has been a significant endeavor aimed at addressing the pricing challenges prevalent in the used car market. Through the application of machine learning algorithms, data preprocessing techniques, and user-friendly interfaces, the project has successfully achieved its objectives of providing accurate price predictions based on vehicle specifications and market dynamics.

The system's deployment offers tangible benefits to both buyers and sellers, empowering them with data-driven insights for making informed decisions and facilitating fair transactions. The project's methodology, including data collection, preprocessing, model development, evaluation, and deployment, has laid a solid foundation for the system's functionality and reliability.

Future Scope:

While the current version of the Used Car Price Predictor system fulfils essential pricing needs, there are several avenues for future enhancement and expansion:

1. Real-time Data Integration: Incorporate real-time data streams for dynamic pricing predictions based on the latest market trends, demand fluctuations, and economic indicators.
2. Enhanced Feature Set: Include additional features such as vehicle specifications (engine size, fuel efficiency, safety ratings), sentiment analysis of customer reviews, and market sentiment analysis for more nuanced price predictions.
3. Geographical Expansion: Extend the system's coverage to encompass multiple geographical regions or global markets, adapting the models and data sources accordingly.
4. Advanced ML Techniques: Explore advanced ML techniques such as deep learning models (e.g., neural networks, recurrent neural networks) for improved predictive accuracy and pattern recognition.
5. User Experience Improvements: Continuously enhance the user interface with intuitive features, personalized recommendations, interactive visualizations, and multi-platform compatibility (web, mobile, desktop).
6. Integration with Automotive Ecosystem: Collaborate with automotive industry stakeholders, dealerships, insurance providers, and regulatory bodies to integrate additional data sources, compliance checks, and value-added services.
7. Predictive Analytics: Utilize predictive analytics for forecasting future trends, price trends, and market insights, enabling proactive decision-making and strategic planning.
8. Ethical Considerations: Address ethical considerations related to data privacy, bias mitigation, fairness in pricing predictions, and regulatory compliance, ensuring transparency and accountability in system operations.

By exploring these future scope areas, the Used Car Price Predictor system can evolve into a comprehensive and indispensable tool for stakeholders in the automotive industry, fostering transparency, efficiency, and trust in the used car market.

# References

[1] https://medium.com/odscjournal/predicting-car-prices-usingmachine-learning-and-data-science-52ed44abab1b

[2]https://www.divaportal.org/smash/get/diva2:1674070/FULLTEXT0 1.pdf

1. https://github.com/suhasmaddali/Car-Prices-Prediction

1. https://youtu.be/iRCaMnR\_bpA?si=T0aCQnDlt37RnN25

1. https://youtube.com/playlist?list=PLPL68eAk13fsESpD9\_-2fl6zb-

PFknvd8&si=hxbhh9QxJnH3ov1\_