

# CARIFY – Predicting Car Maintenance Costs Using Artificial Intelligence

Shreya Pawaskar<sup>1</sup>, Akshata Jedhe<sup>1</sup>, Juee Ashtaputre<sup>1</sup>, Payal Mehta<sup>1</sup>, Riya Kulkarni<sup>1</sup>

<sup>1</sup>Student, Department of Computer Engineering, MKSSS's Cummins College of Engineering, Pune, Maharashtra, India

\*\*\*

**Abstract** - It is difficult to tell which car will cost a lot to maintain because there are so many different makes and models of cars available nowadays. Depending on several factors, such as mileage, the age of the vehicle, the fuel type, the city, and the model of the car, there should be a system that shows the maintenance and health condition of various automotive models or automobile parts. This method indicates when auto parts need to be maintained or replaced. The upkeep of cars varies from state to state and from city to city. For instance, vehicles in mountainous terrain require more maintenance than those in urban areas. Therefore, this examination has to be performed before purchasing an automobile. It is a need of time to develop a system that can accurately predict maintenance and the probability of car parts being changed. Artificial Intelligence has transformed the industry of automobiles. Predictive systems can use machine learning algorithms for high accuracy in the predictive analysis of vehicles.

**Key Words:** Maintenance cost, Maintenance index, Vehicles, Machine learning, Accuracy, Probability

## 1. INTRODUCTION

Many websites compare the costs of vehicles based on their price, mileage, tire size, etc. However, no website compares the cars based on their maintenance costs. Maintenance costs are required to maintain the vehicle's parts and service them monthly, quarterly, or yearly. We have proposed a system, Carify, that calculates the maintenance cost of a car based on its model, the company, the fuel type, the age of the vehicle, city, and mileage. Based on these factors, we determine which car parts need to be replaced or serviced and after how many months. These car parts include an air cleaner filter, engine oil, engine oil filter, sump plug gasket, AC dust filter, climate control air filter, fuel filter, engine coolant, spark plug, oil, brake, windshield wiper, clutch battery, pollen filter, transmission fluid, and drain washer. The maintenance cost varies from city to city, and labor costs, which are included in the maintenance cost, also differ from city to city.

Carify means "We Care for Your Car." This benefits consumers, car manufacturers, the government, and second-hand car buyers.

## 1.1 Consumers

They can decide to buy a car based on the maintenance index. This enables the customer to select the car models according to the patterns and usage in zones, states, cities, and regions, further drilling down concerning the vehicle's aging, usage patterns, and mileage.

## 1.2 Car Manufacturers

With the available data, OEMs will be able to design parts for more robustness. This will reduce warranty and service costs. Companies can analyze the performance of their models concerning their competitors in a particular region.

## 1.3 Consumers

Enable the government to draft the policies and approvals for new model launches in the market.

## 1.4 Second-hand Car Buyers

Customers can decide to buy the most suitable car available for sale.

This project is based on the 3C's: Collect, Calculate, and Compare. It compiles information from reliable sources, determines a vehicle's maintenance costs based on specific criteria, and then contrasts those costs with those of other vehicles. It compares the data city-wise since maintenance costs vary from city to city. Customers will not have to spend as much time browsing and comparing information across different platforms, showing how likely it is that particular components will change.

## 2. LITERATURE SURVEY

Previous research has investigated the significance of optimal car maintenance in the automotive industry and developed models to evaluate maintenance's effect on car performance. Predictive maintenance enabled by machine learning Use cases and challenges in the automotive industry by Andreas Theisslera, Judith Pérez-Velázquez, Marcel Kettelgerdes, and Gordon Elger focus on predictive maintenance (PdM) [1]. The challenge of maintaining functional safety while minimizing maintenance costs has

grown significantly in the automotive industry. Machine learning is perfect for PdM since current automobiles produce considerable operating data.

Krzysztof Danilecki, Jacek Elias, and Piotr Smurawski, "Modeling Inventory and Environmental Impacts of Car Maintenance and Repair: A Case Study of the Ford Focus Passenger Car." Wojciech Stanek and Andrzej Szlak study different maintenance scenarios [2]. As cars age, repairs and maintenance become more frequent and use more resources. The study's primary case study vehicle was the well-known compact car, the Ford Focus II. It examined various maintenance scenarios for passenger cars with internal combustion engines using the Life Cycle Assessment (LCA) method. The analysis took into account a variety of maintenance scenarios (the Ford Maintenance Schedule, reported car breakdowns, etc.). Information from 40 vehicles frequently visiting Ford-approved service centers was used to analyse maintenance and repair scenarios. The inventories were merged to model various auto maintenance scenarios. A vehicle scenario was modeled in ICE diesel and gasoline variants to compare maintenance [3], upkeep, and cost. Each vehicle scenario was modeled in both ICE diesel and gasoline versions—repair options and data from previous studies. The key finding is that, compared to previous studies, the new degree of thoroughness of inventory data has a 58% greater influence on automobile maintenance and repair in a petrol car and a 95% higher effect in a diesel car. The impact is twice as significant as a shortened oil change interval [6].

In Reliability Analysis of Car Maintenance Scheduling and Performance by Ghassan M. Tashtousha, Khalid K. Tashtoush, Mutaz Al-Muhtaseba, and Ahmad T. Mayyash, a statistical model is developed to evaluate the effect of corrective and preventive maintenance schemes on car performance in the presence of system failure where the scheduling objective is to minimize schedule duration[4][5]. When there is no possibility of failure, research has focused primarily on scheduling and performance control issues for auto maintenance, whether with or without maintenance. In real-world situations, this might not be realistic [6][7]. This paper showed that neither strategy is superior to the other, nor their applicability depends on the scheduling environment. They also showed how parameter values could favor preventive maintenance over corrective maintenance. The study's conclusions will help professionals and system administrators plan auto repairs and other jobs [8-10].

### 3. DESIGN METHODOLOGY

#### 3.1 General Layout

Carify has two essential features: knowledge collection and prediction model development. Each step in the process is crucial and would affect the outcome if carried out improperly, would affect the outcome. After data collection, processing plays a vital role in organizing the training and testing datasets [11]. Then the prediction was made using the appropriate machine learning algorithms.

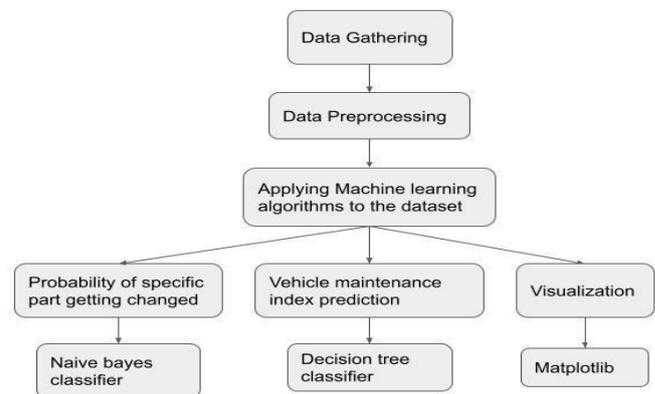


Fig -1: Design Methodology

#### 3.2 Data Source

The data is collected from official car websites of various car models and brands. The data from the website was in a table format, i.e., structured form. It has data like mileage, kilometers traveled, vehicle age, region, etc. The car itself collects driving data such as speed and braking patterns. Take note of multiple vehicle health indicators, such as RPM, tire pressure, ambient air temperature, oil temperature, mileage, fuel level, and oil temperature, to enable more sophisticated correlations. Data regarding car health and maintenance according to region and age of the car, collected from various websites, was unstructured.

#### 3.2 Data Pre-Processing

- Probability of specific part getting changed:

It is the probability that a specific car part(e.g., suspension system, wheels, tires, brakes, etc.) has to be changed after a certain period. To determine the probability, we have used the Naive Bayes Classifier.

- Vehicle maintenance index for cars and each car part:

The vehicle maintenance index is a value that rates a car's health on a scale of 1-10, with one meaning that maintenance is less, and the vehicle is healthy, and ten

means that high maintenance is required to make the car's life better and healthier. We have used a decision tree classifier to predict the vehicle maintenance index.

- Several car model comparisons can be made using data visualization:

Standard illustrations, diagrams, plots, infographics, and animation address data visualizations. They are straightforward visual data representations. They convey complex information relationships and information-driven experiences. With the help of the Python library Matplotlib, data visualization is employed to understand a car's condition better and compare various models.

#### 4. CLASSIFIERS USED

##### 4.1 Naive Bayes Classifier

Using the Bayes Theorem, we can determine how likely a hypothesis is based on the information we already have. Both binary and multi-class classification can be done using it. It can be trained on a small dataset and is straightforward to understand. Due to its high scalability, it grows linearly as more predictor features and data points are added. It can be used with continuous and discrete data and is not affected by minor details. It has been used for many different things, including face recognition and recommendation systems, but it shines regarding NLP issues. It helps predict data and develop hypotheses based on previous results.

##### 4.2 Decision Tree Classifier

This machine learning method, under supervision, constantly divides the data into groups based on a particular parameter. It is inexpensive to develop and quickly classifies unknown records. They can categorize even data for which no attributes are available. They can easily manage attributes that are both categorical and numerical in combination. The tree and its decision-making process can be graphically represented using a tree-like layout where the decisions are made at the leaf nodes and terminal nodes. Even a novice could understand this right away.

### 5. RESULTS

Carify Web Application has five pages.

#### Home page:

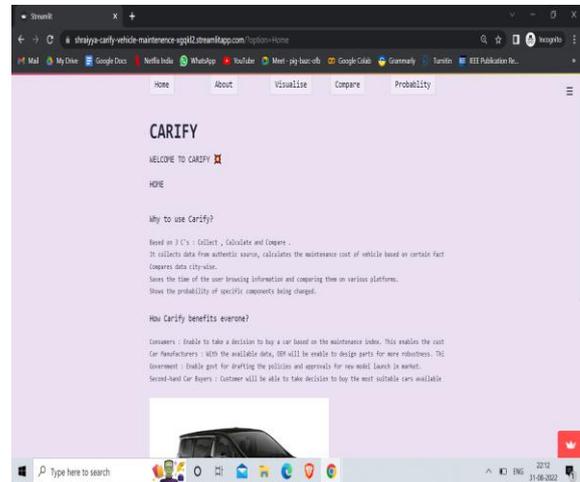


Fig -2: Home Page

**About page:** The about page contains details about the application and defines the need for the application

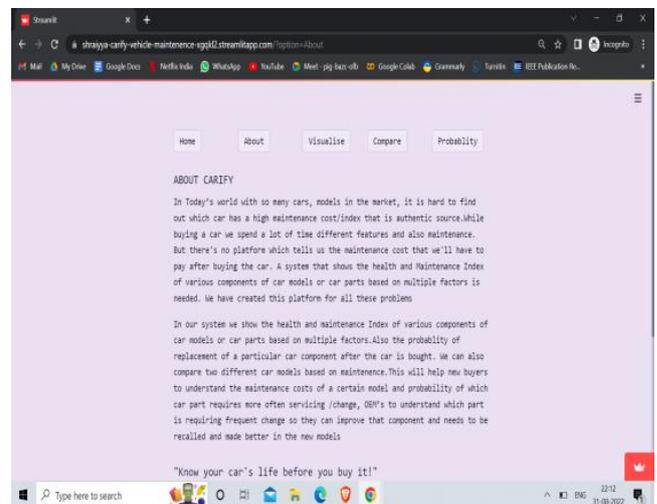


Fig -3: About Page

**Visualize** - The user can see different data visualizations here.

**Compare** - The user can compare two different cars here. In this example, the first car details are a Hyundai Creta Petrol Model of Mumbai city aged 54 months. The second car details are of a Honda Amaze Diesel Model of Mumbai city with an age of 28 months. After clicking on the compare button, the records of each model are shown with the costs.

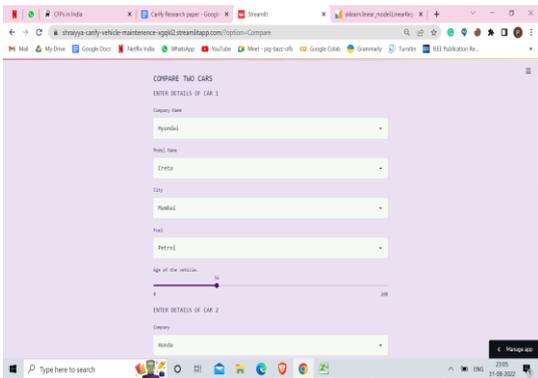


Fig -4: Details of Car One

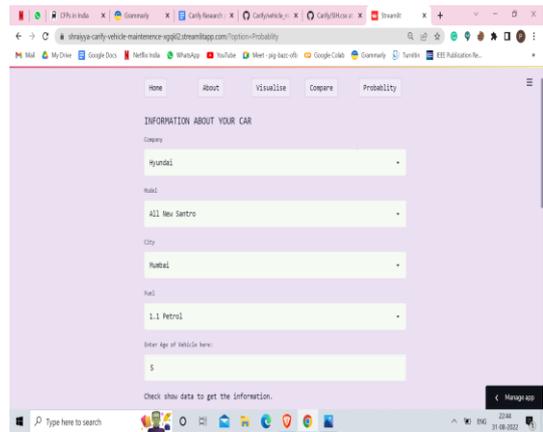


Fig -7: Details of a car to check the probability of part being changed

In the above example, the user chooses a Hyundai 'All new Santro model' of a '1.1 petrol' variant in Mumbai city at five months.

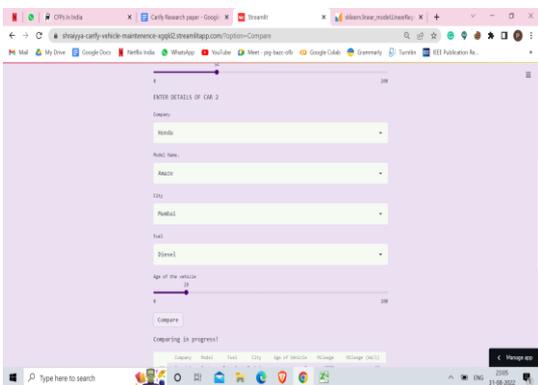


Fig -5: Details of Car Two

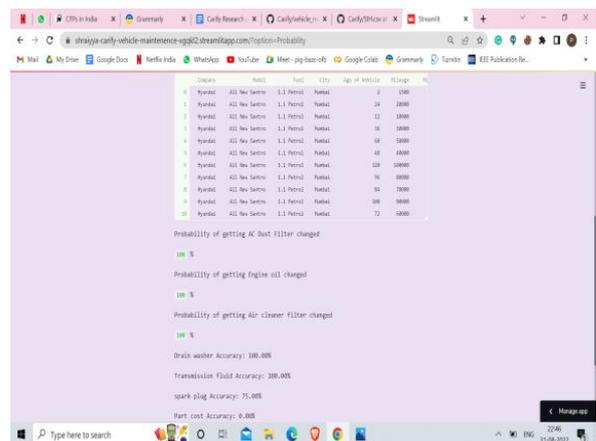


Fig -8: Car Part Change Details

After pressing the enter button, the user can see the database of model details similar to the one he entered. Here the probability of an AC dust filter, Engine Oil, Air Cleaner Filter, Drain washer, and Transmission fluid is 100 percent. The spark plug change accuracy is 75 percent.

## 6. FUTURE WORK

Even though this maintenance index calculator system achieves good accuracy scores using ML algorithms, there can be an improvement by collecting and feeding more data to the algorithm. Data collection is essential for the system described in this paper because results are generated based on data analysis. Web scraping, a technique that can extract many data from a website, can be used, for instance, to gather information from the official websites of different automakers, including Hyundai, Ford, Toyota, and others. This unstructured data

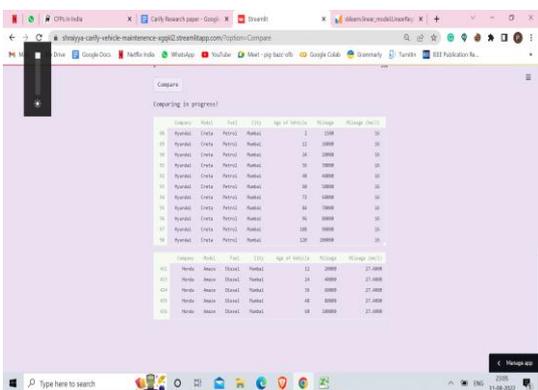


Fig -6: Comparison between two cars

**Probability** - The probability page gives the probability of different parts being changed. The user has to input the company, car model, fuel variant, city, and the number of months. After pressing enter, the user can see details about the car part change probability.

can be transformed into structured data and then fed into the algorithm to improve accuracy. Another method for data collection can be to have a tie-up with the car companies and get official, authorized data.

Deep Learning systems can be used to forecast outcomes with accuracy. Deep learning is a subfield of machine learning where neural network algorithms inspire the workings of the human brain. The Artificial Neural Network (ANN) is another widely used technique for prediction. ANNs are fully connected multi-layer neural networks. Neurons in convolutional neural networks (CNN) have biases and teachable weights.

Additionally, CNNs can aid in enhancing forecasts. The possibilities include creating a machine learning online application in Python and an Android application in Java where the user may input a car model and car manufacturer and receive the most accurate results. Based on the input, the software should accurately predict the maintenance index and the likelihood that specific parts will need to be changed. On the screen, the forecasts should be displayed. For new users, the user interface should be simple and intuitive.

## 7. CONCLUSION

Car maintenance is one of the most critical factors regarding road safety. No matter how advanced a company's engines and machines are, we must inspect each car part for safety. However, with many makes and models available today, it might be challenging to determine which car has a high maintenance cost/index from a trustworthy source. Based on several parameters, the model described in this paper calculates the maintenance index for various automobile models or parts. This system helps consumers and second-hand buyers to find the most suitable car according to the patterns and usage in zones, states, cities, and regions, further drilling down concerning vehicle aging, usage patterns, and mileage. It also saves users time browsing information and comparing it on various platforms. It makes it possible for the government to make policy decisions and for automakers to assess how well their models perform compared to their rivals in a given area. The decision tree classifier obtains the highest accuracy for the maintenance index calculator.

## REFERENCES

[1] "Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry," by Andreas Theissler et al. *Reliability engineering & system safety*, 215 (2021): 107864.

[2] Danilecki, Krzysztof, et al. "Modeling inventory and environmental impacts of car maintenance and repair: A case study of the Ford Focus passenger car." *Journal of Cleaner Production* 315 (2021): 128085.

[3] Yang, Jianxin, Meng, Fanran, Zhang, Zhan, Sun, Xin, and Luo, Xiaoli. (2020). Life cycle assessment of lithium nickel cobalt manganese oxide (NCM) batteries for electric passenger vehicles. *Journal of Cleaner Production*. 273. 123006. [10.1016/j.jclepro.2020.123006](https://doi.org/10.1016/j.jclepro.2020.123006).

[4] Al-Muhtaseb Motaz, Mayyas Ahmad, "Reliability Analysis of Car Maintenance Scheduling and Performance," *Jordan Journal of Mechanical and Industrial Engineering (JJMIE)*, vol. 4, no. 3, pp. 388–393, 2010

[5] B. Pattanaik and S. Chandrasekaran, "Recovery and reliability prediction in fault tolerant automotive embedded system," 2012 International Conference on Emerging Trends in Electrical Engineering and Energy Management (ICETEEEM), 2012, pp. 257-262, doi: [10.1109/ICETEEEM.2012.6494477](https://doi.org/10.1109/ICETEEEM.2012.6494477).

[6] Dolas, Dhananjay R., and Sudhir Deshmukh. "Reliability Analysis of the Cooling System of a Diesel Engine." 3.2 (2015): 57-62.

[7] Arena F, Collotta M, Luca L, Ruggieri M, Termine FG. Predictive Maintenance in the Automotive Sector: A Literature Review. *Mathematical and Computational Applications*. 2022; 27(1):2. <https://doi.org/10.3390/mca27010002>

[8] Dragana Velimirović, Čedomir Duboka, Predrag Damjanović. Automotive maintenance quality of service influencing factors, October 2016, *Tehnicki Vjesnik* Vol. 23(No.5):Article No. TV-20140402074657, DOI:[10.17559/TV-20140402074657](https://doi.org/10.17559/TV-20140402074657)

[9] Ioan Virca, Dorel Badea. Study on the Predictive Maintenance of Vehicles and its Management Using the Specific "Keep the Machine Running" Application.

[10] Fabio Arena , Mario Collotta , Liliana Luca , Marianna Ruggieri and Francesco Gaetano Termine. Predictive Maintenance in the Automotive Sector: A Literature Review, 1 Introduction, 1.2 Predictive maintenance

[11] Mahmuda Akhtar, Sara Moridpour, "A Review of Traffic Congestion Prediction Using Artificial Intelligence," *Journal of Advanced Transportation*, vol. 2021, Article ID 8878011, 18 pages, 2021. <https://doi.org/10.1155/2021/8878011>