Dynamic Pricing Prediction for Cabs Using Machine Learning

[1]A. Prasanthi, [2]Yamini, [3]Hyndavi, [4]K. Vasanthi

[1,2,3,4] UG Student, Department of Computer Science Engineering,

Dr K V Subba Reddy College Of Engineering For Women, Kurnool, Andhra Pradesh, India

Abstract

Ride-on-demand (Rod) services like Uber and OLA cabs are becoming increasingly popular. Rod services employ dynamic pricing to strike a balance between supply and demand in an effort to improve service quality in order to assist both customers and drivers. However, dynamic prices frequently cause issues for passengers: They are unable to quickly make decisions due to frequently "unpredictable" prices. Therefore, it is essential to provide passengers with more information in order to address this issue, and a feasible solution is to forecast dynamic prices. In this paper, we focus on the estimation of dynamic prices by forecasting the price for each individual passenger order using the Rapido dataset as an example. Passengers will be able to alleviate their concerns by understanding, through price predictions, whether they will be able to obtain a lower price in the near future or in nearby locations. The prediction is carried out by learning the relationship between the dataset's features and dynamic prices. We train one linear model and test its output using real service data from various angles as a representative. In addition, on the basis of the model, we examine the contribution of features at various levels and determine which features are most responsible for dynamic prices. Finally, we use evaluation results to predict dynamic prices with an effective linear regression model. As accurate forecas, our hope is that the study contributes to passenger happiness

1. Introduction

In the face of uncertainty, supervised learning constructs a model that makes predictions based on evidence. A supervised learning algorithm trains a model to make reasonable predictions for the response to new data using a known set of input data and known responses to the data (output). If you already know the data for the output you want to predict, use supervised learning. Classification and regression methods are used in supervised learning to create machine learning models. Machine learning is a subset of artificial intelligence (AI). Instead of being explicitly programmed to do so, it focuses on teaching computers to learn from data and improve with experience. Algorithms in machine learning are taught to analyze large data sets for patterns and correlations in order to make the best decisions and predictions. Applications for machine learning get better with use and become more accurate with more data. Machine learning is used in everything from our homes to shopping carts to entertainment media to healthcare.

Concentric subsets of AI include both machine learning and its components, deep learning and neural networks. AI uses data processing to make decisions and forecasts. Without the need for additional programming, AI is able to process that data and use it to learn and become smarter thanks to machine learning algorithms. All of the subsets of machine learning that are below artificial intelligence are derived from it. Machine learning is included in the first subset; Deep learning and neural networks are contained within that.

The term "artificial intelligence" (AI) is used to describe the process of programming machines to behave and think like humans. Any machine that possesses human-like abilities like learning and problem-solving can also be included in this definition.
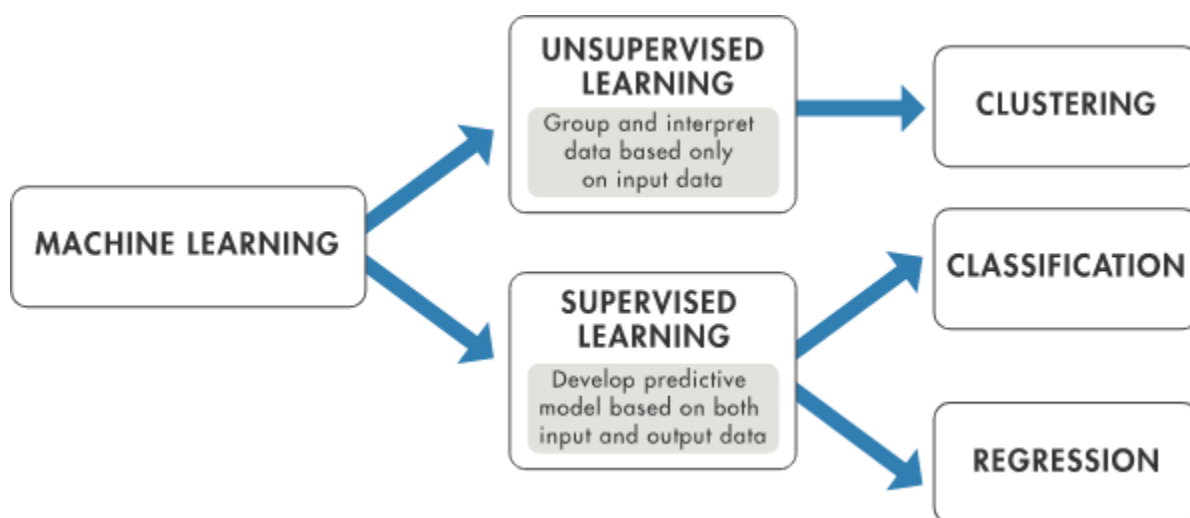

Fig.1 Working of Machine Learning

2. Literature Review

Python is a popular high-level, general-purpose programming language. Guido van Rossum developed it in 1991, and the Python Software Foundation added to it later. Programmers can express their ideas in fewer lines of code thanks to its syntax, which was made with code readability in mind when it was designed.

The programming language Python allows you to work more quickly and efficiently integrate systems. There are two main versions of Python: Python versions 2 and 3 They are very different. Users can create web pages with text, graphics, and hyperlinks (pointers to other web pages) using the world wide web's Hypertext Markup Language (HTML).

Although HTML is not a programming language, it is an application of SGML (Standard Generalized Markup Language) from ISO Standard 8879 that has been tailored for the Web and hypertext. Hypertext's concept of linking points together. We can sort through the information according to our preferences and interests. The elements that should be displayed in HTML are case-insensitive, so a markup language is nothing more than a collection of items enclosed within them. utilizing various fonts, sizes, colors, and graphics. can improve the document's presentation. The document itself includes anything that is not a tag.
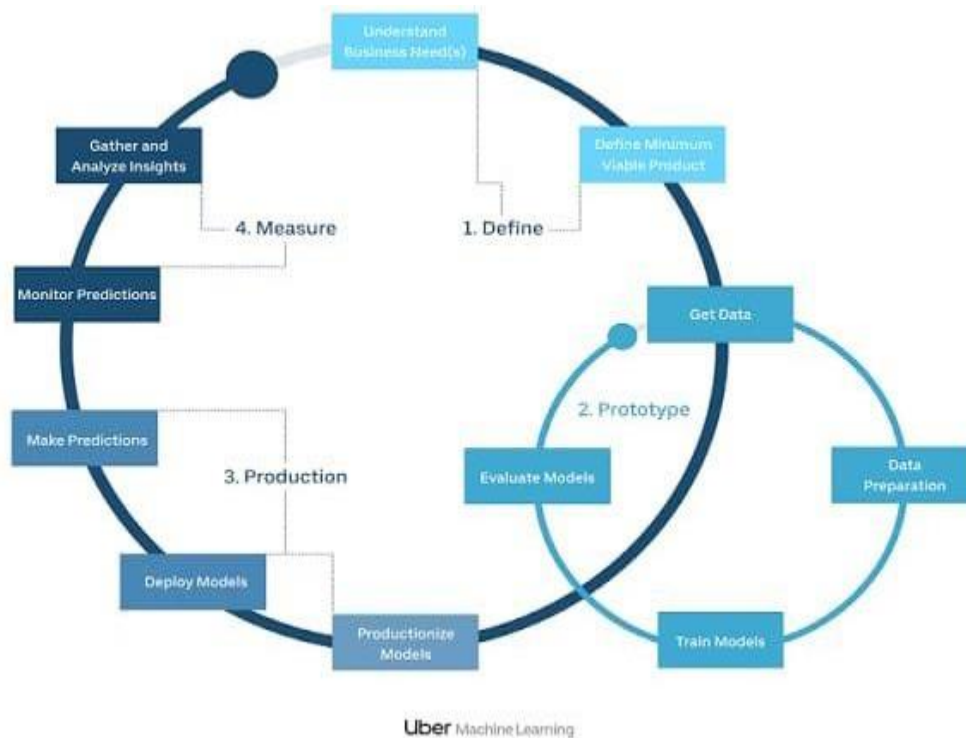
Fig.2 Analysis Diagram

3. Proposed System

Module load Langs/Python/2.7-anaconda is installed on Yale clusters Omega and Grace. The desktop graphical user interface (GUI) Anaconda Navigator is a part of the Anaconda distribution. It lets users manage conda packages, environments, and channels without having to use command-line commands. Navigator is able to find packages on Anaconda Cloud or in a local Anaconda Repository, install them in an environment, run the packages, and update them all at the same time. It is compatible with Linux, macOS, and Windows.

For data science and machine learning-related applications, Anaconda Navigator is a free and open-source distribution of the Python and R programming languages. Conda is a cross-platform, open-source package management system that can be installed on Windows, Linux, and macOS. Tools like JupyterLab, Jupyter Notebook, QtConsole, Spyder, Glueviz, Orange, Rstudio, and Visual Studio Code are included in Anaconda. We will be utilizing Jupiter notebook and spyder for this project. A Jupyter notebook is an electronic file that contains both text descriptions and programming code. Additionally, embedded charts, plots, images, videos, and links can be included in Jupyter notebooks. A web browser like Firefox or Google Chrome is required to operate Jupyter notebooks. Although many different programming languages' code can be found in Jupyter notebooks, Python code is typically found in many of them. The Python code found in a.py file is identical to that found in a Jupyter notebook.

Markdown-formatted explanations and clarifications of the programming code can be found
106

in the text description sections of Jupyter notebooks. The markdown file extension is.md. An open-source Integrated Development Environment is known as Spyder. It was developed by scientists, is written in Python, and is only intended for use by engineers, data analysts, and scientists. Scientific Python Development IDE is another name. It offers an Editor for writing code, a console for evaluating it and viewing the results at any time, a variable defined during evaluation, and a plethora of other features for efficient application or program development.

Spyder is a powerful Python-based scientific environment created by and for scientists, engineers, and data analysts. It is one of a kind because it combines the data exploration, interactive execution, deep inspection, and beautiful visualization capabilities of a scientific package with the advanced editing, analysis, debugging, and profiling capabilities of a comprehensive development tool. In addition, a number of well-known scientific applications, such as NumPy, SciPy, Pandas, Matplotlib, and SymPy, are built into Spyder. Spyder has a lot of built-in features, but third-party plugins can make it even more powerful. Spyder can also be used as a PyQt5 extension library, enabling you to extend its features and embed its parts., such as the interactive console or advanced editor, in your own software
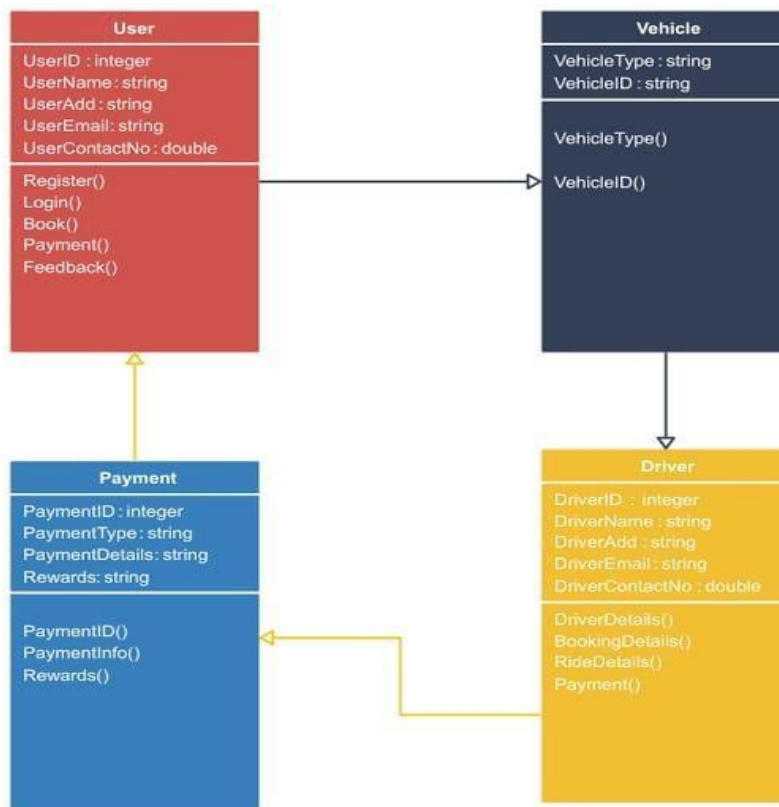

Fig.3 Class Diagram

Random Forest is also a "Tree"-based algorithm that uses the qualities features of multiple Decision Trees for making decisions. Therefore, it can be referred to as a 'Forest' of trees and hence the name "Random Forest". The term 'Random' is due to the fact that this algorithm is a forest of 'Randomly created Decision Trees'. NumPy: NumPy is a Python library used for working with arrays. It also has functions for working in domain of linear algebra, Fourier

107

transform, and matrices. NumPy was created in 2005 by Travis Oliphant. It is an open-source project and you can use it freely. NumPy stands for Numerical Python.

Pandas: Pandas is a Python library used for working with data sets. It has functions for analyzing, cleaning, exploring, and manipulating data. The name "Pandas" has a reference to both "Panel Data", and "Python Data Analysis" and was created by Wes McKinney in 2008.
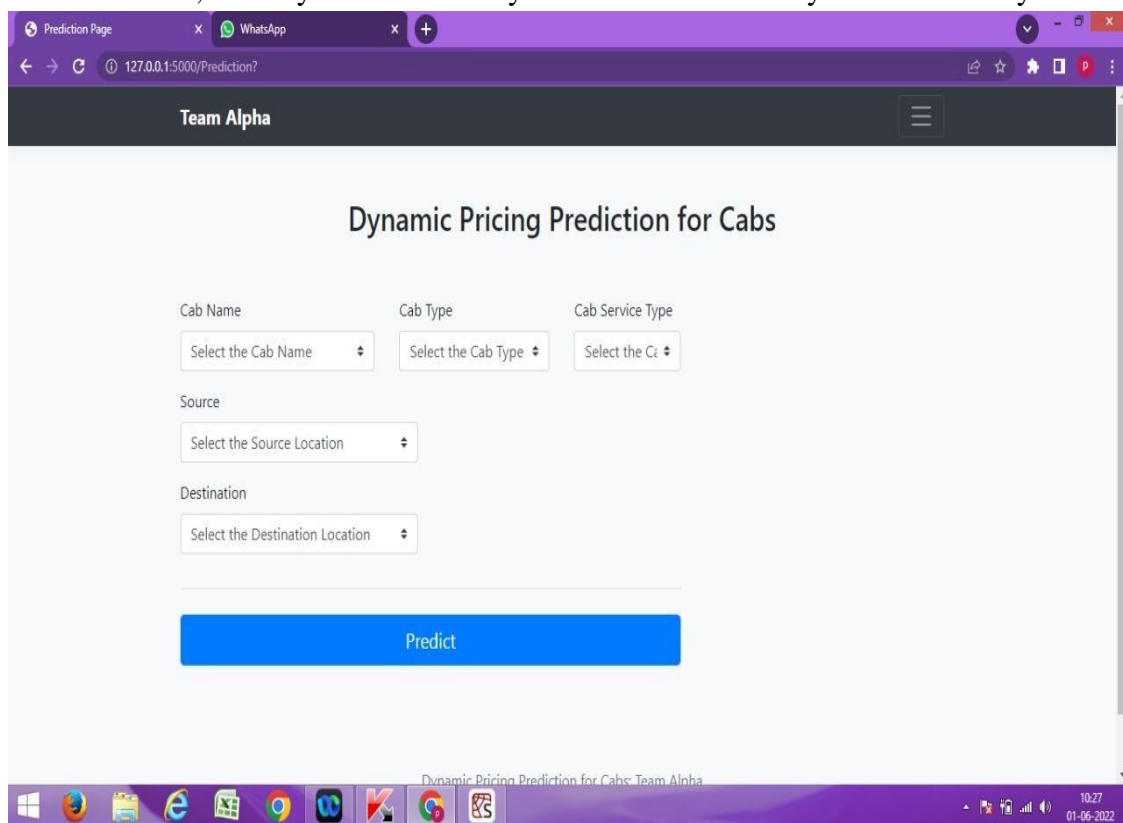

Fig.4 Output Of The Project

5.Conclusion

The Random Forest Regressor was successfully implemented as described in this project. For creating a predictive model, the Random Forest algorithm is excellent for both classification and regression problems. A popular supervised machine learning algorithm is Random Forest. Using various samples, it creates decision trees and uses the majority vote for classification and the average for regression.

Rapido is a Ride-On-Demand (ROD) service that tries to improve service quality by balancing supply and demand through dynamic pricing. The development of a model, metric calculation, and exploratory data analysis (EDA) are the primary focuses of the study. "ride fare" turned out to be the dependent algorithm variable, and "travel distance" and "travel time" are the independent variables. The efficiency of the Linear Regression Model for predicting the dynamic price of trips is 93.40 percent. It performs exceptionally well when used to predict trip fare, which is the target variable. Utilizing a variety of machine learning algorithms and methods, additional work like this one can be done.

108

References

1. Michael Deferred, Kirell Benzie, Pierre Vandergheynst, Xavier Bresson. FMA: A Dataset for Music Analysis. Sound; Information Retrieval. arXiv:1612.01840v3, 2017.

2. Tom LH Li, Antoni B Chan, and A Chun. Automatic musical pattern feature extraction using convolutional neural network. In Proc. Int. Conf. Data Mining and Applications, 2010.

3. https://smartinternz.com/guided-project/dynamic-pricing-prediction-for-cabs. https://youtu.be/M-xEvSEmDus

4. https://www.google.com

5. https://in.search.yahoo.com/?p=dynamic+pricing+prediction+for+cabs&fr2=sb-top&fr=mcafee&vm=r&type=E211IN826G0

6. https://in.search.yahoo.com/?p=dynamic+pricing+prediction+for+cabs&fr2=sb-top&fr=mcafee&vm=r&type=E211IN826G0

7. https://www.kaggle.com 7.https://www.altexsoft.com/blog/car-rental-reservation-system/ https://www.analyticsvidhya.com

8. https://github.com/akhil218/Dynamic-Pricing-Prediction-For-Cabs-Using-IBM-Watson/blob/master/cab_rides.csv