PRIOR KNOWLEDGE

Team ID	PNT2022TMID47865
Project Name	CAR RESALE VALUE PREDECTION

MACHINE LEARNING

Machine learning is a subfield of artificial intelligence (AI) and computer science that uses data and algorithms to replicate how humans learn, gradually improving its accuracy.

Machine learning is a vital component of the rapidly increasing study of data science. Algorithms are utilizing statistical methodologies to make classifications or predictions and to gain important insights in data mining operations. These insights then affect decision making within applications and companies, ideally influencing key growth indicators. As big data expands and grows, so will the market demand for data scientists. They will be required to assist in finding the most pertinent business questions and the data to answer them.

LEARNT OVER

- Supervised and unsupervised learning
- Regression Classification and Clustering
- Random Forest Regressor
- Flask

Has prior knowledge about the concepts

ALGORITHMS USED:

MULTI-LINEAR REGRESSION

Multiple linear regression is a statistical technique that predicts the outcome of a variable based on the values of two or more variables. It is an extension of linear regression and is sometimes referred to as simply multiple regression. The variable we want to predict is known as the dependent variable, and the variables we use to predict the value of the dependent variable are known as independent or explanatory variables.

SUPPORT VECTOR MACHINE

The algorithm represents the data points in three dimensions. With the help of hyperplanes, these data points are then classified. It creates an n-dimensional space for the n number of features in the dataset and then attempts to create hyperplanes that divide the data points with the greatest margin.

DECISION TREE REGRESSION

Decision trees may be used to forecast numerical values (regression) as well as categorize data. A decision tree is a branching series of related choices represented by a tree diagram. One of the benefits of decision trees over neural networks is that they are easier to evaluate and audit.

RANDOM FORESTS REGRESSOR

The machine learning algorithm predicts a value or category in a random forest by combining the results of multiple decision trees. The random forest algorithm is a bagging technique extension that generates an uncorrelated forest of decision trees by combining bagging and feature randomization. Feature randomization, also known as feature bagging or "the random subspace approach," ensures minimal correlation among decision trees by selecting features at random. This distinguishes decision trees from random forests. In contrast to decision trees, random forests only select a subset of the potential feature splits.

K-NEAREST NEIGHBORS

The k-nearest neighbours algorithm, also known as KNN or k-NN, is a non-parametric, supervised learning classifier that uses proximity to classify or predict the grouping of a single data point. While it can be used for either regression or classification problems, it is most commonly used as a classification algorithm, assuming that similar points can be found near one another.

XGBOOST

XGBoost, or Extreme Gradient Boost, is a machine learning technique used to create gradient boosting decision trees. When it comes to unstructured data, such as photos and unstructured text data, ANN models (Artificial neural network) appear to be at the top of the list when it comes to prediction. When it comes to structured/semi- structured data, decision trees are currently the best.

TECHNOLOGIES USED

PYTHON

Python is an agile, dynamically typed, expressive, open source programming language that supports multiple programming philosophies,including procedural, object-oriented, and functional.Python is a popular high-level programming language that is easily extensible through the use of third-party packages and often allows powerful function to be written with few lines of code.

HTML

HTML is a programming language used to organise web content. Its goal is to make online design and development easier by developing a standardised and user-friendly UI markup language. HTML allows you to deconstruct and compartmentalise your pages, as well as build separate components that are not only meant to arrange your site logically, but are also designed to provide your site syndication capabilities. HTML might be referred to as the "information mapping method to website design" since it integrates the concept of information mapping, separating and labelling data to make it easier to use and comprehend. This is the basis for HTML tremendous semantic and

graphical usefulness.

CSS

CSS is an abbreviation for Cascading Style Sheets. It is a style sheet language that is used to specify the look and formatting of a markup

language document. It adds a new functionality to HTML. It is commonly used in conjunction with HTML to alter the appearance of web pages and user interfaces. It may also be used with any type of XML document, such as plain XML, SVG, or XUL.

JAVASCRIPT

JavaScript (js) is a lightweight object-oriented programming language that is used to script webpages on many websites. When applied to an HTML document, it is an interpreted, full-fledged programming language that enables dynamic interactivity on website. Users may use JavaScript to create contemporary web applications that allow them to engage without having to reload the page every time. The typical website makes use of js to give various levels of interactivity and simplicity.

TOOLS USED

FLASK

Flask is a terrific and relatively light Python framework for constructing sophisticated online apps, as well as an amazing tool for generating dynamic and interactive webpages. You can construct fascinating applications using this strategy, even on versions before to
7.X. We can acquire outstanding results quickly using Flask and update management procedures with new apps.

IBM CLOUD

The term "IBM Cloud" refers to IBM's hardware, software, and services for assisting organisations in building private clouds, as well as its

Bluemix public cloud services. The moniker "Bluemix" was formerly reserved for IBM's PaaS services for developers, however Bluemix now now provides certain IaaS services.