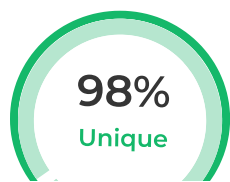
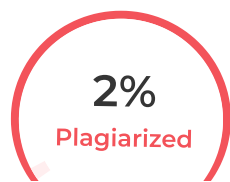


Plagiarism Scan Report



Characters:6513

Words:956

Sentences:44

Speak Time:
8 Min

Excluded URL

None

Content Checked for Plagiarism

1. Project Title Recipe Recommendation Using Machine Learning 2.

Description of the project (Aim/Objective) The objective of this project is to develop a predictive model that, by using different inputs such as past sales data , promotions and holidays , is able to quantitatively model the retail store's performance. Our main task is to create an effective model which can be used to forecast the sales volatility analysis by means of machine learning techniques. With this insight, managers at shopping centers will better informed when it comes to team size, marketing strategies and inventory placements. Apart from that the study is also coming up with an assessment and comparison of different machine learning algorithms in order to identify the success algorithm for the sales prediction of retail setting in this specific situation. The desired result would be to put data based decision making in place to attain profitability and have the operations operating at their peak. 3.

Dataset Description · 20k recipe which are rated, provided the information of the nutritional elements and organized by categories · Recipes extracted from <https://www.kaggle.com/datasets/nehaprabhavalkar/indian-food-101> 4.

Statistical Analysis on the Dataset (EDA) 5. Machine Learning

Implementation a. Pre-processing (e.g: handling null values, handling categorical data etc) b. Apply ML Models (4 Machine Learning models on the same dataset) c. Calculate Evaluation metrics for each ML Model d. Visualize the results 6. Answer all the CO's: CO1, CO2, CO3, CO4, CO5 CO1: Understand different machine learning techniques and its applications with respect to your project/case study problem statement. A simple way to model the relationship between a dependent variable and independent variables is through linear regression. This method assumes that the predictors and the target variable have a linear connection. Linear regression is commonly used for prediction and inference tasks to understand how changes in the independent variables affect the target variable. **K-Nearest Neighbors (KNN)** is a non-parametric technique used for regression and classification tasks. It works by finding the 'k' nearest data points to a specified query point and making predictions based on the average of the 'k' neighbors (for regression) or the majority class (for classification). Interpretability: The link between the predictor and target variables is evident and easy to understand when using simple linear regression. Higher prediction accuracy may be possible with more intricate models, such as Random Forest, however these models frequently lack the openness and interpretability of linear regression. Because

of this, linear regression may be used to determine how certain features in your dataset directly affect the target variable. Baseline Model: For the purpose of evaluating the effectiveness of more complex algorithms such as Random Forest or KNN, a baseline model such as simple linear regression can be used. You may assess whether the extra complexity of these models results in appreciable increases in predict CO2: Understand the importance of simple linear regression in predicting new observations with respect to your project/case study. Can Simple Linear Regression can be used for your problem? Yes/No Explain why? Finding Linear connections: The predictor variables and the target variable can be found to have linear connections thanks to linear regression. Basic linear regression may reveal which factors have a mostly linear impact on the goal, even if your dataset contains non-linear correlations. This can assist direct engineering efforts and feature selection, enhancing the performance of more complicated models. Efficiency and Speed: When compared to ensemble techniques like Random Forest, linear regression is more computationally efficient and can be taught more rapidly. Because of this, it may be used with big datasets or in scenarios with constrained processing resources. Furthermore, real-time or resource-constrained applications might benefit from the practicality of linear regression models due to their lightweight and ease of deployment CO3: Understand the importance of Multiple linear regression in predicting new observations with respect to your project/case study. Can Multiple Linear Regression can be used for your problem? Yes/No Explain why? When there are several independent factors impacting a single dependent variable, as is frequently the case in predictive modeling scenarios, multiple linear regression is a good option. many Linear Regression can enhance existing approaches in your project, such as Random Forest, Decision Trees, Linear Regression, and KNN, by managing many characteristics at once. It helps to comprehend the influence of each independent variable on the target variable by providing comprehensible coefficients for each one. Additionally, it enables you to verify the validity of your models by testing hypotheses like linearity and homoscedasticity. By comparing the model to current methods, Integrating Multiple Linear Regression helps you choose the best strategy for your particular situation. CO4: For your project definition, explain the importance of parameter estimation. Check how it affects the overall results when you change the values of different parameters. parameter estimate is essential as it has a direct impact on your machine learning models' accuracy and performance. Parameters like learning rates, regularization parameters, tree depths, and the number of neighbors in a KNN may all have a big influence on how well the model predicts and generalizes. For example, the convergence speed and overfitting risk of algorithms such as Gradient Boosting Machines or Neural Networks may be tuned by varying parameters like learning rates. Similar to this, changing tree-based models such as Random Forest or Decision Trees may have an impact on how complicated the model is and how well it can identify complex patterns in the data. These parameters include tree depth and the number of trees. Through methodical investigation and refinement of these factors, you may optimize your models

to attain the greatest CO5: Compare different Machine Learning algorithms used for your project/case through evaluation metrics. Show which ML algorithm works best for your problem. • Comparison table showing the results of different ML algorithms.

Sources

2% Plagiarized

K-Nearest Neighbors (KNN) is a non-parametric machine learning algorithm that can be used for both classification and regression tasks.

<https://medium.com/@nandiniverma78988/understanding-k-nearest-neighbors-knn-regression-in-machine-learning-c751a7cf516c>



Home Blog Testimonials About Us Privacy Policy

Copyright © 2024 [Plagiarism Detector](#). All right reserved