■ Sales Prediction Project Report

Introduction

This project focuses on predicting product sales based on advertising expenditures across multiple channels. The dataset provided contains information about advertising budgets for TV, Radio, and Newspaper, along with corresponding sales values. The goal is to build a machine learning model that accurately predicts sales based on these inputs.

Dataset Overview

The dataset consists of 200 rows and 4 columns: TV, Radio, Newspaper, and Sales. Each row represents a specific marketing campaign, with expenditures recorded for each advertising channel and the resulting product sales. TV and Radio expenditures show strong correlations with sales, while Newspaper has a weaker influence.

Methodology

1. Data Preprocessing: The dataset was cleaned and prepared for training. No missing values were present, and the features were used directly. 2. Model Selection: A Linear Regression model was chosen due to the continuous nature of the target variable. This model estimates the relationship between advertising spend and sales. 3. Model Training: The model was trained using all three predictors (TV, Radio, Newspaper). The dataset was split into training and testing sets to evaluate performance. 4. Model Evaluation: Performance was assessed using metrics such as Mean Squared Error (MSE) and R^2 Score.

Results

The Linear Regression model achieved a high R^2 score, indicating that it explains a significant portion of the variance in sales. The coefficients of the model suggest that TV advertising has the strongest impact on sales, followed by Radio. Newspaper expenditure contributes very little to the prediction.

Insights

- TV advertising is the most influential channel for increasing sales. - Radio advertising also contributes positively, though less than TV. - Newspaper advertising shows minimal impact and may not be cost-effective. - Combining TV and Radio budgets can lead to optimized sales outcomes.

Conclusion

The project successfully demonstrates how machine learning can be applied to predict sales based on advertising spend. The Linear Regression model provides interpretable results, making it a valuable tool for marketing strategy decisions. Future improvements could include testing advanced models such as Random Forests or Gradient Boosting to compare performance.