



SALES FORECAST AND PREDICTIONS



GITHUB REPOSITOR
LINK



ABOUT US

PRESENTED BY: ANMOL BANSAL
ENROLLMENT NUMBER: E23CSEU0431

BATCH: 15

GROUP: 04

SUBJECT: STATICAL MACHINE LEARNING
SUPERVISIOM: DR. SUSMITA DAS (ASSISTANT PROF.)

This document provides a detailed explanation of the sales forecasting. The code utilizes historical sales data to create a time series forecasting model. It preprocesses the data, performs feature engineering, and uses machine learning algorithms to predict future sales based on past trends.



INTRODUCTION TO SALES FORECASTING

PREDICTING FUTURE SALES AND REVENUE FOR FUTURE PLANNING



01

Understanding Sales Forecasting

SALES FORECASTING PREDICTS FUTURE SALES BASED ON PAST DATA AND TRENDS.



02

Importance of Accurate Forecasting

ACCURATE FORECASTS ENABLE EFFECTIVE INVENTORY MANAGEMENT AND RESOURCE ALLOCATION.



03

Enhancing Strategic Planning

FORECASTING AIDS IN STRATEGIC PLANNING AND GAIN COMPETITIVE ADVANTAGE.



04

Data-Driven Decision Making

UTILIZING ANALYTICAL TECHNIQUES HELPS IN MAKING INFORMED DECISIONS.



05

Market Trends Analysis

ANALYZING MARKET TRENDS IS CRUCIAL FOR ACCURATE SALES PREDICTIONS.



OUR PROJECT

The script is a practical implementation of machine learning techniques for time series forecasting, specifically designed for sales data. Predicting future sales is a critical task in business operations, as it aids in inventory management, resource allocation, and strategic planning. This script takes historical monthly sales data, preprocesses it into a supervised learning format, and applies regression modeling to predict future sales trends.

The script integrates core data science workflows, including feature engineering, data scaling, and model evaluation, while leveraging Python libraries like pandas, scikit-learn, and matplotlib. By focusing on simplicity and interpretability, sales.py serves as a foundational tool for businesses or students exploring sales forecasting.



Linear Regression in Sales Forecasting

Using statistical methods to predict sales based on influencing factors

Statistical Method

Linear regression is a statistical method for modeling relationships.



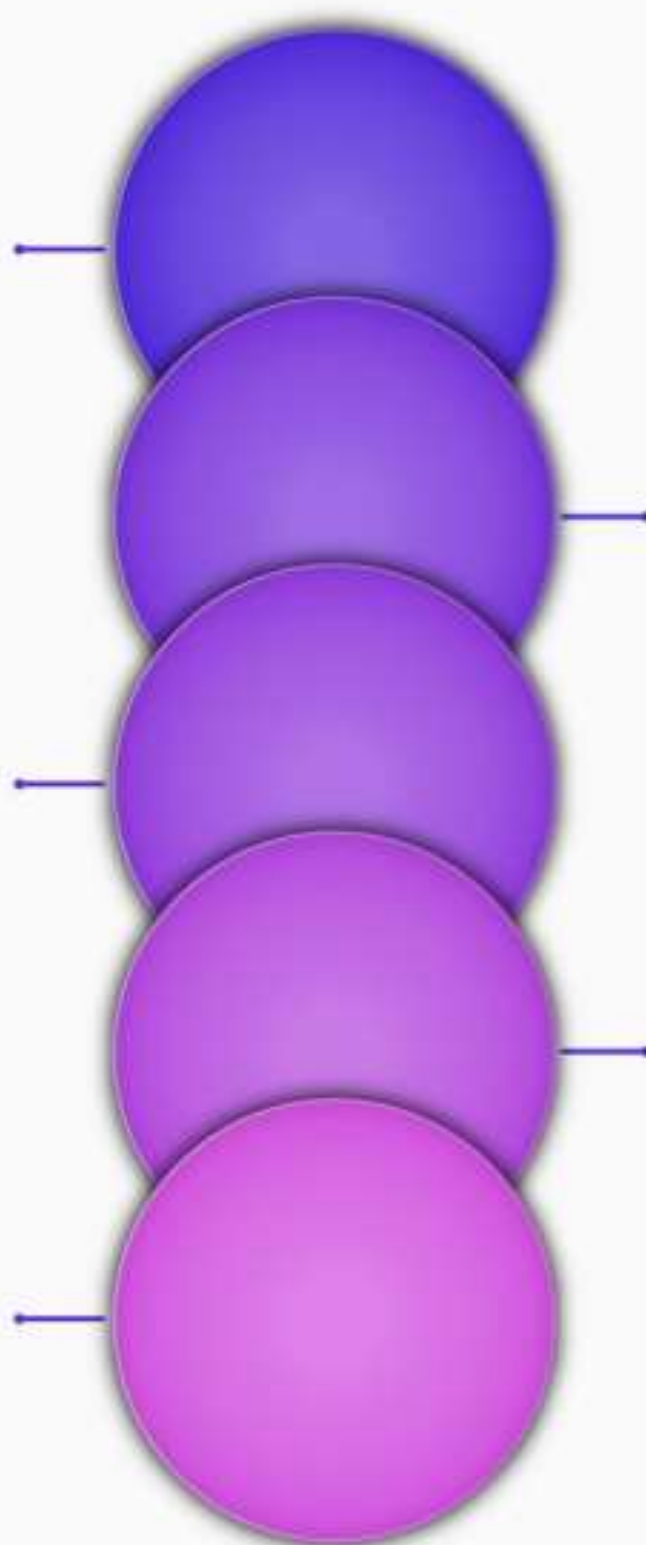
Sales Forecasting

This method is crucial in sales forecasting to predict future sales efficiently.



Influencing Factors

Key influencing factors include price adjustments and advertising efforts.



Dependent and Independent Variables

It involves a dependent variable (sales) and independent variables (factors like price).



Linear Relationship

It establishes a linear relationship, helping to understand how different factors influence sales.





CODE IMPLEMENTATION USING RANDOM FOREST REGRESSOR

IMPLEMENTING A ROBUST MODEL FOR SALES FORECASTING IN PYTHON.

IMPORTING LIBRARIES

BEGIN BY IMPORTING NECESSARY LIBRARIES SUCH AS SCIKIT LEARN TO ACCESS MACHINE LEARNING FUNCTIONS.

DATA PREPROCESSING

PREPARE YOUR DATASET BY HANDLING MISSING VALUES AND SPLITTING IT INTO TRAINING AND TESTING SETS.

MODEL TRAINING

TRAIN THE RANDOM FOREST MODEL ON THE TRAINING DATASET TO LEARN PATTERNS AND RELATIONSHIPS.

MODEL EVALUATION

EVALUATE THE MODEL'S PERFORMANCE USING RMSE AND R-SQUARED METRICS TO ASSESS ACCURACY.

HYPERPARAMETER OPTIMIZATION.

USE GRIDSEARCHCV TO FINE-TUNE THE MODEL'S HYPERPARAMETER FOR POTENTIAL PERFORMANCE.



BACKGROUND & MOTIVATION

Sales forecasting is a critical function for any organization aiming to remain competitive in a data-driven economy.

With the growing availability of historical data and the advent of powerful machine learning (ML) tools, the motivation for developing an ML-based sales forecasting project is rooted in addressing traditional challenges and leveraging modern advancements.





RELATED WORK

SALES FORECASTING USING SARIMAX FOR B2C:

BY KR SENTHIL MURUGAN, R PRATIBHA SELVI, S
DEEKSHA, R DEEPALAKSHMI, B PAUL SETEVE MITHUN

COMPARISON: LOGISTIC REGRESSION, SVM, AND
RANDOM FOREST APPLIED IN PREVIOUS STUDIES
ON WINE QUALITY





SYSTEM REQUIREMENTS



HARDWARE

BASIC COMPUTING DEVICE
(LAPTOP/DESKTOP)



SOFTWARE

FLASK: WEB APP DEVELOPMENT
GOOGLE COLAB: MODEL
DEVELOPMENT



PYTHON LIBRARIES

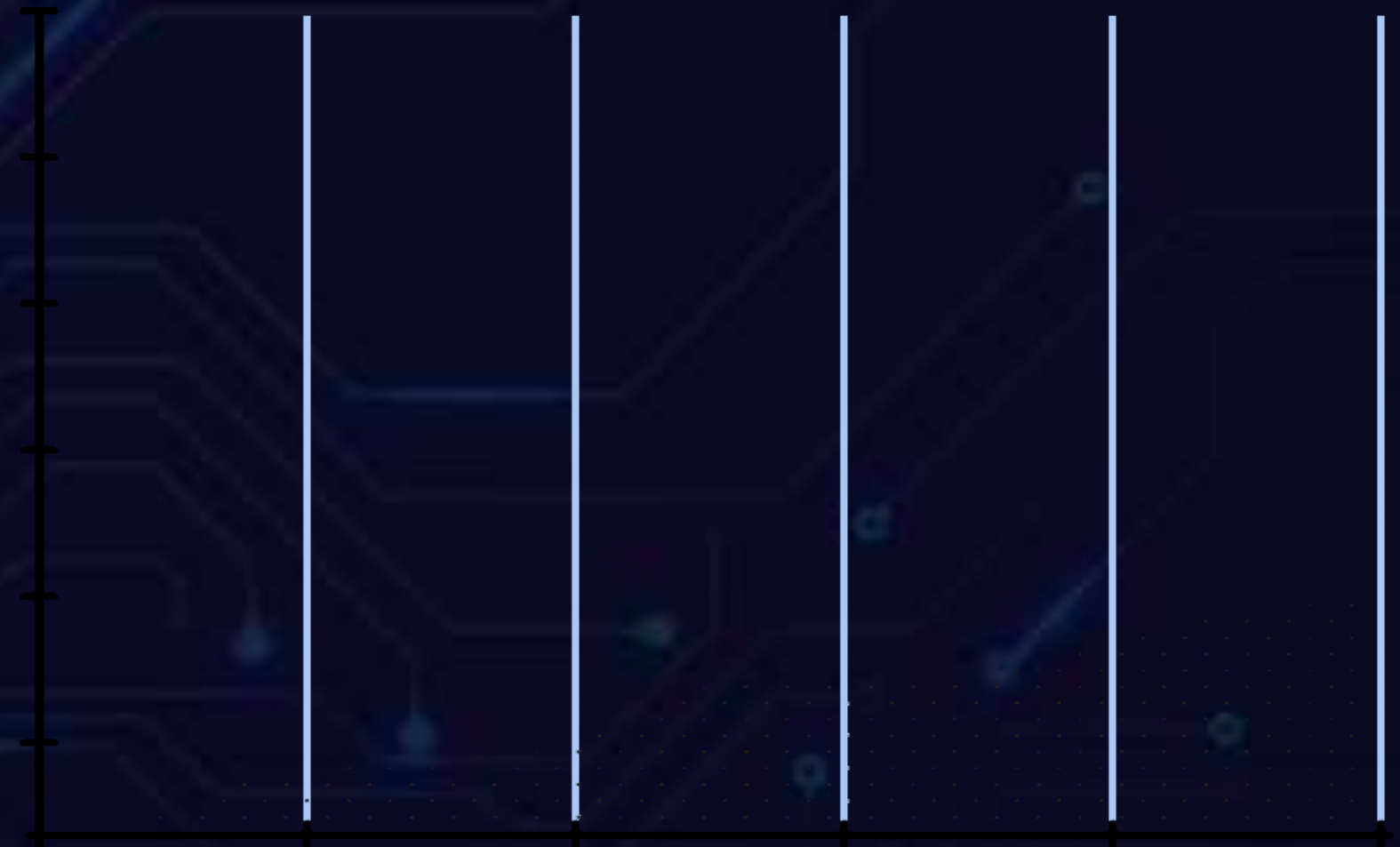
PANDAS, NUMPY,
SKLEARN, MATPLOTLIB.



EXPERIMENTAL RESULTS

MODEL EVALUATION: ONCE PREDICTIONS ARE GENERATED, THEY ARE EVALUATED AGAINST THE ACTUAL SALES DATA USING SEVERAL METRICS: RMSE, MAE AND R-SQUARED.

VISUALIZATION: A LINE PLOT IS CREATED, WITH ACTUAL SALES REPRESENTED AS A CONTINUOUS LINE AND PREDICTIONS AS A DASHED LINE, CLEARLY SHOWING ANY DISCREPANCIES.

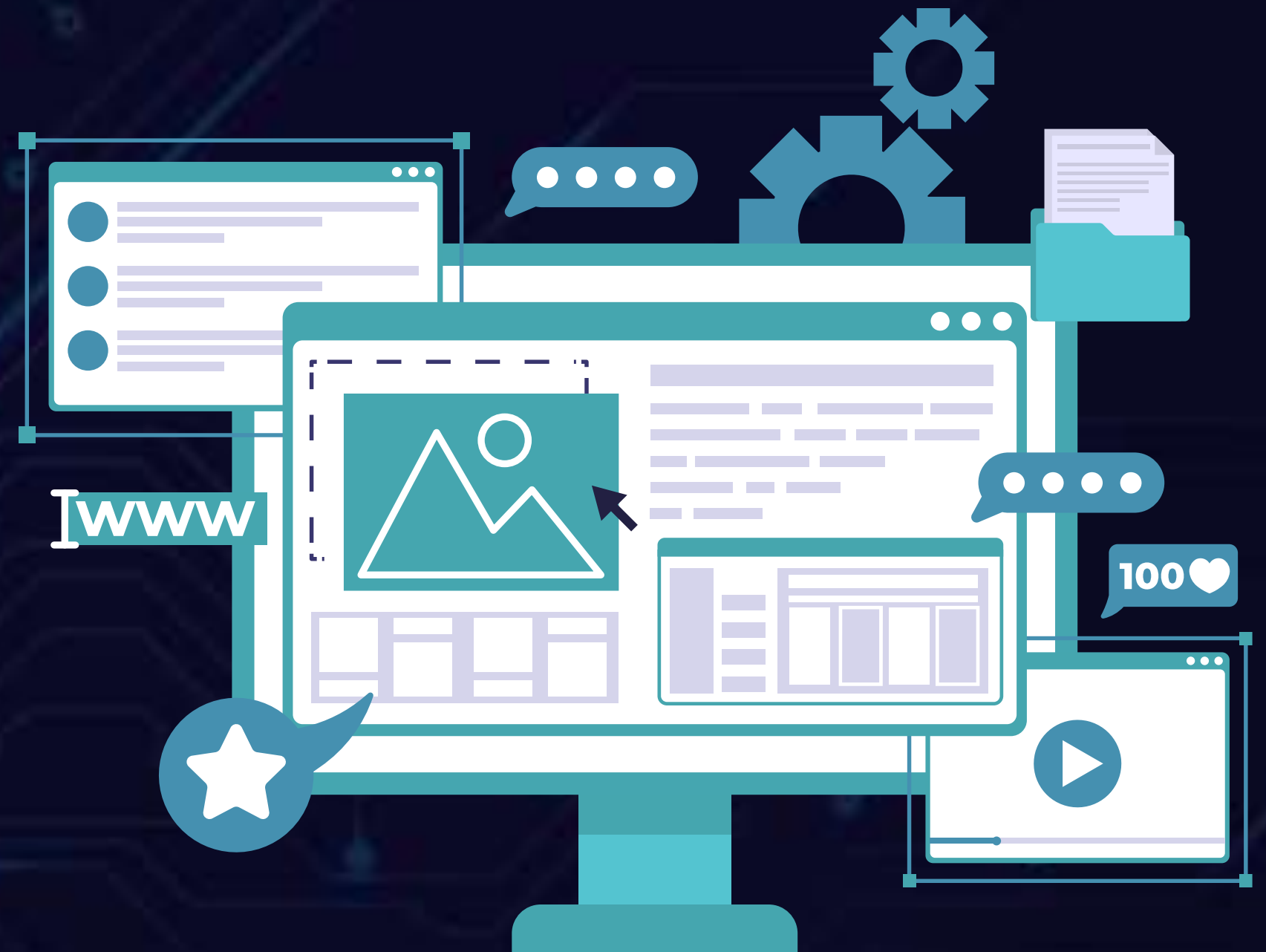




FLASK APPLICATION

OVERVIEW: WEB INTERFACE TO TAKE INPUT OF DATASETS TO
PREDICT FUTURE SALES.

FEATURES: USER-FRIENDLY FORM TO UPLOAD SALES
DATASET. BACK-END MODEL INTEGRATION WITH REAL-TIME
PREDICTIONS.





CONCLUSION

THE SALES FORECASTING PROVIDES A STRUCTURED APPROACH TO PREDICTING FUTURE SALES BASED ON PAST DATA.

IT DEMONSTRATES THE IMPORTANCE OF FEATURE ENGINEERING, DATA SCALING, AND MODEL SELECTION IN BUILDING AN EFFECTIVE TIME SERIES MODEL. THE CODE CAN BE FURTHER ENHANCED BY EXPERIMENTING WITH DIFFERENT MODELS OR OPTIMIZING MODEL PARAMETERS TO IMPROVE FORECASTING ACCURACY.





KEY TAKEAWAYS & FUTURE WORK

STRENGTH LR: LINEAR REGRESSION IS MOST EFFECTIVE FOR IDENTIFYING ADN
MODELING LINEAR RELATIONSHIPS BETWEEN VARIABLES.

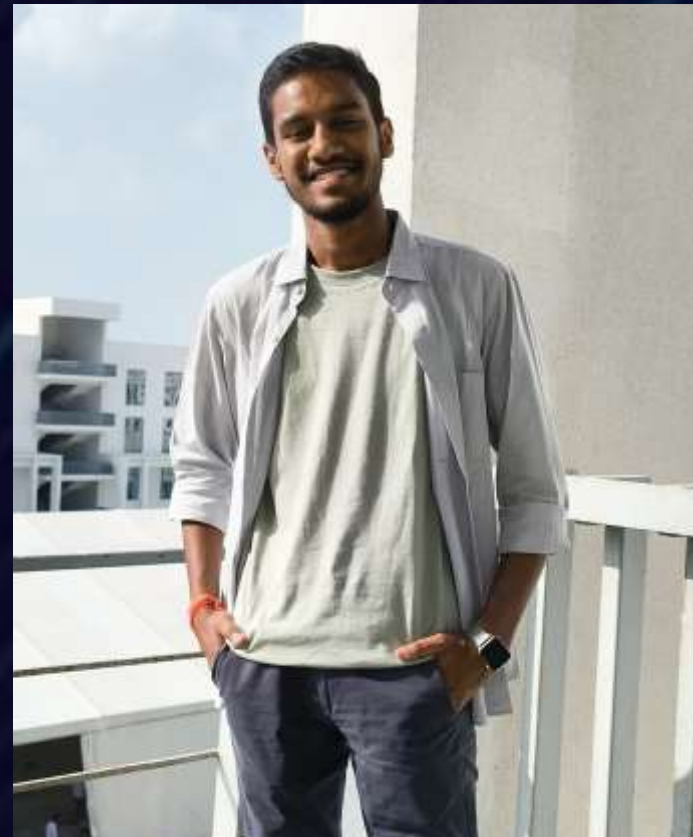
EXPLORING HYBRID MODELS: FUTURE RESEARCH SHOULD FOCUS ON HYBRID
MODELS THAT COMBINE STRENGTHS OF BOTH LINEAR REGRESSION AND
RANDOM FOREST FOR BETTER ACCURACY.

INTEGRATING DATA SOURCES: INCORPORATING DIVERSE DATA SOURCES CAN
SIGNIFICANTLY ENHANCE PREDICTION ACCURACY AND MODEL ROBUSTNESS.





THIS IS ME



ANMOL BANSAL

BTech 2nd year Student
Specialization: Artificial
Intelligence



THANK YOU

GITHUB REPO. LINK

