Final year project

Roll # B21120206007

Financial Computer Simulation (FM 617)

BS Financial Mathematics (4th Year)

Submitted to

Sir Daniyal Hussain

Submitted by:

Aqsa Ali

CODING METHODOLOGY

- Data Generation & Acquisition
- Preprocessing Steps
- Modeling Techniques
- Evaluation Metrics
- Visualization

DATA GENERATION & ACQUISITION

- **Stock Prices**

  with downtrend and high volatility.

- **Credit Risk**

  with income, loan, credit score, and employment type.

- **Revenue/Expense**

  data with growth trends and seasonality.

PREPROCESSING STEPS

Handled missing values using:

- Forward-fill and backfill for time series.
- Median/mode imputation for tabular data.
- Encoded categorical variables using Label Encoder and one-hot encoding.

Created new features:

- Moving averages (SMA_10)
- Log returns and rolling volatility (for stock data)
- Previous values for regression.

Modeling Techniques:

Stock Forecasting

Linear Regression
ARIMA for time series

Credit Risk Modeling:

Logistic Regression
Decision Tree Classifier

Revenue/Expense Forecasting:

Linear Regression
Random Forest Regressor
Gradient Boosting Regressor

Stochastic Processes & Derivatives:

Simulated using
Geometric Brownian Motion
Option pricing with Black-Scholes and Binomial Tree

Evaluation Metrics

Regression:

- RMSE, MAE, R^2 Score

Classification:

- Accuracy, Precision, Recall, F1-score, Confusion Matrix

Visualization

- Plots for actual vs. predicted values
- Confusion matrices for classification
- Heat maps, histograms, time series line plots

FINANCIAL INTERPRETATION

- Stock Price Models
- Credit Risk Models
- Revenue/Expense Forecasting
- Derivatives Pricing

STOCK PRICE MODELS

Linear Regression:

- had moderate predictive ability

RMSE = 2.29, MAE = 1.48, R^2 = 0.37

ARIMA :was less accurate on test data

RMSE = 3.13, R^2 = -0.18 (indicating poor generalization)

Interpretation:

The stock exhibits high volatility; linear regression captures some patterns, but ARIMA's negative R^2 suggests it fails to generalize well.

CREDIT RISK MODELS

Logistic Regression

Precision = 1.0 (all predicted positives were correct) Recall = 0.06 (missed most of the actual positives)

Decision Tree:

Balanced performance: Accuracy = 56.67%, F1 = 0.6061💳

Interpretation:

Logistic regression is overly conservative. The decision tree offers a more balanced assessment of risk.

REVENUE/EXPENSE FORECASTING:

Best model for revenue:

Linear Regression ($R^2 = 0.54$)

Best model for expense:

Linear Regression ($R^2 = 0.44$)

Interpretation:

Revenue and expenses show seasonal and linear growth trends, best captured by linear regression over tree-based models.

DERIVATIVES PRICING:

Call Option

(S=100, K=100, T=1, σ=0.25, r=0.02)

Binomial Tree ≈ \$10.85

Black-Scholes ≈ \$10.87

Put Option: \~\$8.87

Interpretation:

Binomial and Black-Scholes models yield consistent option prices, validating the stochastic model.

Result Submission Summary

Module	Model	Key Metric
Stock Forecasting	Linear Regression	RMSE: 2.29, R^2: 0.37
Credit Risk Modeling	Decision Tree	Accuracy: 56.67%, F1: 0.61
Revenue Forecasting	Linear Regression	R^2: 0.54
Expense Forecasting	Linear Regression	R^2: 0.44
Option Pricing	Black-Scholes / Binomial	Call: 10.87, *Put* : 8.89