

PROFIT PREDICTION USING LINEAR REGRESSION – PROJECT REPORT

1. INTRODUCTION

This project focuses on predicting company profit using important financial indicators such as Administration Spend, R&D; + Marketing Spend and State representation through dummy encoding.

The aim is to identify which parameter drives profit growth the most and to evaluate the efficiency of a Linear Regression model in prediction accuracy.

2. DATASET OVERVIEW

Total Observations: ~993 records

Target Variable: Profit

Features Used:

- Administration_scaled
- R&D;_Marketing_Spend_scaled
- State_California
- State_Florida
- State_New York

Scaling ensures fairness across numerical features, and dummy encoding converts state into binary machine-understandable form.

3. MODEL DEVELOPMENT

Learning Algorithm: Linear Regression

Train–Test Split: 80% Training, 20% Testing

Model trained using OLS (Ordinary Least Squares) Regression.

4. MODEL EVALUATION METRICS

- R² Score measures how much variability in profit is explained by the model.
- RMSE represents average prediction error between actual and predicted outcomes.

5. FEATURE IMPORTANCE

Based on regression coefficient magnitude:

1. R&D;_Marketing_Spend_scaled ↗ Most impactful
2. Administration_scaled ↗ Moderate effect
3. State variables ↘ Least influence on profit

Conclusion: Investment in R&D; + Marketing drives profit significantly higher than administration spend.

6. PREDICTION ANALYSIS

Highest Predicted Profit ₹ 341057.17

Actual Profit for same instance ₹

333962.19

Very small deviation ↗ High confidence and model reliability.

7. VISUAL ANALYSIS SUMMARY

- Actual vs Predicted scatter plot aligns closely = Good prediction accuracy.
- Residual plot displays random spread = No visible pattern ↗ No model bias.
- Feature bar chart proves R&D;+Marketing dominant in driving profit.

8. FINAL CONCLUSION

This model confirms that:

- ✓ Increased R&D; + Marketing spend directly increases profit.
- ✓ Administration cost supports but impacts marginally.
- ✓ Profit is independent of company location state.

Therefore, to maximize scalability and revenue, businesses must allocate higher share towards innovation-driven R&D; and marketing uplift.

