Mini Project Proposal

ICT 4306 – Data Science

Bachelor of Information and Communication Technology (BICT)

Degree Programme

Department of Information and Communication Technology

Faculty of Technology

Rajarata University of Sri Lanka

Mihintale

Details of the Project

Project Title : Profit Prediction with Machine Learning

Group Name : Team One and Zero

Submission Date : 2024/03/04

Group Members :

Name	Index Number	Registration Number
K.S.T. Dilshani	0987	ITT/18/19/030
H.R.D.R.M.Chandrasena	0975	ITT/18/19/018

Contents

1. Introduction

Predicting profits is an important pursuing that helps businesses set achievable targets. Businesses may make intelligent choices and work toward success by accurately determining the profit they can make based on their efforts in marketing, R&D, and other areas. In this project, Python programming and machine learning techniques will be used to examine the topic of profit prediction.

The goal of profit prediction is to create a machine learning model capable of accurately predicting future profit by analyzing previous data. The approach helps businesses set achievable and attainable profit targets by taking into account variables including R&D spend, administrative costs, marketing spend, and the stage of operation. The goal of this project is to develop a profit prediction model using Python programming and machine learning algorithms.

The model is trained using a linear regression approach on the training set of data. By projecting profit values for the testing data, the model's performance is measured. The model's the accuracy is evaluated and the results are examined.

Features like R&D spend, administration costs, marketing spend, status of business, and the associated historical profit made by each firm are all included in the dataset used for this study. There are 50 samples in this data set, and there are no missing values. To ascertain the correlations between the features, correlation analysis and descriptive statistics are used.

In conclusion, machine learning-based profit prediction offers useful information to help organizations establish realistic objectives. Businesses are able to calculate their profit accurately by utilizing machine learning algorithms and Python, taking into account a variety of criteria. The issue definition, aims, technique, data needs, and references for more research on machine learning-based profit prediction will all be covered in this study.

2. Problem statement

Predicting a company's profit for a specific time period based on a variety of variables, including R&D expenditures, administrative expenses, marketing expenditures, and operational status, is the current challenge. Creating a machine learning model that can correctly forecast the profit based on these characteristics is the aim.

3. Objectives

- To train a machine learning model to predict the profit of a company based on historical data.
- To set achievable goals for the company by leveraging the profit prediction model.
- To utilize Python programming language and relevant libraries for implementing the profit prediction model.

4. Methodology

The following steps are part of the machine learning profit prediction methodology:

- Importing the required Python libraries, including seaborn, matplotlib, pandas, and numpy.
- Loading the dataset with historical earnings information from fifty startups.
- Analyzing data exploratory can help you understand it better.
- Dividing the data into training and testing sets after choosing the pertinent characteristics.
- Utilizing the training data to construct a linear regression model.
- Estimating profit values for the testing data in order to assess the model's performance.
- Examining the data and evaluating the model's correctness.

5. Data requirements (description of available data)

The dataset comprises five variables, namely R&D Spend, Administration, Marketing Spend, State, and Profit, with each row representing a particular organization. The dataset appears to be associated with financial information on businesses, specifically with regard to their expenses and earnings across various states (New York, California, and Florida).

6. References

[1]
A. Kharwal, "180 Data Science and Machine Learning Projects with Python," *Coders Camp*, Mar. 18, 2022. https://medium.com/coders-camp/180-data-science-and-machine-learning-projects-with-python-6191bc7b9db9

[2]
A. P. Gulati, "Startups profit prediction using Multiple Linear Regression," *Analytics Vidhya*, Nov. 03, 2021. https://www.analyticsvidhya.com/blog/2021/11/startups-profit-prediction-using-multiple-linear-regression/

[3]
"Profit Prediction with Machine Learning," *Bing*. https://www.bing.com/search?pglt=41&q=Profit+Prediction+with+Machine+Learning&cvid=be86392c6d064fe5abe64b9f070707fc&gs_lcrp=EgZjaHJvbWUyBggAEEUYOTIGCAEQABhAMgYIAhAAGEAyBggDEAAYQDIGCAQQABhAMgYIBRAAGEAyBggGEAAYQDIGCAcQABhAMgYICBBFGDzSAQg0NjQ0ajBqMagCALACAA&FORM=ANNTA1&PC=U531.