Lab 1 Report

Aaditi
Department of Computer Science & Engineering
Amrita School of Computing Bengaluru
Amrita Vishwa Vidyapeetham, India
BL.EN.U4CSE21001

Ananya Avvaru
Department of Computer Science & Engineering
Amrita School of Computing Bengaluru
Amrita Vishwa Vidyapeetham, India
BL.EN.U4CSE21015

Aryan Tandon
Department of Computer Science & Engineering
Amrita School of Computing Bengaluru
Amrita Vishwa Vidyapeetham, India
BL.EN.U4CSE21031

***Abstract*— this report explores the significance of matrix rank in model development for classification tasks. It also delves into regression and classification processes, highlighting their distinctions. Additionally, the report analyzes stock data with the aim of proposing a predictive system for future price and percentage change.**

I. Introduction

In the realm of machine learning and data analysis, matrices and linear algebra play a pivotal role. Linear algebra serves as a foundational toolkit for deciphering intricate datasets using matrix operations and vector spaces. This report demonstrates fundamental operations utilized in data analysis for machine learning in real-world scenarios.

A. Importance of Matrix Rank in Model Building

Matrix rank holds key importance when constructing models for classification tasks. The rank of an observation matrix reflects the dimensionality of the data space it represents. In the context of classification, a low-rank matrix might indicate that the data lies predominantly along a lower-dimensional subspace, potentially leading to overfitting. On the other hand, a full-rank matrix better captures the inherent complexity of the data, allowing for more robust model building.

B. Regression and Classification

1) Regression: Regression tasks involve predicting continuous numerical values based on input features. For instance, predicting the price of a stock given various influencing factors.

2) Classification: Classification tasks entail assigning data points to predefined categories or classes. For example, categorizing stocks as 'buy,' 'hold,' or 'sell' based on certain features. The main distinction lies in the nature of the output: regression predicts values on a continuous spectrum, while classification assigns discrete labels.

II. Differentiating Regression and Classification

Regression and classification are two fundamental concepts in machine learning and statistical analysis. They involve different types of tasks and address distinct prediction scenarios.

A. Regression

Regression is a supervised learning task that involves predicting a continuous output or numerical value. It aims to establish a relationship between input features and a target variable, where the target is a real-valued quantity.

The output of a regression model is a continuous range of values. It can include decimal numbers and fractions. Some examples of this are Predicting stock prices, house prices, temperatures, etc.

Regression results are often visualized using scatter plots, where the predicted values are plotted against the actual values.

B. Classification

Classification is also a supervised learning task that involves categorizing data points into predefined classes or categories.

The output of a classification model is a discrete label representing a class or category. Some examples of this are categorizing emails as spam or not spam, classifying images of animals into different species, sentiment analysis (positive, negative, neutral), etc.

Common evaluation metrics for classification include accuracy, precision, recall, F1-score, confusion matrix, etc. Classification results are often visualized using confusion matrices and ROC curves.

III. Predictive System for Price and Change %

Analyzing the provided stock data, we propose building a predictive system to forecast stock prices and percentage changes. This system would likely incorporate historical stock prices, relevant financial indicators, and possibly sentiment analysis of news related to the company. Machine learning algorithms, such as time series forecasting or regression models, could be employed to capture trends and patterns. Regular updates and refining the model with new data would be essential for maintaining accuracy and adaptability.

IV. Conclusion

This report highlights the significance of matrix rank in classification model building and distinguishes between regression and classification tasks. Furthermore, it outlines the potential approach to develop a predictive system for stock price and percentage change using machine learning techniques. Successfully implementing such a system could provide valuable insights for stock price prediction.

ACKNOWLEDGMENTS

I am deeply grateful to our professor Dr. Peeta Basa Pati for his exceptional guidance and expertise.

REFERENCE

Tom Mitchell. Machine Learning. McGraw Hill; 2017
Richard O. Duda, Peter E. Hart, David G. Stork Pattern Classification. Wiley; Second Edition: 2007