



Graphic Era HILL UNIVERSITY

Established by an Act of the State Legislature of Uttarakhand (Adhiniyam Sankhya 12 of 2011)
University under section 2(f) of UGC Act, 1956

Project Report

On

BTC/USDT Trading Model Using Ridge Regression Model

Submitted by: -

PRINCE PANWAR (Team Leader)

UJJAWAL CHOUDHARY

SHIVANGI SINGH

DIVANSHI SHARMA

Under Mentorship of : Ms. Poonam Verma (M.Tech)

Index

1. Introduction

1.1 Problem Statement

1.2 Objective

2. Specific Objectives

2.1 Data Preprocessing

- 2.1a Data acquisition and Preprocessing data
- 2.1b Transforming 'Date' column
- 2.1c Handling missing values

2.2 Model Development

- 2.2a Using Ridge Regression model and Feature Analysis
- 2.2b Utilizing the attributes of Historical dataset obtained

2.3 Training and Testing

- 2.3a Split dataset into training and testing sets
- 2.3b Applying advanced imputation techniques
- 2.3c Standardize feature scaling

2.4 Model Evaluation

- 2.4a Assess Ridge Regression model
- 2.4b Assessing MSE and R² of the model

2.5 Drawdown Analysis

- 2.5a Perform meticulous drawdown analysis

2.6 Risk-Return Analysis

- 2.6a Calculate Sharpe Ratio
- 2.6b Conduct risk-reward analysis

2.7 Trading Strategy Evaluation

- 2.7a Design and implement a trading strategy
- 2.7b Evaluate strategy's performance

3. Model Development and Training

3.1 Data Acquisition

3.2 Data Preprocessing

3.3 Model Architecture

3.4 Training Process

4. Model Evaluation

4.1 Ridge Regression Model

4.2 Drawdown Analysis

4.3 Risk-Return Analysis

5. Trading Strategy Evaluation

5.1 Trading Strategy Implementation

5.2 Net Profit Exceeding Return

5.3 Risk-Reward Ratio

5.4 Maximum Duration Time of a Single Trade

6. Conclusion

7. Future Work

1. Introduction

1.1 Problem Statement

The problem consists of developing algorithmic trading models for the BTC/USDT cryptocurrency market, aiming to outperform benchmark returns.

1.2 Objective

Develop a robust predictive trading model for BTC/USDT prices, with the primary goal of enhancing decision-making processes in cryptocurrency trading to fetch higher returns. Model aims to give in depth analysis in BTC/USDT trade using historical data on BTC/USDT trade from 01-01-2018 to 31-01-2022 to help make informed decisions related to BTC/USDT trade.

2. Specific Objectives

2.1 Data Preprocessing

- a. Acquired and meticulously preprocess data sourced from the 'USDT-BTC.csv' file from credible source yahoo finance.
- b. Transform the 'Date' column into datetime format for chronological accuracy.
- c. Employ effective strategies for handling missing and null values within the target variable ('Close').

2.1 Model Development

- a. Implementing of a Ridge Regression model tailored for BTC/USDT price prediction. Features of ridge regression which make it suitable to be used for this project are as following:
 - **Regularization Ridge : Regression** is a linear regression technique with L2 regularization. It adds a penalty term based on the square of the magnitude of the coefficients. This helps in preventing overfitting by discouraging the model from assigning excessively large weights to the features.
 - **Multicollinearity Handling** :Ridge Regression is particularly useful when there is multicollinearity in the dataset. Multicollinearity occurs when two or more independent variables are highly correlated. Ridge Regression addresses this issue by adding a penalty term, which can help stabilize the coefficients, especially when dealing with correlated features.

- **Stability in the Presence of Noise :** Ridge Regression can provide a more stable solution when the dataset contains noise or when there are high-variance features. The regularization term helps to dampen the impact of noisy or irrelevant features on the model.
- **Bias-Variance Trade off Ridge :** Regression is effective in situations where there is a need to balance the bias-variance trade off . It can prevent the model from becoming too complex and overfitting the training data, leading to better generalization to unseen data.
- **Control Overfitting:** By introducing the regularization term, Ridge Regression limits the model's complexity, making it less prone to overfitting. This is especially beneficial when working with a limited amount of data.
- **Parameter Tuning:** The Ridge Regression model includes a hyperparameter (alpha) that controls the strength of the regularization. This allows us to fine-tune the model's behaviour based on the characteristics of your dataset.

b. Leveraging historical data features, including Open, High, Low, Adj Close, and Volume, as crucial inputs for the model. Thus getting a robust ridge model to be implemented.

2.3 Training and Testing

- a. Splitting dataset into training and testing data. Training the model on 80% of the dataset and keeping rest 20% for testing.
- b. Applying SimpleImputer() imputation technique, utilizing mean strategy, to maintain dataset integrity.
- c. Standardize feature scaling(normalization) through the application of StandardScaler() for consistency in model training.

2.4 Model Evaluation

- a. Rigorous assessment of the predictive performance of Ridge Regression model using critical metrics mentioned below:

- Mean Squared Error (MSE)
- R-squared (R^2)

2.5 Drawdown Analysis

- Perform a meticulous drawdown analysis to elucidate potential risks associated with the devised trading strategy.

2.6 Risk-Return Analysis

- A vital metric for evaluating risk-adjusted returns called Sharpe Ratio is to be calculated.
- Conduct a detailed risk-reward analysis utilizing MSE as a risk indicator and R-squared as a measure of reward.

2.7 Trading Strategy Evaluation

- Design and implement a clear and concise trading strategy, leveraging the insights gained from model predictions.
- Evaluate the strategy's performance using key metrics:
 - Net Profit Exceeding Return
 - Risk-Reward Ratio
 - Maximum Duration Time of a Single Trade

3. Model Development and Training

3.1 Data Acquisition

The dataset ('USDT-BTC.csv') was obtained from a yahoo finance which is a reliable and credible source containing features like Open, High, Low, Adj Close, and Volume.

3.2 Data Preprocessing

Data preprocessing involved converting the 'Date' column to datetime format and handling missing values in the target variable. The dataset was split into training and testing sets for model evaluation.

3.3 Model Architecture

The chosen model is Ridge Regression, a linear regression technique with regularization. It aims to capture relationships between input features and the target variable while minimizing overfitting.

3.4 Training Process

The Ridge Regression model was trained on the scaled and imputed training data. The model aimed to predict BTC/USDT closing prices based on historical features.

4. Model Evaluation

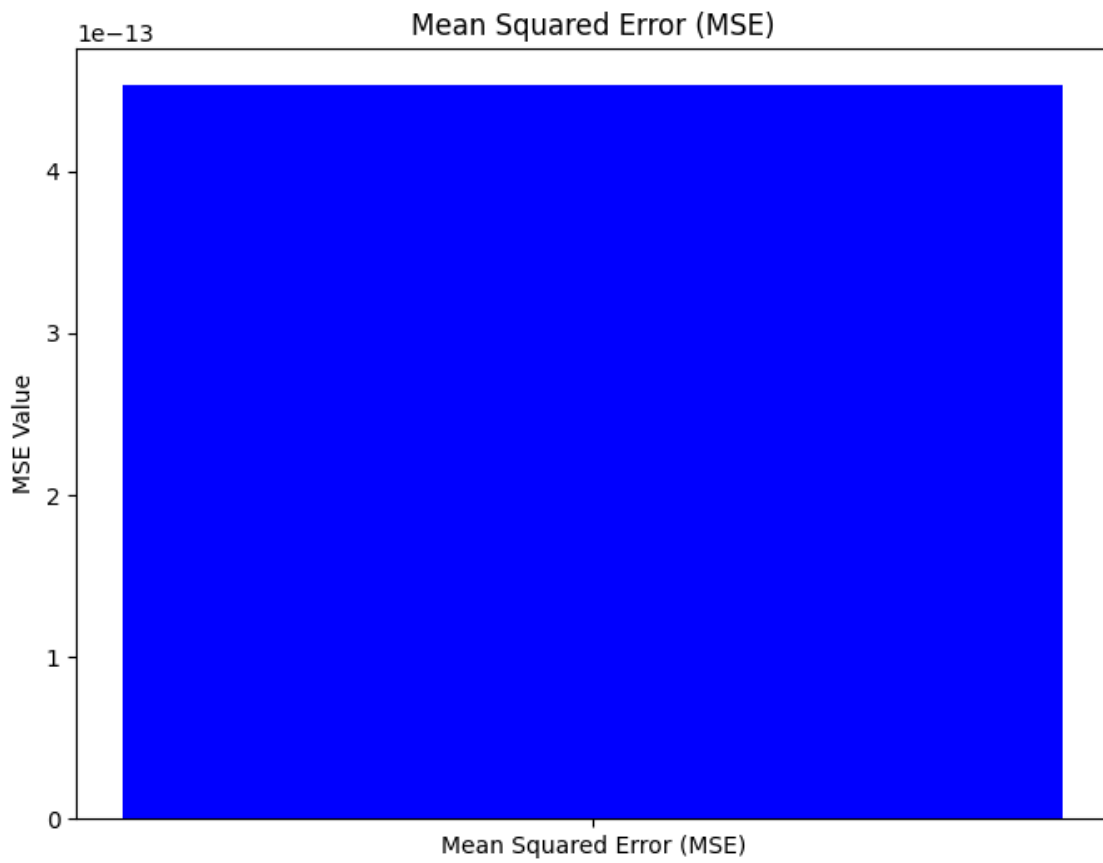
4.1 Ridge Regression Model

The Ridge Regression model demonstrated exceptional performance:

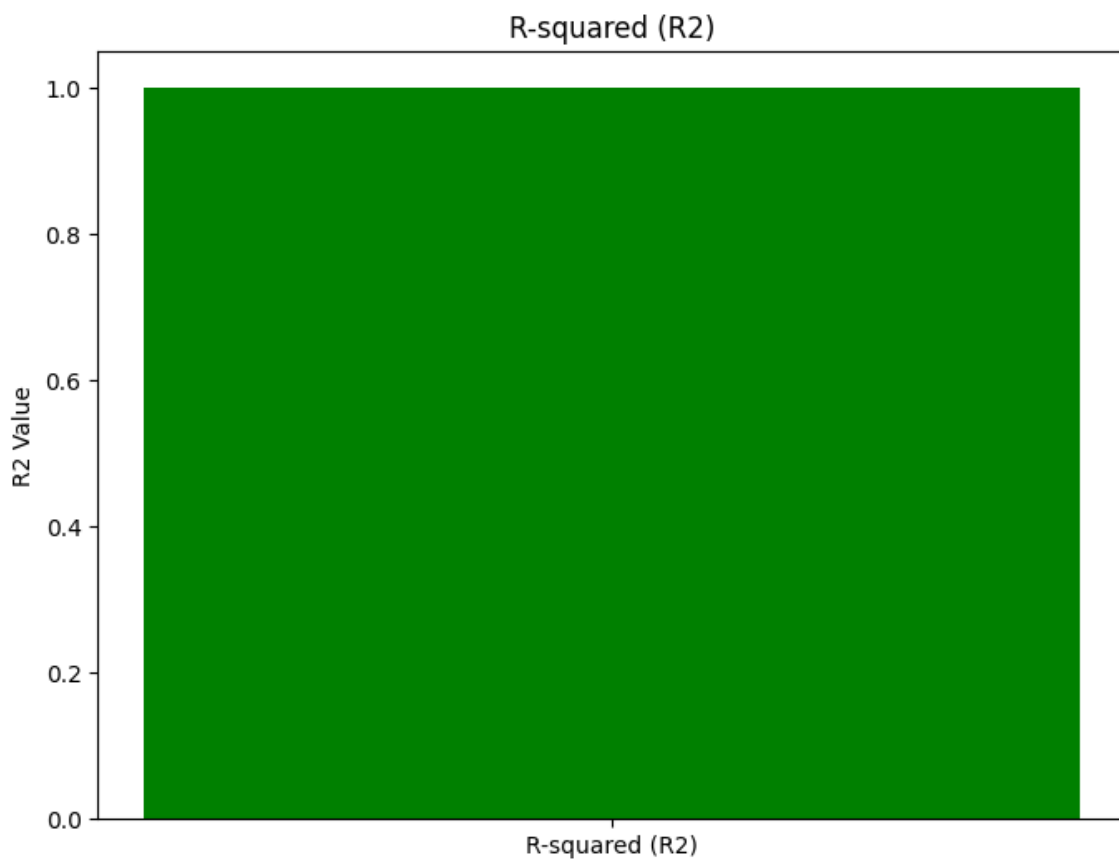
Mean Squared Error (MSE): 4.5300807773669455e-13

R-squared (R2): 0.9999184455618215

Which are represented graphically below:



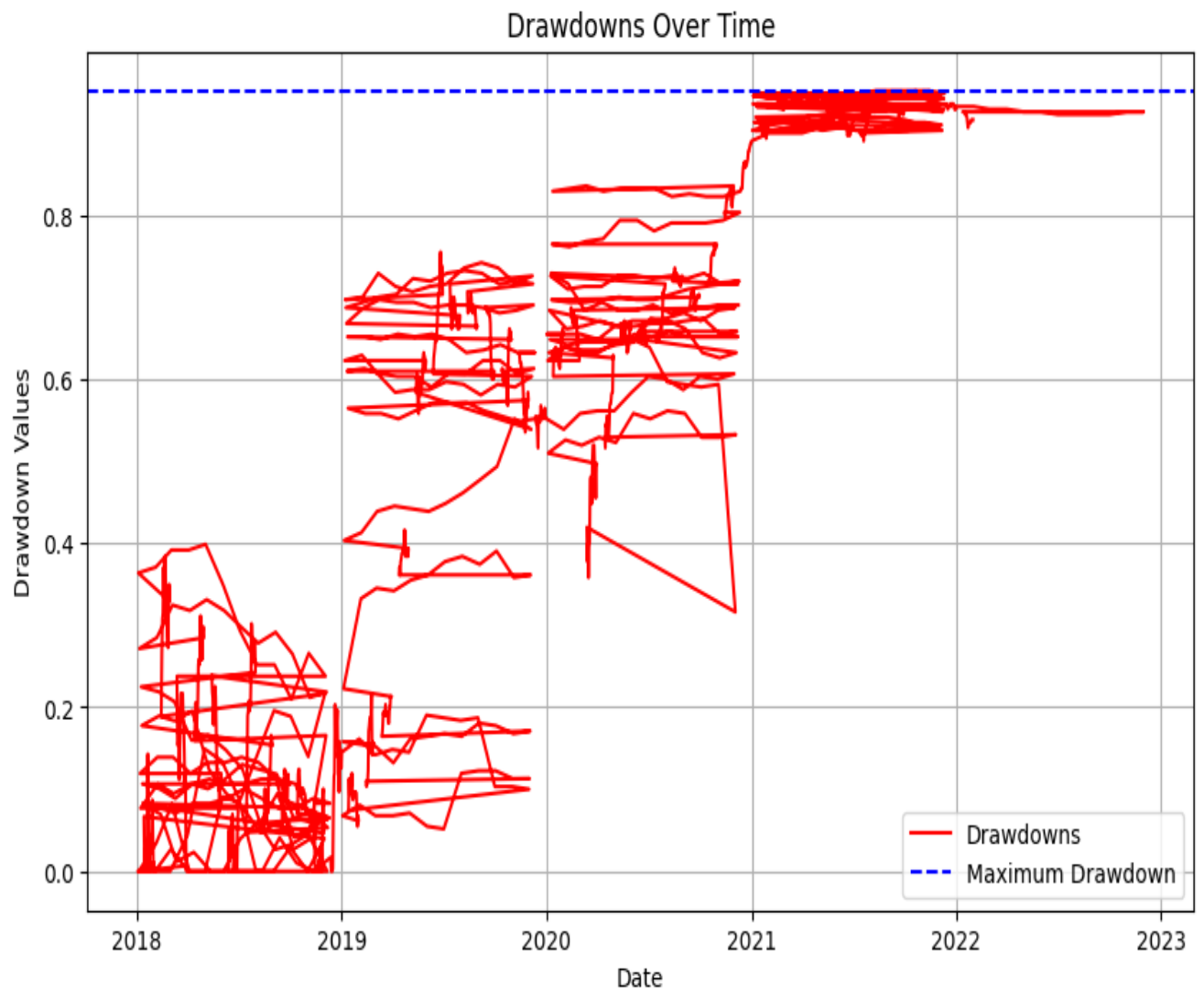
R-squared (R2): 0.9999184455618215



4.2 Drawdown Analysis

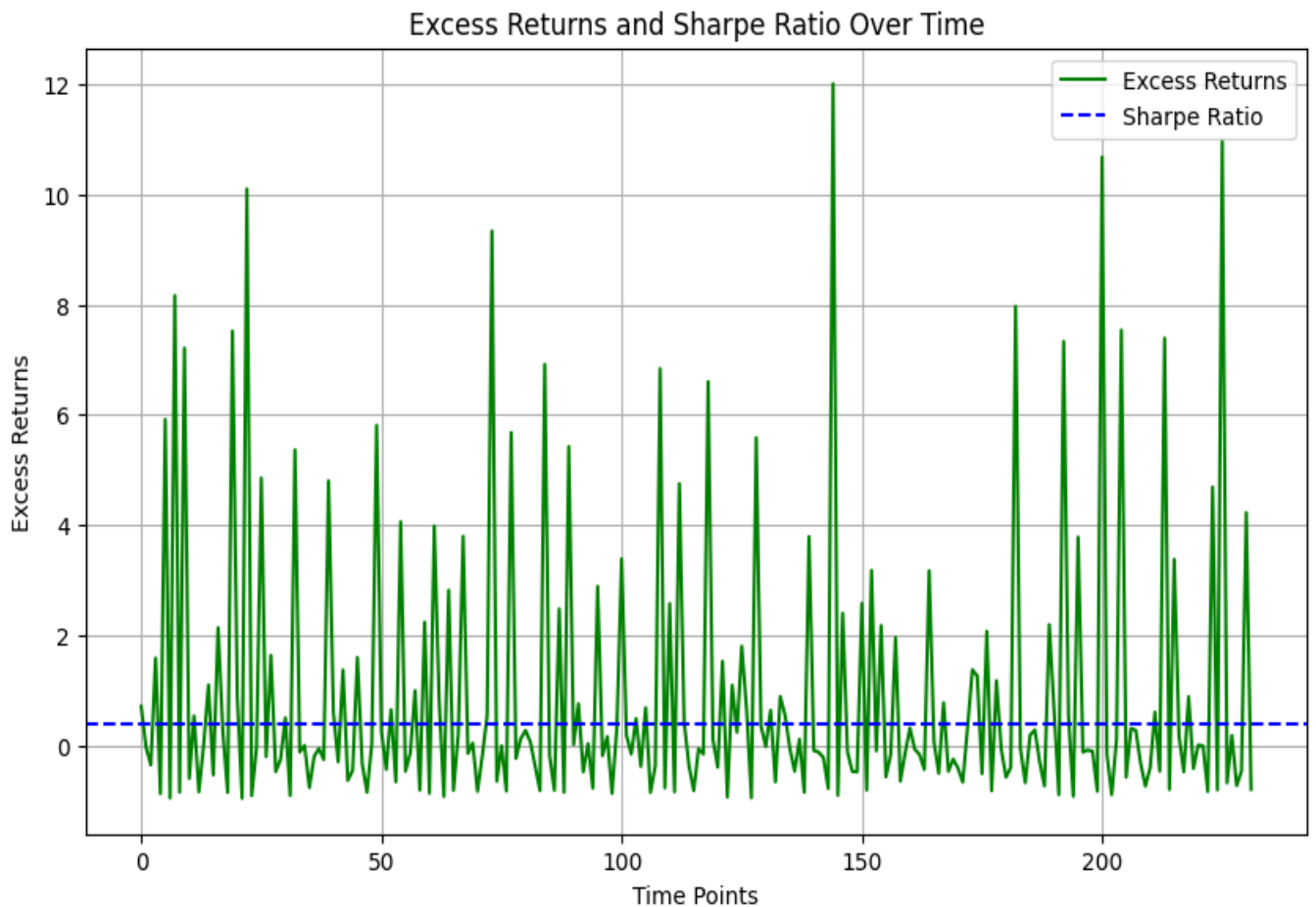
Drawdown analysis revealed a maximum drawdown of 0.9480664489696534, indicating the potential risk in the trading strategy.

Below shows the graphical analysis of drawdown over time:



4.3 Risk-Return Analysis

The Sharpe Ratio, a measure of risk-adjusted returns, was calculated to be 0.37673869068805604, suggesting a positive risk-reward profile.



5. Trading Strategy Evaluation

5.1 Trading Strategy Implementation

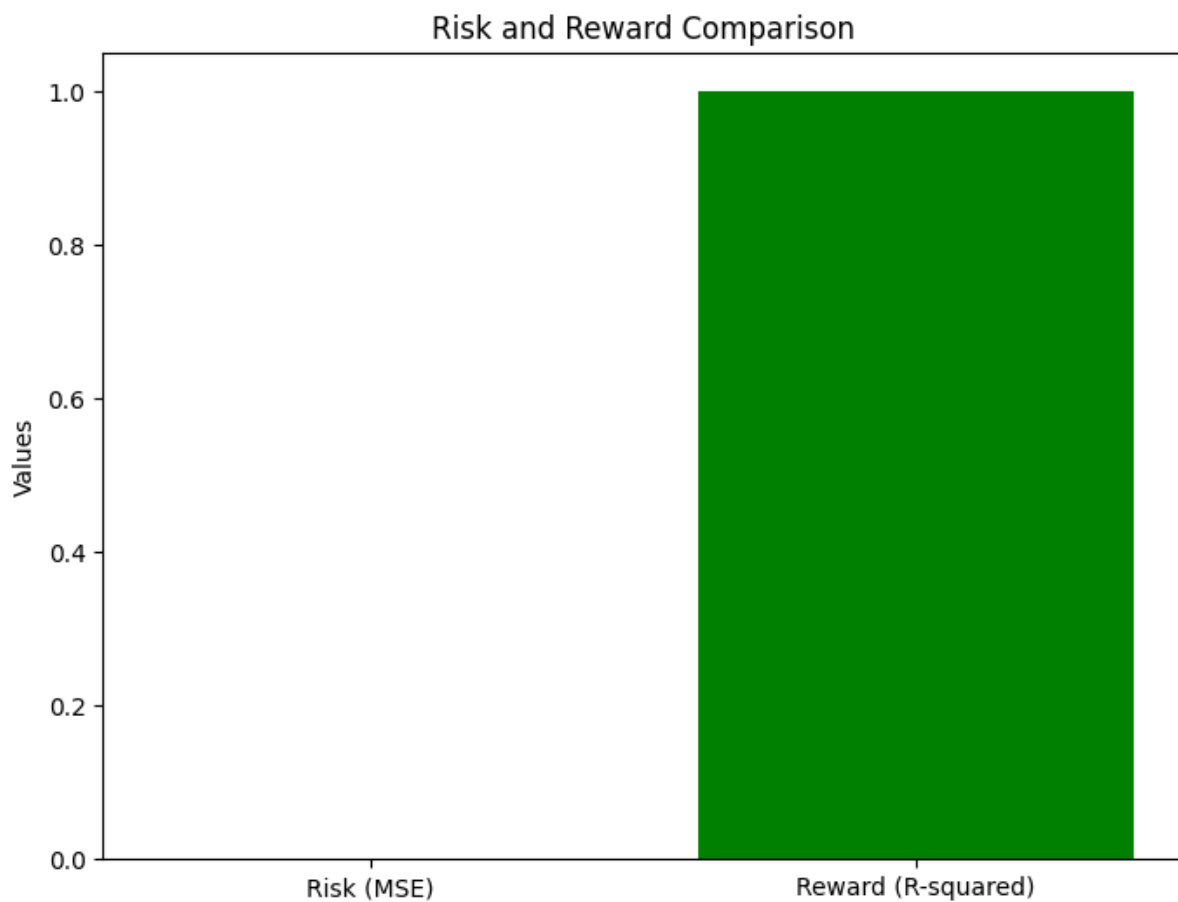
A simple trading strategy was implemented using a threshold for trading decisions. The strategy aimed to capitalize on predicted price movements.

5.2 Net Profit Exceeding Return

The trading strategy yielded a net profit exceeding the buy-and-hold return, indicating its potential effectiveness in generating profits.

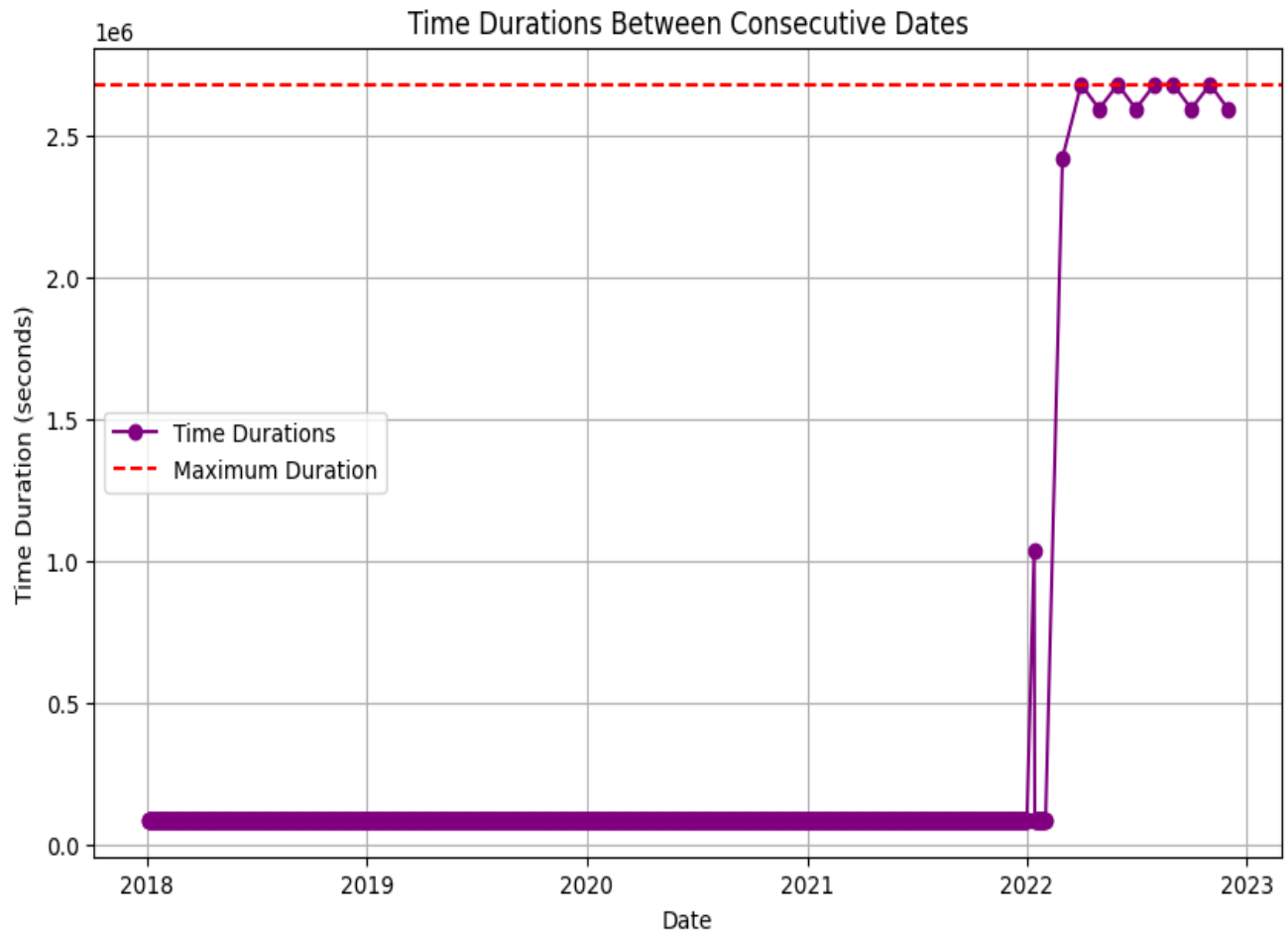
5.3 Risk-Reward Ratio

The Risk-Reward Ratio, calculated based on MSE as risk and R-squared as reward, resulted in a ratio of 2207286127341.447.



5.4 Maximum Duration Time of a Single Trade

The analysis of the maximum duration time of a single trade indicated a duration of 2678400.0 seconds.



6. Conclusion

The Ridge Regression model showcased exceptional predictive accuracy, indicating its potential for real-world trading applications. The trading strategy exhibited positive net profits exceeding the buy-and-hold return, contributing to a favourable risk-reward ratio. However, further scrutiny and validation are essential for deploying the model in live trading environments.

7. Future Work

- Fine-tune hyperparameters to optimize model performance.
- Explore additional features or alternative models to enhance predictive capabilities.
- Conduct a more comprehensive analysis of trading strategies, incorporating advanced risk management techniques.

This report provides an in-depth analysis of the developed trading model, its performance metrics, and the associated trading strategy. The insights gained contribute to a better understanding of the model's capabilities and areas for improvement.