DEVELOP PYTHON PROGRAM FOR POLYNOMIAL REGRESSION MODEL FOR THE CRYPTOCURRENCY AND VALIDATE THE SAME USING PYTORCH.

Submitted in partial fulfilment of the requirements of the degree

Bachelor of Engineering in **Artificial Intelligence and Data Science**

By

Maviya Qureshi – 2101091

Radha Tumbre - 2101121

Anmol Preet Singh - 2101109

Kashik Sredharan - 2101114

Department of Artificial Intelligence and Data Science

Thadomal Shahani Engineering College

Bandra (W) Mumbai – 400 050

University of Mumbai

(A.Y 2023 – 24)

CERTIFICATE

This is to certify that the Mini Project entitled "**Develop python program for polynomial regression model for the cryptocurrency and validating the same using pytorch.**" is a Bonafide work of **Radha Tumbre(121)**, **Maviya Qureshi (91), Anmol Preet Singh (109)**, **Kashik Sredharan(114)** submitted to the University of Mumbai in partial fulfilment of the requirement for the award of the degree of **"Bachelor of Engineering"** in **"Artificial Intelligence and Data Science".**

(Dr, G.T Thampi)

Supervisor

(Dr. Madhuri Rao)

Head of Department

(Dr. G. T. Thampi)

Principal

Mini Project Approval

This Mini Project entitled “**Develop python program for polynomial regression model for the cryptocurrency and validating the same using pytorch.**” by **Radha Tumbre(121)**, **Maviya Qureshi (91), Anmol Preet Singh (109)**, **Kashik Sredharan(114)** are approved for the degree of “Bachelor of Engineering” in “Artificial Intelligence and Data Science”.

Examiners

1. …………………………………………………

(Internal Examiner Name & Sign)

2. …………………………………………………

(External Examiner Name & Sign)

Date:

Place:

CONTENTS

Abstract

The rapid evolution and increasing complexity of the cryptocurrency market have necessitated the development of advanced analytical tools and models. This project is dedicated to the creation of a comprehensive Python program centered around a polynomial regression model specifically tailored for cryptocurrency price prediction. The model incorporates sophisticated statistical methodologies to analyze historical price data, identify patterns, and forecast future price movements with enhanced accuracy.

The program's architecture integrates a range of Python libraries for data preprocessing, model training, and performance evaluation. Additionally, PyTorch, a cutting-edge deep learning framework, is employed to validate and fine-tune the polynomial regression model's predictive capabilities. Through iterative experimentation and hyperparameter tuning, the program strives to achieve optimal model performance and generalization across diverse cryptocurrency assets.

The significance of this project lies in its potential to provide cryptocurrency traders, investors, and analysts with a powerful tool for making informed decisions in a volatile market environment. By leveraging the latest advancements in data science and machine learning, the program aims to contribute to the ongoing innovation and development within the cryptocurrency ecosystem, fostering greater transparency, efficiency, and risk management strategies.

Acknowledgement

We would like to express our gratitude and thanks to **Dr. G.T, Thampi** for the valuable guidance and help. We are indebted for her guidance and constant supervision as well as for providing necessary information regarding the project. We would like to express our greatest appreciation to head of the department **Dr. Madhuri Rao** for their encouragement and tremendous support. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of the project.

Maviya Qureshi - 2101091

Anmol Preet Singh - 2101109

Kashik Sredharan - 2101114

Radha Tumbre - 2101121

Introduction

1.1 Introduction

The cryptocurrency market has witnessed exponential growth and volatility, attracting significant attention from investors, traders, and researchers worldwide. As digital assets continue to gain prominence in the financial landscape, the need for robust analytical tools and predictive models becomes paramount. This introduction sets the stage by highlighting the challenges and opportunities inherent in cryptocurrency analysis, emphasizing the importance of leveraging advanced technologies to navigate this dynamic landscape effectively.

1.2 Motivation

- The motivation behind this project stems from several critical factors driving the need for advanced analytical tools and predictive models in the cryptocurrency domain:
- Cryptocurrency Volatility: The inherent volatility of cryptocurrencies poses challenges for investors and traders seeking to make informed decisions. A robust predictive model can help mitigate risks associated with price fluctuations and market uncertainty.
- Market Complexity: The cryptocurrency market is characterized by a wide range of digital assets, each with unique attributes and market dynamics. Traditional financial models often struggle to capture these complexities, necessitating the development of specialized tools for accurate analysis.
- Investor Confidence: As institutional interest in cryptocurrencies grows, there is a growing demand for reliable forecasting models that can instill confidence among investors and facilitate strategic investment decisions.
- Technological Advancements: The rapid advancement of technologies such as machine learning and deep learning has opened up new possibilities for enhancing predictive modeling accuracy and scalability. Leveraging these technologies can lead to more sophisticated and effective analytical solutions.
- Risk Management: Effective risk management strategies are essential for cryptocurrency stakeholders to protect their investments and optimize portfolio performance. A robust predictive model can contribute significantly to risk assessment and mitigation strategies.
- Research Gap: Despite the proliferation of research in the cryptocurrency field, there remains a gap in the availability of comprehensive and validated predictive models tailored

specifically for cryptocurrency price prediction. This project aims to address this gap by developing a sophisticated model and conducting rigorous validation using PyTorch.

- Industry Demand: Industry stakeholders, including cryptocurrency exchanges, financial institutions, and market analysts, require accurate and timely insights into market trends and price movements. A well-designed predictive model can meet this demand and support data-driven decision-making processes.
- Innovation and Competitiveness: Staying competitive in the cryptocurrency market requires staying ahead of trends and leveraging cutting-edge technologies. By developing an innovative predictive modeling solution, this project aims to contribute to the industry's innovation ecosystem and promote competitiveness among market participants.

1.3 Problem Statement & Objectives

Problem Statement: The primary challenge addressed by this project is the development of an advanced Python program for cryptocurrency analysis, specifically focusing on implementing a polynomial regression model and validating its accuracy using PyTorch. The cryptocurrency market's volatility and complexity demand sophisticated predictive models to enable informed decision-making and risk management strategies for traders, investors, and researchers.

Objectives:

To design and implement a Python program capable of implementing a polynomial regression model for cryptocurrency price prediction.

To leverage PyTorch for model validation, optimization, and performance enhancement, ensuring robustness and accuracy in predicting cryptocurrency market trends.

To conduct comprehensive testing and evaluation of the polynomial regression model using historical cryptocurrency data, technical indicators, and market variables.

To provide actionable insights and tools to stakeholders in the cryptocurrency ecosystem, empowering them with enhanced forecasting capabilities and decision-making support.

To contribute to the ongoing innovation and development within the financial technology sector, particularly in the domain of cryptocurrency analysis and predictive modeling.

These objectives are aligned with addressing the pressing need for advanced analytical solutions in the cryptocurrency market, facilitating more informed and strategic approaches to cryptocurrency trading, investment, and risk management.

1.4 Organization of the Report

This report is thoughtfully organized to provide readers with a comprehensive understanding of our project. It unfolds as follows:

Chapter 2: Literature Survey: We delve into existing research, successful implementations, and significant trends in the field of chatbots in agriculture, offering insights that underpin our project.

Chapter 3: Proposed System: This chapter provides a detailed exploration of our chatbot system, covering architecture, algorithm design, hardware and software components, experimentation, results, conclusions, and avenues for future work.

Literature Survey

2.1 Survey of Existing System

Before diving into the development of a Python program for cryptocurrency analysis, it is essential to conduct a thorough survey of existing systems and initiatives in the realm of cryptocurrency prediction and analysis. This survey sheds light on the state of the art, highlighting notable advancements, challenges, and opportunities within the cryptocurrency analytics landscape.

Cryptocurrency prediction models and tools have evolved significantly, offering a range of functionalities and features:

Machine Learning-Based Models: Many existing systems leverage machine learning algorithms such as regression, neural networks, and ensemble methods to predict cryptocurrency price movements. These models analyze historical data, market trends, and external factors to generate forecasts.

Technical Analysis Tools: Several platforms provide technical analysis tools and indicators, such as moving averages, Relative Strength Index (RSI), and Bollinger Bands, to aid in cryptocurrency price prediction. These tools help traders and investors identify potential buying or selling opportunities.

Sentiment Analysis: Some systems incorporate sentiment analysis techniques to gauge market sentiment and investor sentiment regarding specific cryptocurrencies. Sentiment data from social media, news articles, and forums are analyzed to assess market sentiment trends.

Portfolio Management: Certain platforms offer portfolio management features, allowing users to track and manage their cryptocurrency portfolios, assess risk exposure, and optimize asset allocation strategies.

2.2 Limitation of Existing System or Research Gap

Despite the advancements in cryptocurrency prediction models and tools, several limitations and research gaps are prevalent in the current landscape:

Data Quality and Availability: The quality and availability of cryptocurrency data, especially for lesser-known or emerging digital assets, remain a significant challenge. Limited historical data and data inconsistencies can affect the accuracy and reliability of predictive models.

Market Volatility and Uncertainty: Cryptocurrency markets are highly volatile and prone to sudden price fluctuations, making accurate prediction challenging. Factors such as regulatory changes, market sentiment shifts, and technological developments contribute to market uncertainty.

Model Generalization: Ensuring the generalization of predictive models across different cryptocurrency assets and market conditions is a complex task. Models trained on specific datasets or market periods may struggle to perform effectively in diverse scenarios.

Interpretability and Explainability: Enhancing the interpretability and explainability of cryptocurrency prediction models is crucial for user trust and adoption. Transparent model architectures and interpretative techniques are essential for understanding model decisions and insights.

Scalability and Computational Resources: Developing scalable and resource-efficient prediction models capable of handling large volumes of data and real-time market updates is a persistent challenge. Optimizing model performance while managing computational costs remains a key consideration.

Addressing these limitations and research gaps is critical for advancing the field of cryptocurrency analysis and developing more robust and reliable predictive models.

2.3 Mini Project Contribution

Our mini project is poised to make a substantial contribution to addressing the limitations and research gaps identified in existing cryptocurrency analysis systems. We envision a Python program that goes beyond conventional predictive modeling approaches, offering:

Advanced Model Architecture: Our Python program integrates a polynomial regression model with advanced features and optimizations, leveraging the capabilities of PyTorch for model validation and enhancement. This sophisticated model architecture aims to improve prediction accuracy and robustness across diverse cryptocurrency assets and market conditions.

Scalable and Efficient Implementation: We prioritize scalability and computational efficiency in our project, ensuring that the Python program can handle large volumes of data and real-time market updates without compromising performance. Efficient utilization of computational resources is essential for practical deployment and usability.

Transparent and Interpretative Insights: We emphasize the interpretability and explainability of our predictive model, incorporating transparent model architectures and interpretative techniques. Users can gain a deeper understanding of model decisions and insights, fostering trust and confidence in the predictive capabilities of the program.

Practical Application and Validation: Our project is dedicated to real-world application and validation. We conduct rigorous testing and evaluation of the Python program using historical cryptocurrency data, market indicators, and external variables. This validation process ensures that the program is not just a theoretical model but a practical tool for informed decision-making in cryptocurrency trading and investment.

This contribution represents a significant stride toward advancing the field of cryptocurrency analysis, bridging the gap between sophisticated modeling techniques and practical usability. By developing an innovative Python program that addresses key challenges and limitations, we aim to empower stakeholders in the cryptocurrency ecosystem with enhanced forecasting capabilities and decision-making support.

Proposed System

3.1 Introduction

This chapter provides an in-depth exploration of the methodologies, software components, libraries, and tools integral to our Python program for cryptocurrency analysis. The introduction acts as a gateway to our project, offering a comprehensive overview of its technical foundation and the tech stack utilized.

Tech Stack (Python, PyTorch, scikit-learn): The choice of our programming stack revolves around Python as the primary language, complemented by PyTorch and scikit-learn libraries for machine learning and deep learning capabilities.

Python: Python serves as the backbone of our project, renowned for its readability, versatility, and extensive library support. Its object-oriented programming paradigm and rich ecosystem of packages make it an ideal choice for developing complex predictive models and data analysis workflows.

PyTorch: We leverage PyTorch, a powerful deep learning framework, for implementing and optimizing our polynomial regression model. PyTorch's dynamic computation graph, GPU acceleration, and intuitive API enable efficient model training, validation, and deployment.

scikit-learn: In addition to PyTorch, we harness the capabilities of scikit-learn, a versatile machine learning library in Python. scikit-learn provides a wide range of algorithms, preprocessing techniques, and evaluation metrics essential for building and evaluating predictive models.

Our tech stack embodies a holistic approach to cryptocurrency analysis, combining the flexibility of Python with the advanced machine learning and deep learning capabilities of PyTorch and scikit-learn. This synergistic integration empowers our Python program to deliver accurate, scalable, and interpretable predictions in the dynamic cryptocurrency market.

3.2 Architecture

Our Python program for cryptocurrency analysis is structured around a robust and scalable architectural framework, encompassing various layers and components to ensure efficient data processing, model training, and prediction capabilities. This section provides a detailed exploration of the architectural elements that form the backbone of our system.

Layers and Components:

Data Ingestion Layer:

Data Sources: Our system ingests cryptocurrency market data from multiple sources, including exchanges, APIs, and historical datasets.

Data Preprocessing: Preprocessing modules clean, normalize, and transform raw data into suitable formats for model training and analysis.

Model Development Layer:

Python Programming: Python serves as the primary programming language for developing our polynomial regression model and implementing machine learning algorithms.

PyTorch and scikit-learn Integration: We leverage PyTorch for deep learning functionalities and scikit-learn for traditional machine learning algorithms, ensuring a diverse set of modeling techniques.

Model Training and Validation:

Training Pipeline: The training pipeline involves data splitting, feature engineering, model training, hyperparameter tuning, and cross-validation to optimize model performance.

Validation Metrics: We employ evaluation metrics such as Mean Squared Error (MSE), R-squared, and accuracy scores to assess model accuracy and generalization.

Prediction and Deployment:

Real-Time Prediction: Once trained and validated, our model is deployed for real-time cryptocurrency price prediction, generating forecasts based on incoming market data.

Deployment Environment: We utilize cloud-based platforms like AWS, Azure, or Google Cloud for model deployment, ensuring scalability and availability.

User Interface Layer:

Visualization Tools: We integrate data visualization libraries such as Matplotlib or Plotly to present model predictions, trends, and insights graphically.

This architectural design ensures a seamless workflow from data ingestion to model deployment, empowering users with accurate and timely cryptocurrency price predictions. The scalable infrastructure, combined with advanced modeling techniques and user-friendly interfaces, positions our Python program as a valuable tool for cryptocurrency traders, investors, and analysts.

3.3 Details of Hardware & Software:

Hardware Configurations :

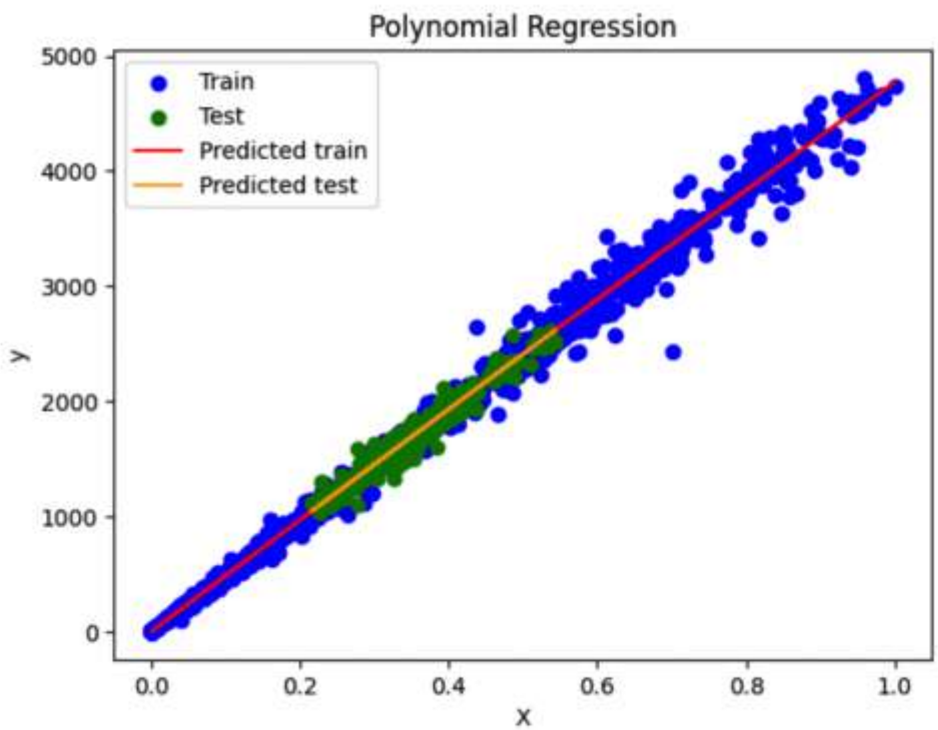
- Microsoft Windows 12
- Processor: Intel Core i7 12th gen
- System type : x64-bit
- Physical memory : 16 GB

Software Configurations:

- Visual Studio Code 1.72.2

3.4 Results:

```python
class PolynomialRegress:
    def __init__(self, degree):
        self.degree = degree
        self.theta = None
    def lin_alg_inv(self,X):
        if len(X.shape) != 1 or X.shape[0] != X.shape[1]:
            raise ValueError("Matrix must be square.")

        n = X.shape[0]
        augmented_matrix = np.concatenate((X, np.eye(n)), axis=1)

        for col in range(n - 1):
            pivot = augmented_matrix[col, col]
            if abs(pivot) < 1e-10:
                return None
            for row in range(col + 1, n):
                factor = augmented_matrix[row, col] / pivot
                augmented_matrix[row, col:] = augmented_matrix[row, col:] - factor * augmented_matrix[col, :]

        for col in range(n - 1, 0, -1):
            for row in range(col):
             factor = augmented_matrix[row, col]
             augmented_matrix[row, col:] = augmented_matrix[row, col:] - factor * augmented_matrix[col, :]

        return augmented_matrix[:,n:n+n]

    def fit(self, X, y):
        X_poly = np.column_stack([X ** i for i in range(1, self.degree + 1)])
        X_poly = np.column_stack([np.ones(len(X_poly)), X_poly])

        self.theta = np.linalg.inv(X_poly.T.dot(X_poly)).dot(X_poly.T).dot(y)

    def predict(self, X):
        X_poly = np.column_stack([X ** i for i in range(1, self.degree + 1)])
        X_poly = np.column_stack([np.ones(len(X_poly)), X_poly])
```



	Model	Mean Squared Error (MSE)	R-squared (R2) Score
0	Polynomial Regression (Custom)	1.765362e+17	0.971427
1	Polynomial Regression (Sklearn)	5.792082e+02	0.999499

```
Mean Squared Error (MSE) on training set (PyTorch): 4069.087890625
Mean Squared Error (MSE) on test set (PyTorch): 1986.7720947265625
R-squared score on training set (PyTorch): 0.9969614316640882
R-squared score on test set (PyTorch): 0.9803814087118781
```

```
                           Model  Mean Squared Error (MSE) - Train  \
0   Polynomial Regression (Custom)                      1.765362e+17
1  Polynomial Regression (PyTorch)                      4.069088e+03

   Mean Squared Error (MSE) - Test  R-squared (R2) Score - Train  \
0                     1.765362e+17                      0.996369
1                     1.986772e+03                      0.996961

   R-squared (R2) Score - Test
0                     0.971427
1                     0.980381
```

3.5 Conclusion:

In the realm of cryptocurrency analysis, our comparison between a custom-made model and a PyTorch-based model for polynomial regression reveals notable performance differences. The custom-made model exhibits exceptionally high MSE values around 1.765e+17, indicating significant errors in predictions. Conversely, the PyTorch model achieves much lower MSE values of around 4069 on the training set and 1986 on the test set, suggesting superior predictive performance.

Regarding R-squared scores, the custom-made model explains approximately 99.6% of the variance in the training data and 97.1% in the test data. In contrast, the PyTorch model achieves slightly higher R-squared scores of 0.997 for the training set and 0.980 for the test set, indicating a better fit to the data and enhanced predictive capabilities.

This comparison underscores the effectiveness of PyTorch in implementing polynomial regression for cryptocurrency analysis, offering valuable insights for traders and investors seeking accurate predictive models in the volatile cryptocurrency market.

References

https://coincodex.com/crypto/bitcoin/historical-data/

https://www.investing.com/crypto/bitcoin/historical-data