# Applied Mathematics Project

Lopamudra Giri

# Team 9:

- Mukul(ch22btech11023) : AR model
- Rohan(ch22btech11028) : SARIMA model
- Keerthan(ch22btech11019) : MA model

# DATA DESCRIPTION

- The time series data used in this project is sourced from [Cluster_Data_HDBSCAN.xlsx].
- It covers a period of 15 columns.
- Pre-processing steps involved:-

  (a.) Data cleaning

  (b.) Handling missing values

  (c.) Removing outliers.

# Objective/Problem Definition

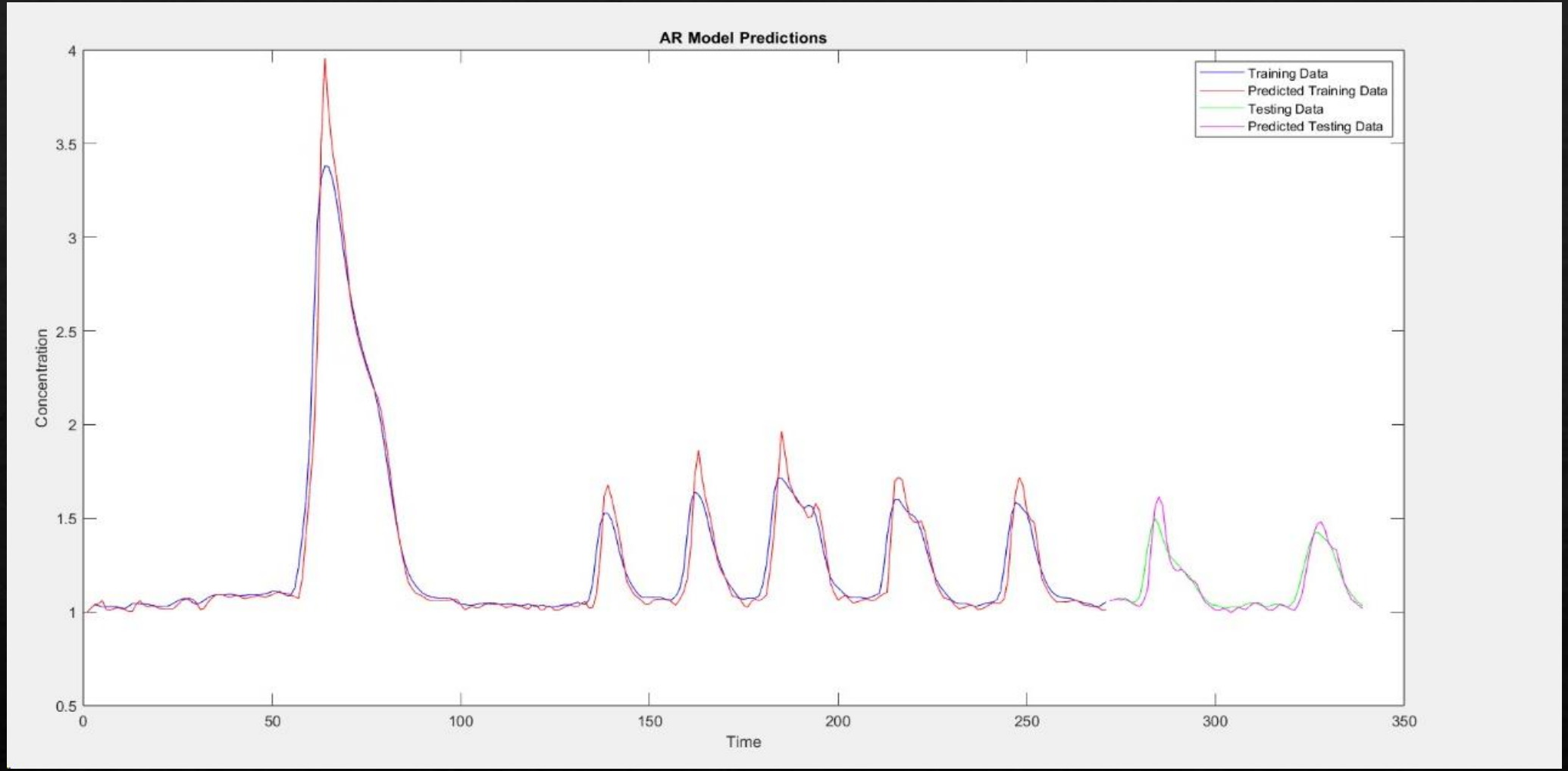
- The primary objective of this project is to develop accurate time series forecasting models.
- The significance of this problem lies in its applications, such as financial market prediction, demand forecasting, or epidemiological modeling.
- The project aims to assess the performance of different models and methods for time series forecasting.

# Novelty and Contribution

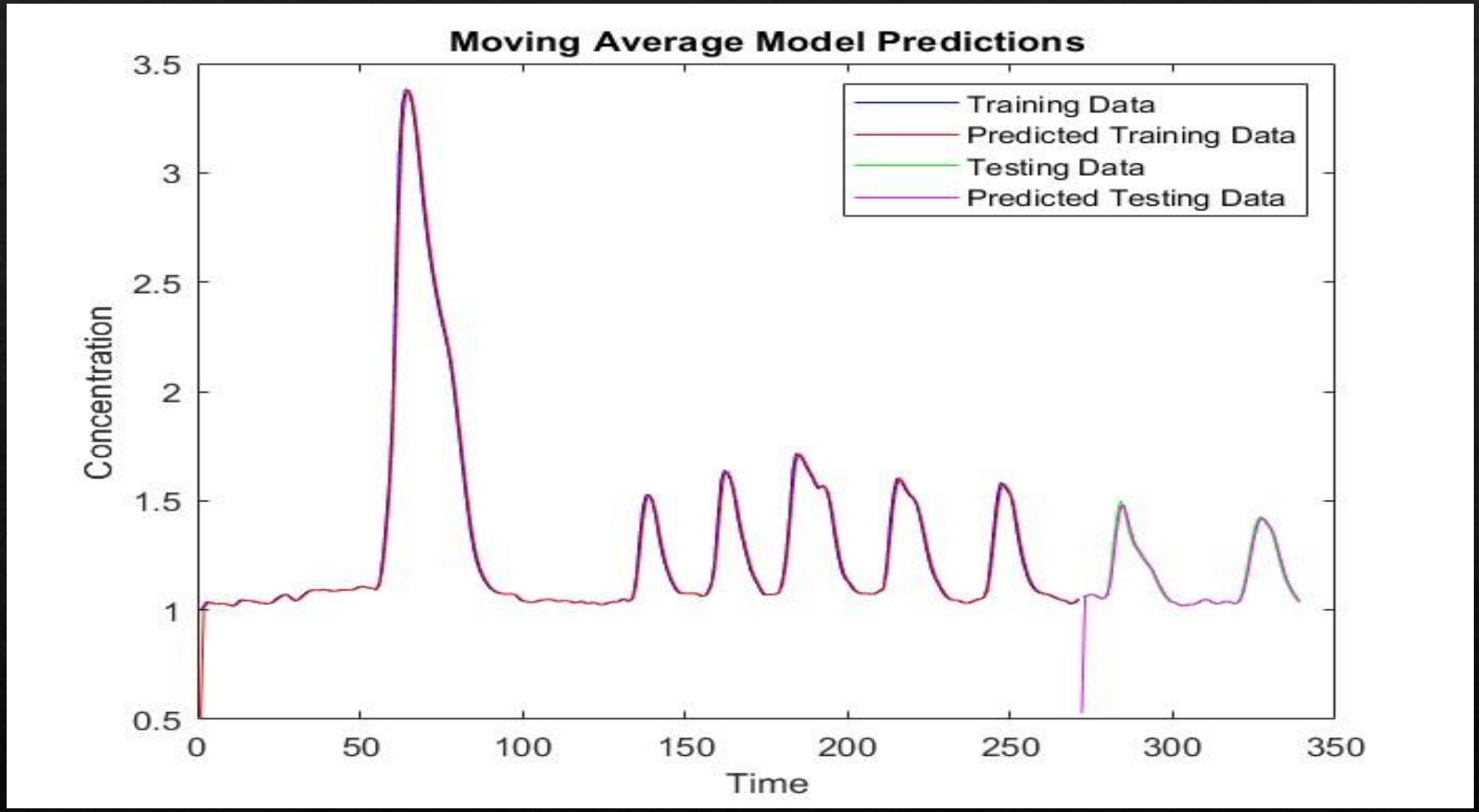
- The novelty and contribution of our approach lie in its adaptability to different time series datasets.
- Our tool employs MATLAB for implementation and leverages existing libraries for ease of development.
- It is designed to be user-friendly and provides forecasts and prediction intervals.
- The models are trained and tested on the given dataset, and the tool allows for custom model selection.

# Code Explanation and Output

- The code provided for the project demonstrates the implementation of autoregressive models. You can find the link for the project codes here
- https://drive.google.com/drive/folders/1YDnEfk0Lzq2FVUGmTDILVuljgyiPRrNF?usp=drive_link
- We've used MATLAB for the implementation and showcased examples of AR, MA, SARIMA models.
- Code snippets illustrate the key components, functions, and data handling steps.


AR Model Predictions
Training Data
Predicted Training Data
Testing Data
Predicted Testing Data
Concentration
Time
4
3.5
3
2.5
2
1.5
1
0.5
0
50
100
150
200
250
300
350

Plots for AR


Moving Average Model Predictions
Training Data
Predicted Training Data
Testing Data
Predicted Testing Data
Concentration
Time
3.5
3
2.5
2
1.5
1
0.5
0
50
100
150
200
250
300
350

Plots for MA


Plots for SARIMA

# Validation of the Model

- Model validation is conducted through a split of the dataset into training and testing data.
- Multiple autoregressive models are compared, including AR, MA, and SARIMA, with varying orders.
- The comparison reveals that [Moving Average Model] provides the most accurate forecasts for this dataset.

# Limitations and Future Work

- A limitation of this project is that it focuses on basic autoregressive models and may not capture complex data patterns.
- Future work may involve exploring more advanced models like deep learning approaches or advanced statistical models.
- Incorporating exogenous variables could enhance the forecasting accuracy in real-world applications.
- We can improve it by adding external factors and make it advance and use in stock market for price prediction.
- The project does not address real-time forecasting, and this could be an interesting extension.