Time Series Forecasting with ARIMA and ANN

# Combining Linear and Non-Linear Models for Enhanced Forecasting

## Muhammad Maemoon Farooq

# Introduction

Objective: To forecast time series data using a combination of ARIMA and ANN models.  
Dataset: Hourly energy consumption data (AEP\_hourly.csv), S&P500 index (sp500\_index.csv), CO2 concentration data.

# Dataset Overview

## AEP\_hourly.csv:

• Datetime: Timestamps of the readings.  
• AEP\_MW: Energy consumption in megawatts.

## sp500\_index.csv:

• Date: Date of the readings.  
• S&P500: S&P 500 index value.

## CO2\_concentration.csv:

• Columns: year, month, day, cycle, trend, etc.

# Data Preparation

• Loading Data: Data is loaded from CSV files.  
• Datetime Conversion: Converting Date and Datetime columns to datetime format.  
• Handling NaNs: Forward and backward filling of missing values.

# ARIMA Model

Objective: Capture linear patterns in time series data.  
Methodology:  
• Split data into training (80%) and testing (20%).  
• Fit ARIMA model to the training data.  
• Forecast and calculate residuals on the testing data.

## ARIMA Model Explanation

ARIMA: AutoRegressive Integrated Moving Average.  
Components:  
• AR (p): Autoregression.  
• I (d): Integration (differencing).  
• MA (q): Moving Average.  
Process:  
• Identify the order of the model (p, d, q).  
• Fit the model to the historical data.  
• Generate forecasts and calculate residuals.

# ANN Model

Objective: Capture non-linear patterns in ARIMA residuals.  
Methodology:  
• Normalize residuals.  
• Create sequences for training the ANN.  
• Train and test the ANN on residuals.

## ANN Model Explanation

ANN: Artificial Neural Network.  
Components:  
• Input Layer: Takes in the normalized residuals.  
• Hidden Layers: Processes the inputs through neurons.  
• Output Layer: Generates the forecast.  
Process:  
• Normalize the residuals.  
• Create input sequences.  
• Train the network on the sequences.  
• Test the network on unseen data.

# Combining Forecasts

Objective: Improve forecasting accuracy by combining ARIMA and ANN forecasts.  
Methodology: Average the ARIMA and ANN forecasts to form the final forecast.

## Combining Forecasts Explanation

Why Combine:  
• ARIMA captures linear trends.  
• ANN captures non-linear patterns.  
• Combining leverages strengths of both.  
Method:  
• Align forecasts from both models.  
• Average the forecasts to smooth out errors.

# Applying to Datasets

## AEP Dataset:

• Apply ARIMA and ANN models.  
• Combine forecasts and plot results.

## S&P 500 Dataset:

• Apply ARIMA model.  
• Plot results.

## CO2 Dataset:

• Apply ARIMA and ANN models.  
• Combine forecasts and plot results.

# Results

## AEP Dataset:

• ARIMA MAE, MSE, R-squared.  
• ANN MAE, MSE, R-squared.  
• Combined Forecast Plot.

## S&P 500 Dataset:

• ARIMA Model Summary.  
• Forecast Plot.

## CO2 Dataset:

• ARIMA MAE, MSE, R-squared.  
• ANN MAE, MSE, R-squared.  
• Combined Forecast Plot.

# Conclusion

## Summary:

• ARIMA captures linear patterns effectively.  
• ANN captures non-linear patterns in residuals.  
• Combining both improves forecast accuracy.

## Future Work:

• Explore other models like LSTM for time series.  
• Apply models to different datasets for robustness.

# Q&A

Questions: Open the floor for any questions from the audience.