Submitted By: Muhammad Bachal
Roll No.: 232563
Section: BSDS_V

Reproduction Report: Legaki & Koutsouri Method (M4 Competition)

Introduction

The M4 Competition includes 100,000 time series from different domains and frequencies. This report reproduces the method by Legaki & Koutsouri, which ranked 8th in the competition. The aim is to understand its workflow and check how easily the results can be replicated.

Method Overview

The method uses a statistical workflow that includes decomposition, feature extraction, model selection, and a weighted ensemble of forecasting models. It focuses on accuracy and full reproducibility.

Workflow Steps

1. Load the time series dataset.
2. Preprocess data and handle missing values.
3. Apply STL decomposition.
4. Extract statistical and seasonal features.
5. Select suitable models (ETS, ARIMA, Theta).
6. Perform rolling-origin evaluation.
7. Convert errors into model weights.
8. Generate forecasts from each model.
9. Combine forecasts using weighted ensemble.
10. Save final results and plots.

Challenges & Resolutions

Some series had formatting issues and missing values, which required cleaning before analysis. A few ARIMA/ETS models failed to fit, so fallback methods like seasonal naïve were used. Forecast horizons also differed across frequencies, so consistent rules were applied. Using checks, fallback strategies, and unified horizon mapping helped ensure smooth reproduction and accurate forecasts.