Research Proposal

Title: Electricity Demand Forecasting Using Time Series Analysis and Machine Learning Models

Team Name: Snipers

Team Members:

1. Abdallah Ahmed Abdallah
2. Ammar Yasser
3. Yara
4. Mohamed Yasser Ahmed
5. Mohamed Yasser Mohamed
6. Mario Rafat

Under Supervision of

Prof. Dr. Maha A. Hassanein

Problem Definition: Electricity demand is influenced by daily cycles, seasonal trends, and external factors like temperature. Accurate forecasting is essential for utility providers to ensure grid stability, optimize resources, and prevent power disruptions. With the growing integration of renewable energy, managing demand fluctuations has become even more critical, as renewables introduce supply variability that must be balanced with precise demand predictions.

This research proposes a model to forecast electricity demand using both time series analysis (ARIMA) and machine learning models (LSTM networks). By incorporating historical demand and weather data, the model aims to capture usage patterns and correlations with temperature, enabling utility providers to anticipate peak periods and adjust resource allocation. The study will evaluate the effectiveness of ARIMA and LSTM models in predicting demand and identify the optimal approach for supporting grid efficiency and reliability.

Statistical Questions:

1. **Are there significant differences in electricity demand across different seasons?**

    - **Test:** Use **ANOVA** (Analysis of Variance) to determine if there are statistically significant differences in demand across seasons.

2. **Is there a statistically significant relationship between temperature and electricity demand?**

    - **Test:** Apply a **Pearson correlation** or **Spearman rank correlation** to quantify the strength and direction of the relationship between temperature and demand.

3. **Does the inclusion of weather data, beyond temperature, improve the model's forecasting accuracy?**

    - **Test:** Perform a **regression analysis** with and without additional weather variables (e.g., humidity, wind speed) to determine if these factors significantly improve forecast accuracy.

4. **Are there significant differences in forecasting accuracy between the ARIMA and LSTM models?**

    - **Test:** Use a **paired t-test** or **Wilcoxon signed-rank test** to compare the error distributions (e.g., Mean Absolute Error or Root Mean Squared Error) of each model.

5. **How well does the model perform in predicting electricity demand during extreme temperature days compared to average days?**

- **Test:** Conduct an **independent t-test** or **Mann-Whitney U test** to compare model errors on days with extreme temperatures (e.g., very hot or cold days) against average temperature days.

6. **What is the lag effect of temperature on electricity demand, and how many days (or hours) of historical data provide the most accurate predictions?**

   - **Test:** Use **cross-correlation analysis** to determine the lag at which temperature has the most significant impact on demand.

7. **Does demand prediction accuracy vary significantly during holidays and weekends compared to regular weekdays?**

   - **Test:** Use an **independent t-test** to assess differences in prediction errors between regular days and holidays/weekends.

8. **What is the confidence interval of demand predictions made by the models, and how stable is it across different time frames?**

   - **Test:** Calculate **confidence intervals** around forecasted values to gauge the certainty of predictions, and compare intervals across different times (e.g., day, month) for stability.

9. **Do temperature anomalies (e.g., sudden spikes or drops) lead to greater forecast errors?**

   - **Test:** Perform an **outlier analysis** on temperature and examine if these anomalies coincide with significant deviations in forecast accuracy.

10. **Can ensemble methods that combine ARIMA and LSTM improve forecast accuracy?**

    - **Test:** Compare the errors of ensemble models with standalone ARIMA and LSTM models using an **F-test for variances** or a **paired t-test**.

Motivation for the Problem: As electricity demand continues to grow and more renewable sources are integrated, reliable demand forecasting becomes essential to maintaining grid stability and optimizing energy distribution. Accurate forecasts allow utility providers to schedule energy distribution, minimize operational costs, and reduce reliance on backup resources. This research will explore both statistical and machine learning models to find practical, accurate solutions for electricity demand forecasting. By incorporating patterns in daily and seasonal cycles as well as correlations with temperature, this research aims to enhance grid reliability and contribute to the sustainability of modern energy systems.