Project Title Household Electric Power Consumption Forecasting Using Time Series Models

Chapter 1: INTRODUCTION

1.1 Background

Electric power consumption is one of the most important factors in modern society, reflecting the lifestyle, economic activities, and technological growth of a household. Household electricity usage is influenced by multiple factors such as the number of family members, appliances in use, weather conditions (temperature, humidity), time of day, and seasonal variations. Electricity consumption directly impacts household expenses, energy demand forecasting, and sustainable resource management.

In recent years, with the growing population and modernization of households, the demand for electricity has rapidly increased. As a result, efficient monitoring and forecasting of power consumption have become essential for both energy providers and consumers. Reliable consumption forecasts help in optimizing energy distribution, avoiding wastage, reducing power cuts, and maintaining the stability of the national grid.

Time series analysis plays a crucial role in understanding and predicting household electricity consumption. It helps capture consumption patterns such as daily, weekly, and seasonal cycles. With advancements in Machine Learning (ML) and Artificial Intelligence (AI**)**, predictive models can be developed to learn historical data patterns and make accurate forecasts of future electricity usage.

Machine learning algorithms are broadly divided into three categories:

- Supervised Learning: Uses labeled data (e.g., past power consumption with timestamps) to predict future consumption. Methods include Linear Regression, Random Forest, Neural Networks, and Support Vector Machines (SVM).
- Unsupervised Learning**:** Clusters similar consumption patterns without predefined labels, e.g., grouping households by similar energy usage behavior. Methods include K-Means, Hierarchical Clustering, etc.
- Reinforcement Learning**:** Learns optimal strategies through feedback, e.g., smart energy-saving systems that reduce peak consumption.

Thus, household power consumption analysis using machine learning not only helps in understanding usage trends but also assists in smart energy management and sustainable development.

1.2 Statement of Problem

With the increase in urbanization and modernization, the demand for electricity in households has significantly risen. In many developing countries, including Nepal, the supply of electricity often struggles to keep up with the growing demand, leading to issues like load-shedding, voltage fluctuations, and higher energy costs.

Household electricity usage is irregular due to several factors such as lifestyle, weather, seasonal variations, and appliance usage patterns. Without proper forecasting and analysis, both consumers and power providers face challenges:

- Consumers: Difficulty in managing electricity bills and optimizing appliance usage.
- Energy Providers: Difficulty in balancing demand and supply, leading to energy shortages or wastage.

Hence, an accurate household power consumption forecasting system is necessary. Time series-based prediction models can help monitor and forecast energy demand, ensuring better planning, efficient usage, and reduction in unnecessary wastage.

1.3 Objectives

- To analyze household electricity consumption patterns over time.
- To identify the most influencing factors that affect electricity usage.
- To develop and evaluate predictive models for household power consumption forecasting.
- To select the best-performing model for future energy demand prediction

1.4 Scopes

The project focuses on building a time series forecasting model for household power consumption using historical data. The scope includes:

- Analyzing historical household power consumption trends (hourly, daily, seasonal variations).
- Building predictive models using time series techniques and machine learning approaches.
- Providing insights for energy providers to manage load distribution effectively.
- Assisting households in better energy usage planning and cost savings.
- While limited to household-level data, the methodology can be scaled to community, city, or national-level electricity forecasting.

1.5 Applications

- Smart Energy Management**:** Predictive models help households and smart grids optimize electricity usage and avoid wastage.
- Policy-Making Guidance: Government and electricity boards can use insights for future energy planning, tariff adjustments, and sustainable energy policies.

- Public Awareness**:** Consumers can track their usage patterns and adopt energy-saving practices.
- Research & Development: Provides a foundation for academic research in smart grids, renewable energy integration, and AI-driven energy systems.
- Industrial Application**:** Energy companies can use the model for real-time load forecasting and power supply optimization.

Chapter 2: LITERATURE REVIEW

In recent years, many prediction studies have been conducted in the field of energy and household power consumption. Since electricity demand forecasting is a major concern for researchers, several works have been carried out using different statistical, machine learning, and deep learning approaches.

A study presented a project whose objective was to forecast household electricity demand using time series models such as ARIMA and SARIMA. The results showed that these models could capture daily and seasonal consumption patterns but were limited in handling sudden changes in usage.

A study focused on household load forecasting using machine learning algorithms like Random Forest, Support Vector Regression, and Gradient Boosting. The dataset included electricity consumption data along with external factors such as temperature and holidays. The Random Forest model showed better performance in short-term forecasting compared to traditional approaches.

A work compared SARIMA, SVM, and LSTM models for short-term electricity demand prediction in residential areas. The research highlighted that LSTM provided more accurate predictions as it was able to learn long-term dependencies in time series data.

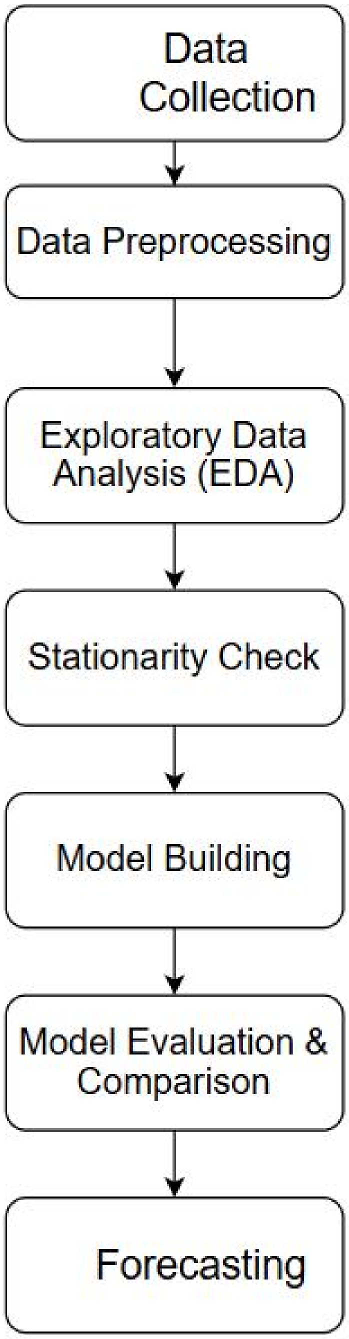
Paper published by CR and team employed logistic regression and autoregression methods to classify consumption levels (low, medium, high) and to predict future household electricity usage. Their system demonstrated that simple regression-based approaches can still provide useful insights for small-scale datasets.

Another platform developed a web-based energy monitoring and forecasting system where household users could check their current, hourly, and daily consumption. The study compared different machine learning algorithms including Neural Networks, Random Forest, and LSTM, and found that Neural Networks were highly effective for hourly prediction, while LSTM performed well in long-term forecasting.

After reviewing these works, it is clear that forecasting household electricity consumption is important for both energy providers and consumers. Machine learning and deep learning models such as Random Forest, ANN, and LSTM have shown great potential in providing accurate predictions compared to traditional statistical models. Our research aims to analyze household power consumption using time series techniques and to identify the most effective model for forecasting, which can help in energy planning, efficient usage, and cost management.

Chapter 3: METHODOLOGY

3.1 Workflow Diagram



3.2 Workflow

3.2.1 Data Collection

For this project, the dataset used was the Individual Household Electric Power Consumption dataset, which was obtained from the Kaggle open-source repository. The dataset is quite large, containing about 2,075,259 entries**,** recorded at a minute-level frequency**.** It includes several important features such as Global Active Power (kW*)*, Global Reactive Power (kVAR)*,* Voltage (V)*,* Global Intensity (A)*,* and three different sub-metering variables *(*Sub_metering_1, Sub_metering_2, Sub_metering_3*)*, which represent the energy consumed by different household appliances. This dataset provides a rich source of information to study and forecast household electricity consumption trends.

3.2.2 Data Preprocessing

Before analysis, the raw dataset required significant preprocessing. First, the Date and Time columns were merged to form a single datetime index**,** which helped in time series analysis. All the features were then converted into numeric types for consistency. Since the dataset was originally at minute-level frequency, it was resampled and standardized to ensure uniformity. Handling missing values was another crucial step linear interpolation was applied to fill in the gaps. Outliers were also detected using the Z-score method and replaced with median values to avoid skewed results. Furthermore, feature engineering was performed by generating additional variables such as hour of the day and day of the week*,* which were later used to capture daily and weekly consumption patterns.

3.2.3 Exploratory Data Analysis (EDA)

Exploratory analysis was carried out to better understand the dataset. A line plot of Global Active Power provided insights into the overall electricity consumption trend over time. To break down the structure of the series, seasonal decomposition was performed, which revealed three components trend, seasonality, and residuals. This analysis highlighted the existence of periodic daily and weekly usage cycles. Additionally, the Augmented Dickey-Fuller (ADF) test was applied to check for stationarity. The results confirmed that certain parts of the dataset needed transformation for effective time series modeling.

3.2.4 Model Building

For model development, the dataset was divided into train and test sets using an 80/20 split. A baseline model was first created using a naïve forecasting approach, where the last observed value was carried forward. This served as a benchmark for comparison. Subsequently, more advanced forecasting models were applied:

- ARIMA (AutoRegressive Integrated Moving Average): Used with different parameter orders determined through ACF and PACF plots and grid search.
- Exponential Smoothing (Holt-Winters)**:** Implemented with an additive trend to capture seasonality.

- Prophet (by Facebook): Used to capture daily and weekly seasonality patterns, providing flexibility for handling missing data and outliers.

These models were trained and tuned to find the best-performing method for household electricity consumption forecasting.

3.2.5 Model Evaluation

To evaluate model performance, error metrics such as Root Mean Squared Error (RMSE)**,** Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) were used. In addition, cross-validation with 5 folds using the TimeSeriesSplit method was carried out to test the robustness of the models. The comparison revealed that Prophet outperformed ARIMA, ETS, and the Naïve baseline, providing more accurate and stable predictions across different evaluation metrics.

3.2.6 Forecasting

Finally, the Prophet model was selected as the best predictive model for household electricity consumption forecasting. It was used to generate 7-day ahead forecasts at an hourly resolution, along with uncertainty intervals. The forecasted results were saved in CSV format and visualized through plots, which clearly demonstrated the expected consumption trends and seasonal effects in household electricity usage. These forecasts can be used to plan energy consumption, manage household costs, and aid policymakers in demand-side energy management.

Chapter 4: Results and Analysis

This chapter presents the results obtained from the time series forecasting model developed using Prophet and deployed through a Streamlit-based web application. The model was trained on household electric power consumption data and used to forecast future consumption patterns.

4.1 Forecasting Results

The final forecasting model selected was Facebook Prophet, as it demonstrated better performance compared to ARIMA, ETS, and baseline naïve models during evaluation. The model captured both daily and weekly seasonality patterns effectively.

- A 7-day ahead forecast (with hourly resolution) was generated.
- The forecast plot showed close alignment between actual consumption values and predicted values.
- Prophet also provided uncertainty intervals (upper and lower bounds) that highlighted the confidence of predictions.

These results indicate that the model can reasonably capture the electricity consumption trends, which can be extended to short-term demand forecasting applications.

4.2 Residual Analysis

Residuals were calculated as the difference between the actual observed values and the model's predictions.

- The residual plot revealed that most residuals were centered around zero, suggesting the model's predictions were unbiased.
- A few deviations were observed, likely due to irregular consumption spikes not captured by the seasonality trends.
- The residuals did not show a strong pattern, confirming that the model assumptions hold and no major systematic error exists.

4.3 Evaluation Metrics

The model performance was evaluated using RMSE, MAE, and MAPE:

- Prophet outperformed ARIMA and ETS with lower error values.
- Time Series Split cross-validation confirmed the stability of the model's accuracy across multiple folds.

Thus, the Prophet model was considered robust and reliable for this dataset.

Chapter 5: Conclusion

This study focused on forecasting household electric power consumption using various time series models, with the final deployment carried out using the Prophet model integrated into a Streamlit web application for real-time visualization. The results showed that Prophet was able to effectively capture both the trend and seasonality present in the data, making it suitable for short-term demand forecasting. The predicted values closely matched the actual consumption patterns and were accompanied by reasonable confidence intervals, which added reliability to the forecasts. Furthermore, residual analysis demonstrated that the model's predictions were unbiased, with errors distributed randomly rather than showing systematic deviations. Finally, the developed Streamlit application provided an interactive and user-friendly platform that allows users to generate forecasts, visualize results, and analyze residuals efficiently, thereby showcasing the practical applicability of the proposed system.

Future Scope

The future scope of this project suggests several directions for improvement and practical application. Incorporating external features such as temperature, holidays, and demographic information could significantly enhance the accuracy of the forecasts by accounting for factors that directly influence power consumption. Similarly, extending the model to predict longer forecast horizons, such as monthly or annual consumption, would provide valuable insights for strategic planning and energy management. Additionally, integrating the model with real-time data streams could make the application more dynamic and useful for household monitoring or smart-grid systems, enabling timely decision-making and optimization of resources. Overall, this project highlights the potential of combining machine learning and time series forecasting techniques for energy demand prediction, which can play a crucial role in sustainable energy management and policy development.