Prediction of Prosumers Energy Consumption Patterns

Minor Project (AIML- 451)

योगः कर्मसु कौशलम्
IN PURSUIT OF PERFECTION

Under the supervision of

Dr. Monika Bansal
&
Dr. Bhawna Rawat

by

Krishna Goel, 02717711621
Anant Chauhan, 01517711621
Akanksha, 08017711621

School of Engineering & Technology
Vivekananda Institute of Professional Studies – Technical Campus
Pitampura, Delhi - 110034

Table of Contents

Introduction

The growing adoption of renewable energy sources, especially solar power, has led to the rise of energy prosumers—individuals who both produce and consume energy. This shift in the energy landscape presents new challenges for predicting energy production and consumption patterns, which are influenced by various factors such as weather conditions and energy prices.
Accurately predicting these patterns is crucial for minimizing imbalance costs in energy grids and ensuring optimal energy utilization.

The global energy landscape is shifting towards decentralization, largely due to the increasing adoption of renewable energy sources like solar and wind power. This shift has led to the rise of "prosumers" — entities that not only consume energy but also produce it using installed renewable energy systems, such as solar panels. However, this prosumer behavior introduces significant variability in energy production and consumption, which poses challenges for energy grid operators in maintaining balance and minimizing associated costs. Predicting energy consumption and production accurately is crucial for grid stability, energy cost reduction, and optimizing the energy trade between prosumers and the grid.

Energy prosumers are individuals, businesses, or organizations that both consume and produce energy. This concept represents a shift from the traditional model where consumers simply purchase energy from utilities and rely on centralized power generation sources. Energy prosumers are actively involved in the energy ecosystem by generating their own electricity, typically through renewable energy sources like solar panels (or wind turbines, small-scale hydropower etc.). They also consume energy from the grid when their own generation is insufficient to meet their needs

This project aims to develop a forecasting model that can predict the energy behavior of Estonian prosumers who have installed solar panels. The availability of large datasets, including weather conditions, photovoltaic capacities, and energy prices, provides a solid foundation for building an accurate forecasting model.

Previous studies and global forecasting competitions, such as the Global Energy Forecasting Competition 2012 and the European Energy Markets Competitions, have provided valuable insights into the importance of using time-series data and advanced machine learning algorithms in energy prediction tasks . These competitions demonstrated the relevance of using weather data and time-dependent features (such as seasonality) to improve forecasting accuracy.

Problem Statement & Feasibility Study

Problem Statement

The problem this project addresses is the challenge of accurately predicting the energy behavior of prosumers—those who produce energy via solar panels while consuming energy from the grid. The unpredictable nature of weather conditions, combined with fluctuating energy prices, makes it difficult to balance energy supply and demand. Traditional methods struggle to accurately predict hourly energy production and consumption, leading to increased imbalance costs. The goal is to use machine learning techniques to create models that can predict these energy patterns with higher accuracy.

The primary objective of this project is to predict the energy consumption and production of prosumers in Estonia who have installed solar panels. By analyzing the available data — including weather patterns, energy prices, and photovoltaic capacities — the project aims to develop a robust forecasting model that can accurately predict energy patterns at hourly intervals. The key challenge lies in handling the inherent variability in renewable energy production due to factors like weather changes and the non-linear nature of prosumer behavior.

Feasibility Study

- **Data Availability:** The dataset provided by the competition is rich in features, including weather data, energy prices, and photovoltaic capacity, making it feasible to build a predictive model. The data is time-stamped on an hourly basis, which is ideal for time-series forecasting.
- **Machine Learning Tools:** The availability of libraries such as XGBoost, Scikit-learn, and Pandas in Python makes the task of building and testing machine learning models feasible. XGBoost, in particular, is well-suited for this task because of its speed, accuracy, and handling of time-series data with missing values.
- **Computational Resources:** Modern personal computers with at least 16GB of RAM and a multi-core processor should be sufficient to handle the dataset and run the model. For faster computation and hyperparameter tuning, *GPU acceleration* can be leveraged.
- **Feasibility of Implementation:** Given the availability of well-documented libraries and tools, combined with the structured dataset, this project is practically feasible. The primary challenge will be tuning the machine learning model to achieve optimal accuracy, but this is manageable through iterative testing and cross-validation
- A feasibility study confirms that similar projects using weather and energy data have led to significant improvements in grid management and energy cost reduction. By predicting energy usage and production, grid operators can better plan for energy demand and supply imbalances, resulting in more efficient energy distribution. Given the rich dataset and the availability of machine learning tools, this project is both feasible and impactful.

Hardware and Software Requirements

Hardware Requirements:

- **Computer:** A computer with a modern processor (Intel i5 or equivalent) and at least 8GB of RAM to run SQL and Power BI efficiently.
- **Storage:** Adequate storage capacity (minimum 256GB SSD) to handle temporary files and project data.

Software requirements:

- **Programming Language:**
  - **Python 3.x:** Python will be the primary language used for data analysis, model development, and evaluation.
- **Libraries and Dependencies:**
  - **XGBoost:** The primary library for model implementation
  - **Pandas and NumPy:** For data manipulation and handling
  - **Scikit-learn:** For data preprocessing and model evaluation
  - **Matplotlib and Seaborn:** For data visualization and feature analysis
  - **Jupyter Notebook/Google Colab/Kaggle Notebooks:** For interactive development and testing.
  - **Kaggle API:** For data access and leaderboard submission.
- **Development Environment**:
  - **Anaconda (optional):** A Python distribution with pre-installed libraries
  - **Kaggle Notebooks or Google Colab:** for cloud-based computation, eliminating the need for high-end local hardware

References

1. *Hong, T., Pinson, P., & Fan, S. (2016).* "Global Energy Forecasting Competition 2012 and beyond." International Journal of Forecasting, 32(3), 596-608. DOI: [10.1016/j.ijforecast.2015.11.014](https://doi.org/10.1016/j.ijforecast.2015.11.014)

2. *Fan, S., & Hyndman, R. J. (2012).* "Short-term load forecasting based on a semi-parametric additive model." IEEE Transactions on Power Systems, 27(1), 134-141. DOI: [10.1109/TPWRS.2011.2162082](https://doi.org/10.1109/TPWRS.2011.2162082)

3. *Chen, T., & Guestrin, C. (2016).* "XGBoost: A scalable tree boosting system." In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge and Data Mining, pp. 785-794.DOI: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)

4. *Li, L., Yu, J., & Yang, Z. (2015).* "Renewable energy forecasting based on time-series method and machine learning." Renewable and Sustainable Energy Reviews, 52, 273-284. DOI: [10.1016/j.rser.2015.07.078](https://doi.org/10.1016/j.rser.2015.07.078)