Master's Thesis Project Plan

Electricity Consumption Forecasting Using High-Resolution Customer Data and External Data Sources

Tyra Wodén (021215-7440)

id21twn@cs.umu.se

January 26, 2026
Semester: VT 26
External partner: Umeå Energi

1 Introduction

In recent years, the electricity market has been affected by large variations in price and periodically very high electricity prices. This has increased the need for customers to understand and influence their electricity consumption and future electricity costs. To address this need, Umeå Energi has worked on developing and improving their interfaces for customers, for example by developing an application for private customers. Customers wish to, before the invoice arrives, receive an estimation of the electricity cost. A first step towards a cost forecast is forecasting the electricity consumption. Currently, Umeå Energi has developed a linear regression model to forecast the consumption for the current month. The model is based on historical consumption and temperature, aggregated by month. For certain customer groups, Umeå Energi has access to high definition measurement data where electricity consumption is registered on a 15-minute level, which creates opportunities for more advanced modeling. There may also be other factors, apart from historical consumption, that affect the electricity consumption, for example weather data, calendar information (holidays etc.), and energy prices.

The goal with this Master's thesis is to develop and evaluate forecasting models that show the predicted energy consumption for the current month. Previous research has identified several models that are able to accurately predict energy consumption for different time intervals, utilizing various combinations of data sources [1, 2]. However, quantifying and explaining the contribution of each feature is an underexplored area [citation needed]. Explainable AI (XAI) will be used to evaluate the contributions of different

variables to the forecast. The forecast will be updated continuously during the month, and as actual observed values of electricity consumption become available, these values will be used to improve the forecast.

2 Problem Formulation

This project aims solve the problem of accurately predicting customers' monthly electricity consumption based on consumption data and external data sources. The main focus of the project will be to quantify and explain the contribution of each variable in the forecast. Several research questions are posed in the following priority order:

1. Which external sources of data contribute to the accuracy of the forecast, and which sources provide little to no improvement?
2. How can the accuracy of the forecasts be validated over time after the models are implemented in production?
3. What type of forecasting model gives the best results for predicting electricity usage of the current month?

3 Method

To answer the research questions, the following main steps will be performed:

- **Literature study**: Review existing work in the field of electricity consumption forecasting, in order to gain valuable theoretical background and insights on suitable statistical/ML models.
- **Data preparation**: Retrieve 15-minute level electricity consumption data, as well as relevant external data sources (such as weather data, calendar data, and electricity prices).
- **Primary model implementation**: Choose an ML/DL model (e.g. ANN, SVM, LSTM) that will be the main focus of the thesis, and implement a forecasting model based on the consumption data and the external sources.
- **Evaluation of data sources**: Analyze and evaluate how the different data sources contribute to the accuracy of the forecasts using Explainable AI. The goal is to identify data sources that add significant value, and sources that are not able to motivate the increased complexity.
- **Continuous forecast updates**: Develop a method for improving the forecasts during the current month by continuously replacing predicted values with actual observations.

- **Integration perspectives**: Develop a prototype that demonstrates how the forecasting models can be integrated into Umeå Energi's existing environments (Microsoft SQL Server). The prototype includes reading of input data from database tables, execution of forecasting models in Python, and storing of results.
- **Quality assurance**: Develop a suggestion on how to continuously follow up the accuracy of the models after implementing them in production.

In case there is time, more than one model will be implemented (or possibly ensemble learning [3]) in order to allow more extensive comparisons or accuracy improvements. Reviewing of literature will be used to identify relevant models.

3.1 Evaluation Methods

Evaluation will mainly be performed on the data sources' contributions, and in case of time, the different models' performance. As a baseline comparison, an existing linear regression electricity consumption model (based on historical monthly consumption and temperature data) will be utilized. XAI methods such as feature importance and Shapley values will be used to evaluate the contributions of the different data sources. Different combinations of data sources may also be evaluated.

4 Literature

Two comprehensive reviews on electricity consumption forecasting have been identified:

- An extensive review and comparison of both statistical and ML/DL techniques for forecasting is found in [1]. They also review combinations of different techniques, i.e. hybrid models. Claims that ANN has more advantages than statistical models, and has better performance for nonlinear problems. Highlights that hybrid models can be beneficial to capture complexities in building energy and operational data.
- Another review of statistical, AI, and hybrid methods for forecasting is found in [2]. They also highlight the strength in AI models for dealing with nonlinear patterns. Claims that hybrids between AI and Swarm Intelligence (SI) methods show potential for increased accuracy. Provides a clear overview of different studies regarding prediction time intervals, included features, building types etc.

Papers where experiments have been performed:

- Support Vector Machine (SVM) for forecasting energy consumption [4].
- Monthly electricity consumption forecasting based on decomposition methods and ARIMA: [5].
- Forecasting cooling energy using ANN (for three university buildings, weekly/monthly): [6].
- Forecasting high voltage consumers' electricity consumption using LSTM, GRU, TCN: [7].

Papers specifically on XAI and electricity forecasting:

-

5 Implementation

The implementation of the project consists of models for predicting electricity consumption, as well as integration of resulting model predictions in the existing database. The models will predict based on electricity consumption data from Umeå Energi's database, idk. Part of the implementation is also to use Explainable AI methods to evaluate the contributions of each variable to the prediction.

6 Work Structure and Time Plan

The work structure of the project will follow agile principles, in order to allow flexibility and adjustments of plans as the project develops. Additionally, the implementation may happen iteratively; first implementing one model and evaluating it, and then implementing a second model...

7 Risk Analysis

- Risk 1: x

8 Supervision

8.1 UmU Supervision

Supervision with my internal supervisor will be arranged as needed, roughly once a week or once every two weeks. Email will be used to book time slots, and meetings can happen either in person on campus or digitally via zoom.

Table 1: Contact information for internal supervisor.

Name	Esteban Guerrero Rosero
Email	esteban.guerrero@umu.se

8.2 External Supervision

I will be spending most of my time at the Umeå Energi office, so arranging supervision with my external supervisor will be easy. If needed, I am also able to get help from other team members at Umeå Energi.

Table 2: Contact information for external supervisor.

Name	Elin Eriksson
Email	elin.eriksson@umeaenergi.se

References

[1] Deb, C., Zhang, F., Yang, J., Lee, S.E., Shah, K.W.: A review on time series forecasting techniques for building energy consumption. Renewable and Sustainable Energy Reviews **74** (July 2017) 902–924

[2] Mat Daut, M.A., Hassan, M.Y., Abdullah, H., Rahman, H.A., Abdullah, M.P., Hussin, F.: Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review. Renewable and Sustainable Energy Reviews **70** (April 2017) 1108–1118

[3] Sagi, O., Rokach, L.: Ensemble learning: A survey. WIREs Data Mining and Knowledge Discovery **8**(4) (2018) e1249

[4] Dong, B., Cao, C., Lee, S.E.: Applying support vector machines to predict building energy consumption in tropical region. Energy and Buildings **37**(5) (May 2005) 545–553

[5] Sun, T., Zhang, T., Teng, Y., Chen, Z., Fang, J.: Monthly Electricity Consumption Forecasting Method Based on X12 and STL Decomposition Model in an Integrated Energy System. Mathematical Problems in Engineering **2019**(1) (2019) 9012543 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1155/2019/9012543.

[6] Deb, C., Eang, L.S., Yang, J., Santamouris, M.: Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks. Energy and Buildings **121** (June 2016) 284–297

[7] Hadjout, D., Torres, J., Troncoso, A., Sebaa, A., Martínez-Álvarez, F.: Electricity consumption forecasting based on ensemble deep learning with application to the Algerian market. Energy **243** (March 2022) 123060