A Data Science Approach to Forecast Electricity Consumption in Australia

Svetlana Temirov (z5472623),

Yuanjie Zhang (z5125136),

Zhengda Zhong (z5158361),

Satya Mavuri (z5405935).

19/03/2024

Abstract

Our team believes that there is enough variance in electricity demand, that forecasting demand on a short-term basis (24 hours) is necessary to capture the behaviour of electricity consumption. Our plan is to execute a forecasting model for energy demand based on the short-term metrics of cultural and societal nature. Particularly, school, and public holidays, weekends, and popular vacation times. In combination with outside temperature, which has an established correlation with energy demand. We propose to use advanced machine learning models such as ARIMA, SARIMA, Neural Networks and Gradient Boosting and compare their performance to determine the preferred model.

Table of Contents

1 Introduction and Motivation

If you stop and think for a minute, you will see that almost every one of our regular daily activities depends on electric power. Access to a constant flow of electricity is a staple of our modern society. The task of supplying just the right amount, is one that hinges on the ability to accurately forecast the expected demand. We were provided with data on actual and predicted energy demand figures across Australia, as well as the data and information on the correlation between the outside temperature and energy demand courtesy of Endgame Economics.

Short term forecasting is a crucial aspect of the industry, where accurate forecasting can save millions [1]. To facilitate more accurate modelling, our team wants to explore the human element of electricity demand and see if social parameters such as holidays, weekends, school vacations and popular dates for people to take time off work have an effect on the demand. And if these parameters complement the forecast done with temperature alone.

We aim to discover patterns in electricity demands based on these human elements to provide a better forecasting tool for future electricity demand. Moreover, we are also motivated to delve into the patterns these metrics can highlight about our societal schedules and elctricity usage.

2 Brief Literature Review

Electrical power demand is influenced by a range of factors including weather conditions, socioeconomic activities, calender effect, urbanisation and demand with surge of technologies (such as electric vehicles). Energy use, and hence, its demand is linked directly to social routines, habits and cultural activities such as working, Weekend and public holidays [2], [3], [4], [5].

Close match of Energy demand patterns with socio-cultural boundaries was demonstrated by Román MO, Stokes EC. [2]. Long public holidays like Christmas, New years and Holy month of Ramadan were tracked and further analysed by multiple linear regression modeling demonstrated demand the patterns. Another relevant study on electricity use specific to Christmas day showed a reduced demand on the Christmas and similar public holiday periods like Ester [6].

These findings lead us to believe that there is merit in our motivation to further investigate and establish a data link between societal and cultural elements and electricity demand. A case study in Germanyhas looked large data set with emphasis on the specific contribution of of Public holidays, weekends on the electricity demand [3]. Various univariate and multivariate models were used to predict the effect of holidays in Germany's electrical demand forecasting. This study has found that incorporating holiday effects into the forecasting model has raised the forecasting accuracy substantially, improving the estimates on public holidays by 80% and reduced the error to approximately 10%.

3 Methods, Software and Data Description

Our approach to short-term forecasting will expand on the dataset provided by Endgame Economics. We will generate data on Australian dates of school and public holidays as well as obtain data on dates of weekends vs. weekdays, and on times of the year most Australian residents take time off work. We will need to gather this data over the period of 2010 to 2021, as this is the span of the existing dataset. This data will need to be cleaned and processed and then combined with the existing data from Engame Economics

The data pre-processing and will be primarily done in Python, utilising the pandas library. Additionally, for visualizing the data and results of our analyses, the Matplotlib and Seaborn libraries. For final reporting and presentation quality visualisations we aim to use Tableau.

3.0.1 Forecasting Model

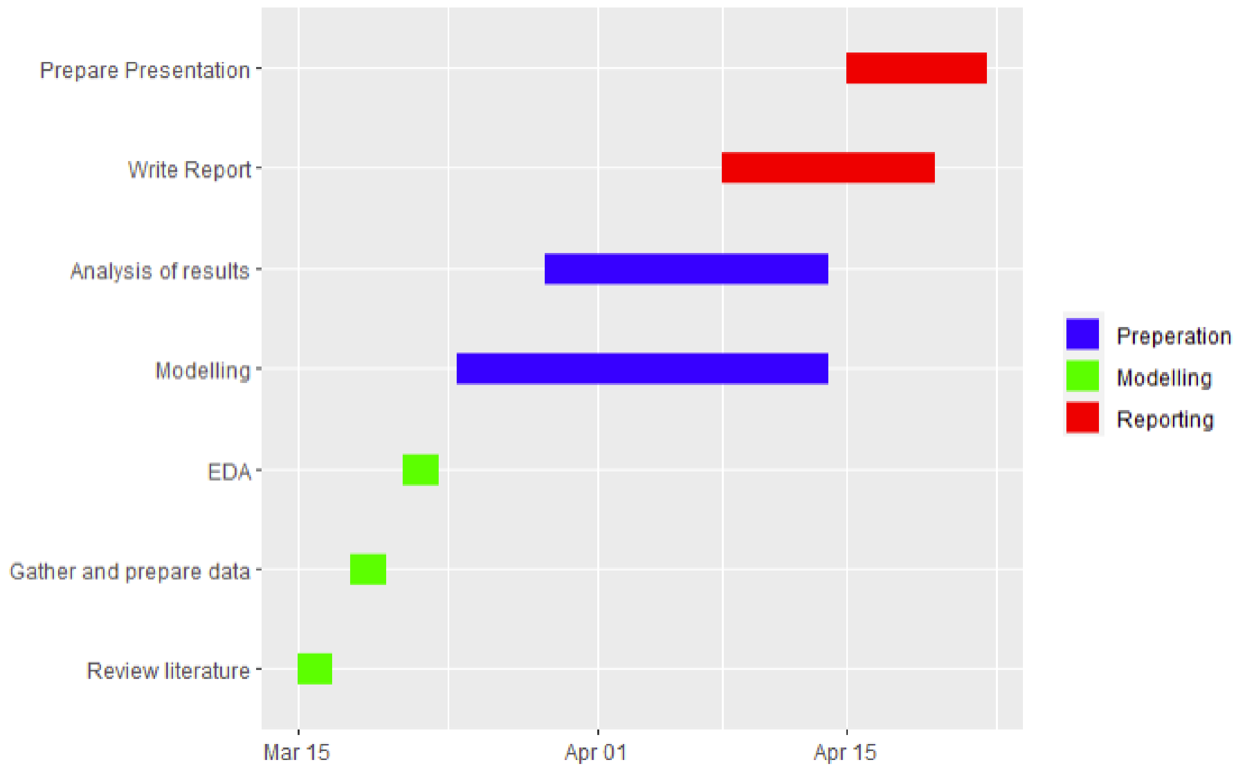
With a clean and combined dataset, we will employ a combination of time-series analysis techniques and machine learning models. Techniques such as ARIMA (Auto Regressive Integrated Moving Average) and SARIMA (Seasonal ARIMA) will be utilized for their ability to model and forecast time-series data effectively. Additionally, Neural Network and Gradient Boosting Machines will be explored for their capacity to handle complex interactions between multiple variables and their robustness in forecasting.

The literature review conducted highlighted that similar research into electricity demand predictions utilised similar models [5], [1]. Therefore, we believe they will be a suitable choice for models to explore.

To build our models, we will primarily use Python and take advantage of the following libraries: For the time-series analysis involving ARIMA and SARIMA models, we plan to use statsmodels. For the machine learning components, including Neural Networks and Gradient Boosting Machines, we will primarily use the scikit-learn library. In addition, for more complex neural network architectures, we will utilize the TensorFlow and Keras libraries.

4 Activities and Schedule

The chart on the next page shows the Gantt chart of the key milestones for our project. Within each milestone, sub-tasks will be allocated and divided among team members. And all Members will participate in each Milestone task.



We divided the project into 3 stages. The preparation stage where we conduct literature review, clean and organise the data and formally define and fine tune our research question. The second stage is our modelling stage where we build and tune the machine learning models and analyse our results. This is our longest stage as we expect it to be iterative with the first pass of the models giving us insights into how to tune them further. The last stage of the project is for formal reporting and presenting our results in an engaging manner to a non-technical audience.

The list below shows tasks for which team members would like to take more major roles. All other team members will be able to assist, add and change tasks as appropriate with the progress of the project.

- Svetlana Temirov (z5472623) - literature review, modelling the ARIMA and SARIMA modesl and will conduct content and cohesion editing to written and visual deliverables. Point person for submitting deliverables.
- Yuanjie Zhang (z5125136) - literature review, Neural Network and Gradient Boosting modelling, testing the combination of both models
- Zhengda Zhong (z5158361) - literature review, Neural Network and Gradient Boosting modelling, testing the combination of both models
- Satya Mavuri (z5405935) - Detailed literature reivew, data exploratory analysis, data cleaning and organisation, assist with the ARIMA and SARIMA models and special visualisations needed for deliveables.

References

1. Al-Musaylh MS, Deo RC, Adamowski JF, Li Y. Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in queensland, australia. Advanced Engineering Informatics. 2018;35:1–16.

2. Román MO, Stokes EC. Holidays in lights: Tracking cultural patterns in demand for energy services. Earth’s Future. 2015;3:182–205.

3. ZIEL F. Modeling public holidays in load forecasting: A german case study. Journal of Modern Power Systems and Clean Energy. 2018;6:191–207.

4. Ben Taieb S, Hyndman RJ. A gradient boosting approach to the kaggle load forecasting competition. International Journal of Forecasting. 2014;30:382–94.

5. Pesantez JE, Li B, Lee C, Zhao Z, Butala M, Stillwell AS. A comparison study of predictive models for electricity demand in a diverse urban environment. Energy. 2023;283:129142.

6. Hugh Saddler. What can we learn from looking at electricity use on christmas day. 2013.