

ASSESSMENT 1 PART A
GROUP PROJECT PLAN BY TEAM L

A DATA SCIENCE APPROACH TO FORECAST
ELECTRICITY CONSUMPTION IN AUSTRALIA

Mark Budrewicz (z5353932)

Hao Chiou (z5131909)

Charlet Jeyapaul (z5375906)

Louis Long (z5162359)

School of Mathematics and Statistics
UNSW Sydney

September 2023

Table of Contents

Table of Contents	2
1. Introduction and Motivation.....	3
2. Brief Literature Review.....	3
3. Methods, Software and Data Description	4
4. Activities and Schedule	5
References.....	6

1. Introduction and Motivation

Motivation

Electricity demand forecasting is a topic that is within the interest of government and market bodies, with the consensus of accurate forecasts of energy demand driving profitability. Electricity in Australia is generated primarily using oil, coal, and gas as fuel (Bhattacharya, Inekwe & Sadorsky 2020). Also, with oil prices increasing, this indicates that the cost to generate energy will also increase (Kpodar & Liu 2022). Therefore, profit maximization is achieved when energy production is minimized to be near actual energy demand. Thus, if electricity demand can be modelled accurately, government and market bodies can then utilize such information to tailor policies as well as optimize pricing.

Introduction

The aim of this project is to provide an electricity generator client with a rolling 1 day forward demand forecast for the NSW grid at 30-minute intervals to help minimise operational costs, giving a competitive edge in pricing which will benefit its profitability and flow down to lower costs for end consumers.

In this scenario, it is important for an electricity supplier to be able to estimate how much electricity is required by the grid so that it can strategically plan generation output to minimise its startup and shut down costs and maximise profit. Times of high load requirement in the grid inevitable lead to higher spot prices, so it is in the interests of an electricity supplier to be able to accurately forecast and capitalise on such situations, inline with its operational constraints (startup and shutdown times, maintenance requirements). While an electricity demand forecast is already generated by the Australian Energy Market Operator (AEMO) (AEMO 2023) for participants in the grid, this analysis aims to create a forecasting model that will outperform AEMO's demand forecast, allowing a supplier to take advantage of potential misalignment in AEMO's forecast to actual load demand. The model will leverage novel machine learning techniques such as Support Vector Machines to capture the complex non linear relationships in the electricity demand time series to forecast rolling 1 day forward NSW electricity demand at half-hourly intervals.

In this analysis, the key variables of interest will be temperature, historical energy demand, historical energy forecast, seasonality, time of day, and meteorology data such as weather condition and humidity (Liu et al. 2021). Transformations will be made to the variables to ensure data used will be fit for purpose. Various time series and machine learning models will be explored, while taking consideration of analyses and models that have been completed in the past. The models will then be further refined via training with simulated data. The final model will be selected after comparison of various generated models and features, specifically through consideration of feature selection, accuracy, and interpretability and performance relative to AEMO's forecast.

2. Brief Literature Review

Initial review of the literature on electricity demand forecasting pointed to a well understood relationship between temperature and electricity demand that has already been incorporated by the AEMO's forecasting procedures (AEMO 2023). This relationship has been widely explored with a specific focus on higher frequency time intervals. McCulloch & Ignatieva 2020 showed a strong relationship when modelling Australian 5-minute electricity demand with intraday temperature using a Generalised Additive Model. This same research also indicated that sensitivity of demand to temperature was heavily dependent on the time of day with higher demand sensitivity during hours of high human activity. In Zhang & Guo 2020, their support vector machine (SVM) model found other meteorological features such as wind speed, relative humidity, precipitation and air pressure impacted electricity demand. Interestingly the behaviour of these variables on demand was non linear, which the authors hypothesised could be due to their impact on relative temperature (ie high wind can act to help cool on warmer days, but make it feel colder on cooler days). Other literature identified other forms of seasonality and cyclicity (from daily to monthly) in the electricity demand time series that needed to be factored into the modelling. One approach by Jiang et al 2020 was to remove the seasonality as a preprocessing step to not affect the SVM algorithm, while other approaches such as Alonso et al 2020, chose to let their Long Short-Term Memory (LSTM) recurrent neural network to capture this idiosyncrasy naturally.

As the complexity of electricity demand has evolved so has the research into techniques for its forecasting. Earlier models exploring Box and Jenkins time series modelling (Hagan & Behr 1987), followed on with adaptive autoregressive moving-average (ARMA) modelling of day and week look ahead forecasts showing strong results using simpler methods (Chen et al 1995). However, the research has shifted to techniques that can model the complex nonlinear relationship of electricity demand to temperature and seasonality. One of the most recent papers in the literature, Bashari & Rahimi-Kian 2020, forecasts 24 hour ahead electricity load for the city of Toronto by combining a LSTM network using historical electrical demand with a deep Feedforward Neural Network (FNN) using forecasted meteorological inputs (temperature). The combination of the two artificial neural networks outperformed each individual network as measured by Root Mean Square Error (RMSE). Jiang et al 2020 utilised SVM algorithms with 30min time series electricity demand data to generate a forecast for both NSW and Singapore electricity demand, performing better than their benchmarking models.

Following from the success in the literature above, we look to explore the forecasting capability of meteorological features such temperature, wind, humidity, precipitation and cloud cover on NSW short term electricity demand which has been shown to have predicative capability in other grids. SVM algorithms will be used to help model the non linear relationship, with an investigation into how to manage the seasonality inherent in the data.

3. Methods, Software and Data Description

Method	Software	Potential research data
<ul style="list-style-type: none"> • Data preparation: A few approaches will be considered depending on the model, including splitting into training/holdout data sets, and bootstrapping • Hyperparameter tuning: Specifically for machine learning models, tuning specific parameters to find the most appropriate model will be integral • Evaluation criteria: A variety of metrics will be considered when assessing the effectiveness of models, including performance metrics such as accuracy and MSE, as well as quantitative metrics such as complexity and interpretability 	<ul style="list-style-type: none"> • Coding: We have decided to use RStudio due to its suitability in the project through its capability in data manipulation, statistical analysis, modelling and visualisation • Communication: MS Teams will be our main channel for communication due to its ease of use and familiarity within the team, as well as Microsoft integration • Planning: Team environments require organisation as there are many moving parts involved. Therefore, we will use MS Planner to track current and future action items • Central storage: The team has chosen OneDrive to be our central file storage application, due to its seamless integration with MS Teams • GitHub: This is an essential application to enable collaboration between programmers through version control. 	<ul style="list-style-type: none"> • Historical demand data: Historical electricity consumption data will be essential as a baseline to forecast future demand • Seasonality/time of day: Literature review has highlighted the importance of seasonality as well as time of day/week and its relation to energy demand • Weather: Temperature is the greatest indicator of energy demand and will therefore be a factor in our models. In addition, weather data such as weather conditions (e.g., cloudy vs. sunny) and humidity are potential indicators to forecast demand as well.

4. Activities and Schedule

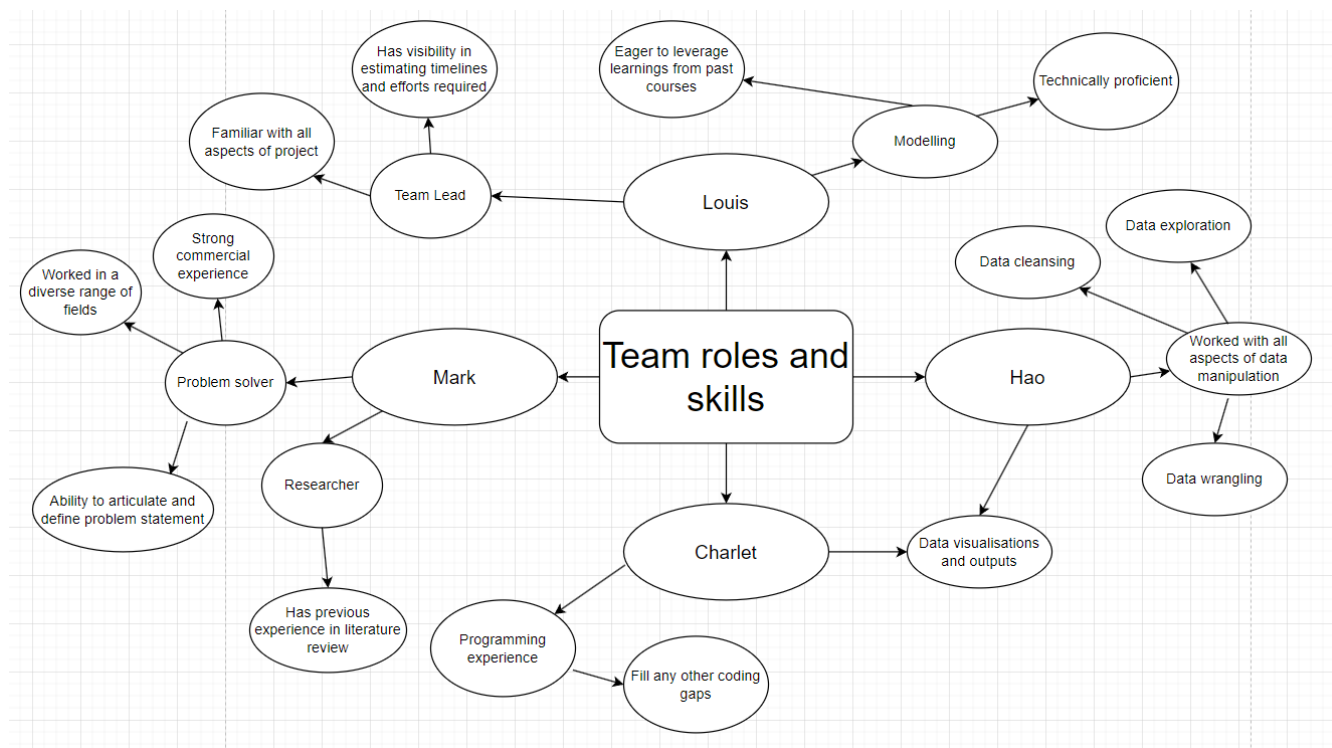
Activities

Planning and discovery	Technical development	Create outputs
<ul style="list-style-type: none"> Define key problem statement Align on tools and applications Literature review Data exploration 	<ul style="list-style-type: none"> Data preparation Modelling Data visualisation and outputs 	<ul style="list-style-type: none"> Create report Create and rehearse presentation

Schedule

Activity	Task Owner	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7
Establish comms	All							
Choose applications/tools	All							
Write Group Project Plan	All							
Research literature	Mark							
Data exploration + prep	Hao							
Modelling	Louis							
Visualisation + outputs	Charlet							
Create report	Louis + Hao							
Create presentation	Mark							
Rehearse	All							
Group video presentation	All							

Team roles



References

- AEMO 2023, 'Forecasting Approach Electricity Demand Forecasting Methodology - August 2023', Australian Energy Market Operator, Version 1.3, accessed 3 September 2023 <https://aemo.com.au/-/media/files/electricity/nem/planning_and_forecasting/nem_esoo/2023/forecasting-approach_electricity-demand-forecasting-methodology_final.pdf?la=en>
- Alonso, AM, Nogales, FJ & Ruiz, C 2020, 'A single scalable LSTM model for short-term forecasting of massive electricity time series', *Energies (Basel)*, vol. 13, no. 20, p. 5328, accessed 3 September 2023, <https://primoa.library.unsw.edu.au/permalink/f/11jha62/TN_cdi_doaj_primary_oai_doaj_org_article_81185b65a0a94c6c84b07f7d8daf5284>
- Bashari, M & Rahimi-Kian, A 2020, 'Forecasting Electric Load by Aggregating Meteorological and History-based Deep Learning Modules', in 2020 IEEE Power & Energy Society General Meeting (PESGM), IEEE, pp. 1–5, accessed 3 September 2023, <https://primoa.library.unsw.edu.au/primo-explore/openurl?rft_id=info:doi%2F10.1109%2FPESGM41954.2020.9282124&vid=UNSW_SERVICES_PAG&institution=61UNSW_INST&url_ctx_val=&url_ctx_fmt=null&isServicesPage=true>
- Bhattacharya, M. Inekwe JN. Sadowsky, P 2020, 'Convergence of energy productivity in Australian states and territories: Determinants and forecasts', *Energy Economics*, Volume 85, no. 104538, accessed 3 September 2023, <<https://www.sciencedirect.com/science/article/pii/S0140988319303330>>
- Chen, J-F, Wang, W-M & Huang, C-M 1995, 'Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting', *Electric power systems research*, vol. 34, no. 3, pp. 187–196. accessed 3 September 2023, <https://primoa.library.unsw.edu.au/permalink/f/11jha62/TN_cdi_crossref_primary_10_1016_0378_7796_95_00977_1>
- Hagan, MT & Behr, SM 1987, 'The Time Series Approach to Short Term Load Forecasting', *IEEE transactions on power systems*, vol. 2, no. 3, pp. 785–791, accessed 3 September 2023 <https://primoa.library.unsw.edu.au/permalink/f/11jha62/TN_cdi_pascalfrancis_primary_7630203>
- Jiang, P, Li, R, Liu, N & Gao, Y 2020, 'A novel composite electricity demand forecasting framework by data processing and optimized support vector machine', *Applied energy*, vol. 260, p. 114243, accessed 4 September 2023, <https://primoa.library.unsw.edu.au/permalink/f/11jha62/TN_cdi_crossref_primary_10_1016_j_apenergy_2019_114243>
- Jiang, P, Li, R, Lu, H & Zhang, X 2020, 'Modeling of electricity demand forecast for power system', *Neural computing & applications*, vol. 32, no. 11, pp. 6857–6875, accessed 3 September 2023 <https://primoa.library.unsw.edu.au/permalink/f/11jha62/TN_cdi_springer_journals_10_1007_s00521_019_04153_5>
- Kpodar, K. Liu, B 2022, 'The distributional implications of the impact of fuel price increases on inflation', *Energy Economics*, Volume 108, no. 105909, accessed 3 September 2023, <<https://www.sciencedirect.com/science/article/pii/S0140988322000895>>
- Liu, XQ. Zhang, C. Zhou, Y. Liao, H 2021, 'Temperature change and electricity consumption of the group living: A case study of college students', *Science of The Total Environment*, Volume 781, no. 146574, accessed 3 September 2023, <<https://www.sciencedirect.com/science/article/pii/S0048969721016429>>

McCulloch, J. Ignatieva, K 2020, 'Intra-day Electricity Demand and Temperature', The Energy Journal, Vol. 41, No. 3, accessed 3 September 2023, <<https://www.iaee.org/en/publications/init2.aspx?id=0>>

Zhang, G & Guo, J 2020, 'A Novel Method for Hourly Electricity Demand Forecasting', IEEE transactions on power systems, vol. 35, no. 2, pp. 1351–1363, accessed 4 September 2023, <https://primoa.library.unsw.edu.au/permalink/f/11jha62/TN_cdi_crossref_primary_10_1109_TPWRS_2019_2941277>