i

PROJECT PLAN BY GROUP E

A DATA SCIENCE APPROACH TO FORECAST ELECTRICITY CONSUMPTION IN AUSTRALIA

Abdul El-Hamawi (z5019165), Chris Strods (z5329477),
David Anderson (z5343521)
Jamie Twiss (z5353394), Shubhankar Dutta (z5304573),
Sonal Chawla (z5092985)

School of Mathematics and Statistics
UNSW Sydney

March 2022

Abstract

The ability to forecast energy demand is essential to ensuring the stability of Australia's electricity grid, and therefore to supporting the Australian economy. In the past, this has been done through a combination of simple linear modelling and operator intuition. Advances in computing make it likely that more sophisticated approaches using artificial intelligence can improve the accuracy of existing models. Building on research already conducted on the use of AI in forecasting, this project will build a new demand-forecasting model, using previous predictions, temperature, and recent demand as inputs. A variety of methodologies will be used to seek to create a new model that can generate more accurate forecasts and demand scenarios.

Contents

1. Introduction and Motivation

Accurate forecasts of energy demand are important for accuracy and price stability. Outages or price spikes can have a sharp negative economic effect, as seen in the 1973 oil crisis (Steven, 2020). This project aims to show the utility of machine learning (ML) modelling techniques in forecasting energy demand.

Point-load forecasting has recently come under criticism due to their inherent inaccuracy, in light of the chaotic nature of most standard model inputs (Hong & Fan, 2016) (Lucas Segarra, et al., 2020). Probabilistic load forecasting (PLF), calculating probability for a range of values, may be a better alternative. The analysis will consider various forecasting models aiming to find the most appropriate one.

The purpose of our investigation into energy forecasting models will aim at providing a piece of analysis that could be used by a business or market body that operates within the energy sector to add value or ease their decision-making process. To achieve the best value from our analysis we must consider the most appropriate model outputs for the end customer. The thought of providing a single point load forecast of expected energy usage per unit time is recently becoming criticized (Hong & Fan, 2016) (Lucas Segarra, et al., 2020). Point load forecasting models are inherently inaccurate due to the chaotic nature of most standard model inputs i.e., temperature and weather. A suggested method of addressing this uncertainty is by utilizing probabilistic load forecasting (PLF), an approach that provides a range of different forecasts with their prescribed probabilities. This can be seen to provide much more utility and benefit to the end clients of this project than a single point forecast model. Most current energy forecast models approaches can be split into 2 categories:(Hong & Fan, 2016) (Lucas Segarra, et al., 2020)

There are two common approaches to building energy-forecast models. The econometric approach identifies the statistical relationship between the dependent variable (energy usage) and the independent variables (in this case, forecast demand, previous demand, and temperature. The end-use approach disaggregates energy demand into social, behavioural, economic, and technological factors and analyses them separately, to determine trends and interrelationships before formalization of a structured mathematical relationship.

(Craig, Gadgil, & Koomey, 2002) shows that long-term forecasts in the United States were more than twice the actual rates due to underestimating uncertainties in underlying social or economic trends. A model that can identify and account for such uncertainties is going to be more accurate than one that does not.

2. Brief Literature Review

Simple forecasting models such as (Codoni, Park, & Ramani, 1985) examine energy demand in relation to consumer income. Investigations by (Armstrong, 2001) and (Craig, Gadgil, & Koomey, 2002) show that these simple models can be just as accurate as more complex ones. The data for these models is not difficult to acquire and in the scope of our project, the integration of simpler methods into more complex models could be useful.

They could also serve as an evaluation metric in comparison to forecast accuracy of more complicated models.

Some of the more complex energy forecasting models:

- (Xu, Wang, & Tang, 2019) use artificial neural networks and temperature data to create a PLF output, showing the best accuracies on short term models forecasting (one day ahead).
- (Vanting, Ma, & Jørgen, 2021) use neural networks to build hybrid models and consider the approach of (Vesa, et al., 2019)
- Bottom-up approaches such as (Alajmi & Phelan, 2020), modelling consumption in residential buildings in Kuwait, show promising accuracy.
- (Jing, Yaoqi, & Xiaojuan, 2018) and (Fan, Liao, Zhou, Zhou, & Ding, 2020) use the Monte Carlo method to pre-process chaotic weather data

3. Methods, Software and Data Description

As mentioned, the goal of this project is to develop an effective energy demand forecasting model. To do this we will be primarily analysing data regarding energy demand and temperature in NSW.

The data was forked from GitHub and then downloaded to individual machines. There were twelve separate files, once recombined and extracted, containing each of temperature, forecast demand, and actual demand across each of four states (NSW, VIC, QLD, and SA). The forecast demand data is the largest, by a wide measure, given that it contains multiple predictions for each DATETIME object.

There is one anomaly in the data, which is that the SA temperature file has an index column in the original .csv, which the others do not. This was removed. Otherwise, the data is extremely clean.

The data analysis will be done with, the Python programming language, Tableau, and the R programming language. Python will be used for data cleaning and modelling using various data and Machine learning libraries, such as pandas, scikit-learn, keras and others. Tableau is ideal for creating visualisations of the data and models which will instrumental in the analysis. R is will also be used for creating some visualisations as well as some data modelling and analysis. Other tools may be used if required.
Our analysis will look at finding relationships between the various data, specifically;

- How energy demand relates to temperature?
- How temperature relates to the current forecast models?
- How well current forecasting, models demand?

Various ML modelling techniques will be tested to find best one.

4. Activities and Schedule

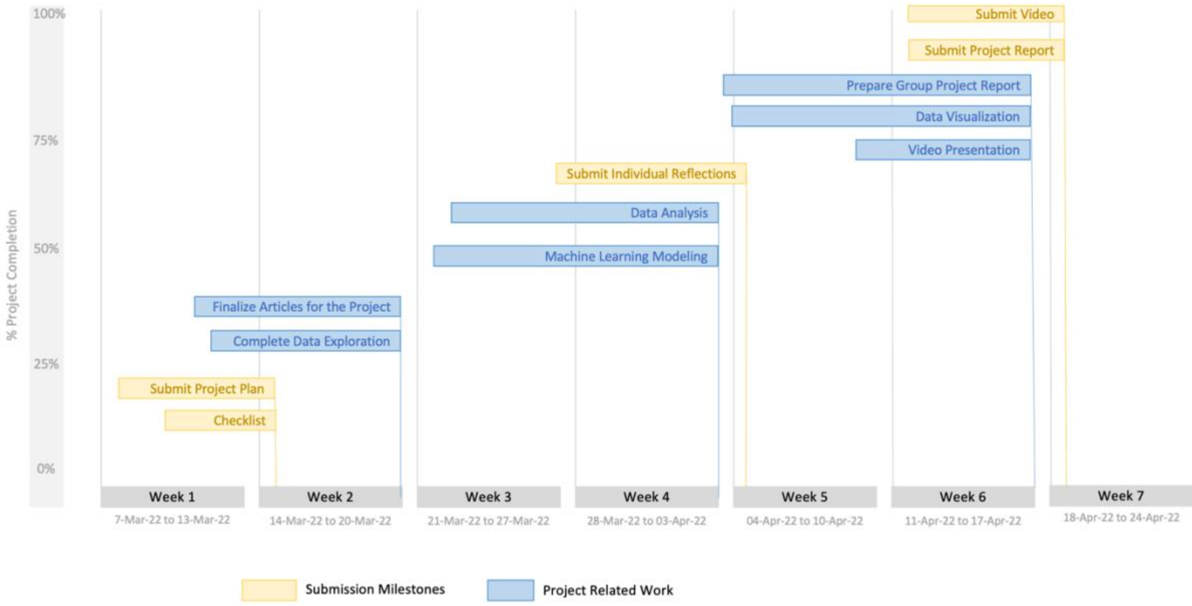
The project will be conducted over a six-week period (of which the first week and one day is already complete. The high-level project plan is shown in Appendix 1. This project plan highlights the high-level work packages and the planned milestones of the project.

The high-level work packages are then further subdivided into multiple activities, and a detailed plan is created using the high-level project plan as the baseline.

The detailed plan is set out in Appendix 1 as well. Note that this plan will keep changing as the project moves forward; it is based on our current understanding of the necessary activities and the natural sequencing within those activities, as well as estimates of how long it will take to complete them.

Appendix 1: Plan Details

High-level plan:



Detailed plan:

TASK	ASSIGNED TO	Planned in We	PROGRESS	START	END
Team Introduction	Team	Week 1	Complete	6-Mar	6-Mar
Selection of Project Tools	Team	Week 1	Complete	7-Mar	13-Mar
Set up MS Teams	David	Week 1	Complete		
Set Up Project Planner	Shuba	Week 1	Complete		
Minutes of meeting	Team	Week 1	Complete		
Team Member Roles	Team	Week 1	Complete	7-Mar	13-Mar
Project Implementation Checklist		Week 1	In Progress	7-Mar	13-Mar
Project Gantt Chart	Shuba	Week 1	In Progress	7-Mar	13-Mar
Project Plan Report Template	Shuba, Jamie	Week 1	Complete	7-Mar	13-Mar
Project Plan Report Contents	David, Abdul,Jamie	Week 1	In Progress	9-Mar	14-Mar
Set up GitHub - Forked from the UNSW Project	Shuba, Chris	Week 1	Complete	12-Mar	13-Mar
Project Plan Submission	Jamie	Week 1	In Progress	7-Mar	13-Mar
Week 2 - Meeting Agenda	Shuba , Sonal	Week 2	Complete	13-Mar	14-Mar
Data Exploration		Week 2			
Research		Week 2			
Finalize Articles for Project					
Review Data Analytics Techniques					
Finalize Statistical Tools					
Perform Dimensionality Reduction					
Develop Machine Learning Algorithms					
Data Analysis					
Machine Learning					
Structuring of Report					
Develop Visualization					
Prepare Group Project Report					
Prepare Presentation Video					

Appendix 2: References

Alajmi, T. & Phelan, P., 2020. *Modeling and Forecasting End-Use Energy Consumption for Residential Buildings in Kuwait Using a Bottom-Up Approach,* Kuwait: Energy and Building Research Center, Kuwait Institute for Scientific Research.

Armstrong, J., 2001. *Principles of forecasting : a handbook for researchers and practitioners.* Boston: Kluwer Academic.

Codoni, R., Park, H.-c. & Ramani, K. V., 1985. *Integrated energy planning : a manual.* 967-99954-5-3 ed. Kuala Lumpur: Kuala Lumpur.

Craig, P. P., Gadgil, A. & Koomey, J. G., 2002. *www.annualreviews.org.* [Online]
Available at:
https://www.annualreviews.org/doi/full/10.1146/annurev.energy.27.122001.083425
[Accessed 11 March 2022].

Fan, C. et al., 2020. Improving cooling load prediction reliability for HVAC system using Monte-Carlo simulation to deal with uncertainties in input variables. *Energy and Buildings,* 226(110372).

Hong, T. & Fan, S., 2016. Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting,* 33(3), pp. 914-938.

Jing, Z., Yaoqi, D. & Xiaojuan, L., 2018. Uncertainty Analysis of Weather Forecast Data for Cooling Load Forecasting Based on the Monte Carlo Method. *Energies,* 11(SP - 1900).

Lucas Segarra, E., Ramos Ruiz, G. & Fernández Bandera, C., 2020. *Probabilistic Load Forecasting for Building Energy Models,* Pamplona, Spain: School of Architecture, University of Navarra.

Steven, K., 2020. *Encyclopedia Britannica.* [Online]
Available at: https://www.britannica.com/topic/oil-crisis
[Accessed 2022 March 11].

Vanting, N. B., Ma, Z. & Jørgen, B. N., 2021. *A scoping review of deep neural networks for electric load forecasting.* [Online]
Available at: https://energyinformatics.springeropen.com/articles/10.1186/s42162-021-00148-6
[Accessed 11 March 2022].

Vesa, A. V. et al., 2019. *A Stacking Multi-Learning Ensemble Model for Predicting Near Real Time Energy Consumption Demand of Residential Buildings,* Cluj-Napoca: Department of Computer Science, Technical University.

Xu, L., Wang, S. & Tang, R., 2019. Probabilistic load forecasting for buildings considering weather forecasting uncertainty and uncertain peak load. *Applied Energy,* 237(ISSN 0306-2619), pp. 180-195.