Energy Forecasting: Phase 3 – Ensemble Modeling

VOLT Analytics (Orange 12)
Tobi Yañez, Omry Brewster, Vishali Chauhan,
Tyler Farr, Lu Shen

October 28, 2023

Table of Contents

Overview 1
Methodology & Analysis 1
Data Used 1
Ensemble Modeling 2
Results 2
Recommendations 4
Conclusion 4

Energy Forecasting: Phase 3 – Ensemble Modeling

Overview

American Electric Power Co. (AEP) partnered with VOLT Analytics to develop a model to forecast the Appalachian Power territory's hourly energy load in megawatts (MW). In the project's initial phase (RFP #: EF – F2.H1), we tested exponential smoothing and seasonal ARIMA models from the provided data. We concluded that the ARIMA model forecasted better due to a lower error. In the second phase (RFP #: EF – F2.H2), we built a Prophet and neural network model, with the neural network model achieving the lowest error.

In this phase, we improved our model further with a weighted ensemble model combining the ARIMA, neural network, and Prophet model. This model boasted a mean absolute percent error (MAPE) of 3.67% on the validation data set. The corresponding mean absolute error (MAE) was 146.25 megawatts. We recommend using the ensemble model for hourly forecasting due to the lower MAE and MAPE values compared to the previous individual models. This model will improve forecasting accuracy and better meet the customers' needs while reducing operational expenses.

Methodology & Analysis

Data Used

The data provided contains information about the hourly megawatts (MW) load for the Appalachian Power territory of AEP from 2016 to 2023. We found that the data before January 1, 2020, did not capture the desired signal of hourly energy usage. As a result, we decided to subset the data from January 1, 2020, to October 25, 2023, comprising 33,455 hourly observations. The data contained no missing values, so imputation was not required. Figure 1 shows the subsetted data regarding its energy use.

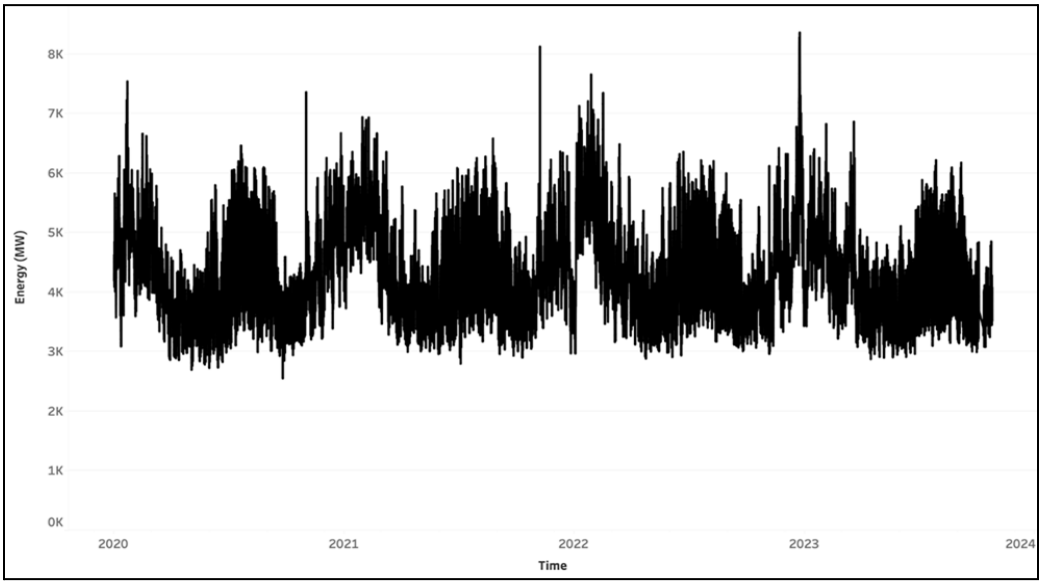

Figure 1: Hourly energy usage for data in megawatts from January 1, 2020, to October 25, 2023.

As we only attempt to predict energy usage from October 27, 2023, to November 2, 2023, we decided to evaluate our models on the most similar week of data. We assessed our models using MAPE and MAE on our validation set encompassing 168 hourly observations from October 19, 2023, to October 25, 2023. Ultimately, we compared the ensemble model to our previous models from RFP #: EF – F2.H1 and RFP #: EF – F2.H1 with updated validation data.

Ensemble Modeling

Our weighted ensemble model was developed by integrating an ARIMA model, a neural network model, and a Prophet model. We built these models from the findings of our previous work on RFP #: EF – F2.H1 and RFP #: EF – F2.H1. We incorporated an intervention variable into both the ARIMA and Prophet models to account for a winter storm that occurred between December 23, 2022, and December 24, 2022. These models have remained unchanged since their respective reports. However, we slightly modified the neural network model to enhance its ability to differentiate the captured signal from the other models. By adjusting the autoregressive terms to three (from zero) and the seasonal autoregressive terms to 15 (from ten), we observed an improvement in the predictive capability of the ensemble model.

Instead of taking the simple average of each model to form our final ensemble model, we weighted the individual models before combining them to improve our forecast. Our final ensemble comprised 50% neural net, 33.3% seasonal ARIMA, and 16.7% Prophet. We found that while Prophet captured different signals, it was highly variable between validation sets, and thus, it weighed less than the other models. On the other hand, the neural net performed well across validation sets in previous reports, leading to us assigning it the most prominent weight. In comparison, the seasonal ARIMA performed consistently well, although not as well as the neural net, and thus, it was assigned an intermediate weight.

Results

Our team found that the selected weighted ensemble model outperformed other individual/ensemble models when evaluating MAPE and MAE. The weighted ensemble model of seasonal ARIMA, Prophet, and neural network had a MAPE of 3.67% and an MAE of 146.25 megawatts on the validation set. The next closest-performing model was the average ensemble model of seasonal ARIMA, Prophet, and neural network, which had a MAPE of 3.74% and an MAE of 148.82 megawatts on the validation set. Table 1 shows the MAPE and MAE values on the validation set for the individual models (multiplicative Holt-Winters, seasonal ARIMA, Prophet, neural network) and ensembles.

Table 1: MAPE and MAE for validation data set

Model	MAE	MAPE
Multiplicative Holt-Winters	275.11	6.84%
Seasonal ARIMA	181.07	4.63%
Prophet	222.30	5.73%
Neural Network	185.76	4.73%
Ensemble of Seasonal ARIMA, Prophet, and Neural Network	148.82	3.74%
Weighted Ensemble of Seasonal ARIMA, Prophet, and Neural Network	146.25	3.67%

Figure 2 illustrates the predicted and actual energy usage observed from October 19, 2023, to October 25, 2023, for the weighted ensemble model, respectively. The predicted values are blue, while the actual values are black. The figures demonstrate that the weighted ensemble model excels in capturing intricate details within the seasonality pattern, attributed to the models capturing various signals.

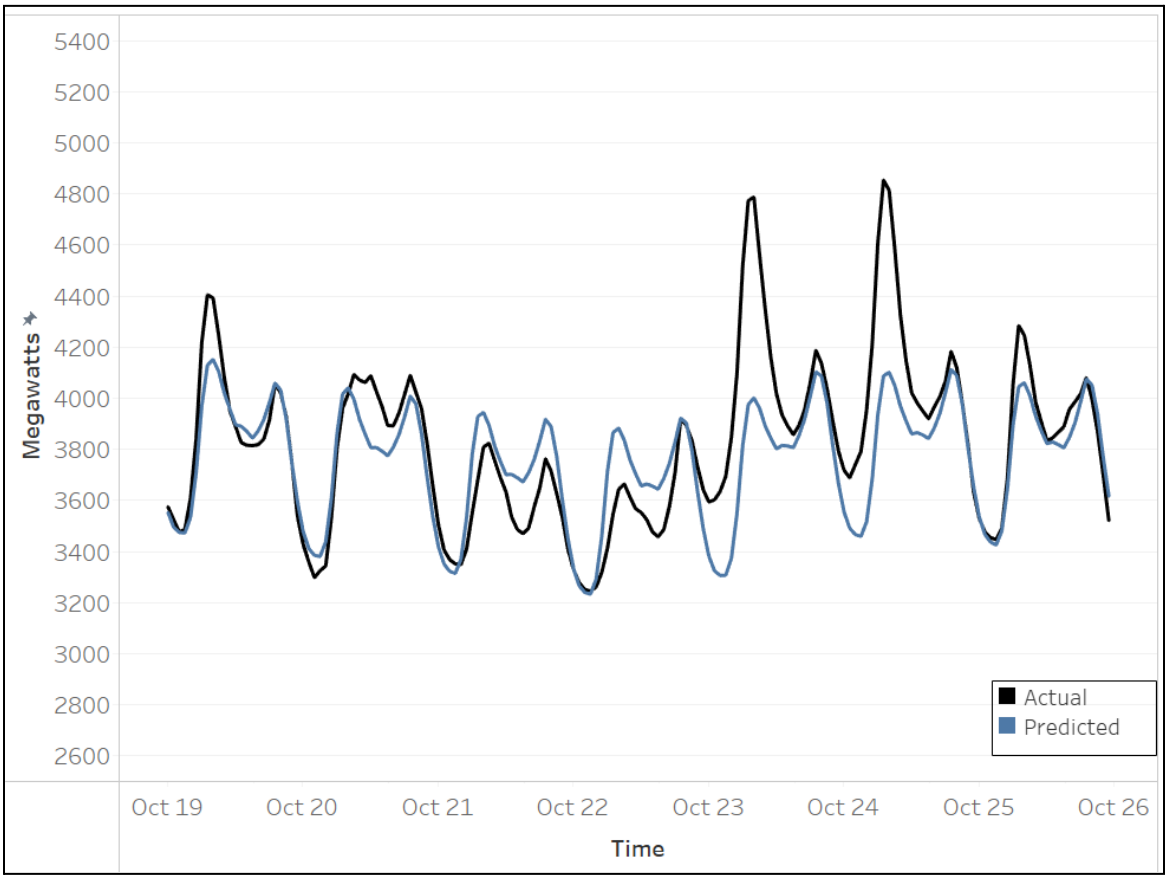

Figure 2: Time plot for the predicted energy usage of the ensemble model and the actual energy usage from October 19, 2023, to October 25, 2023.

Recommendations

Based on the results of our analysis, we recommend that AEP take the following steps:

- Use the ensemble model for near-term forecasting, as it has the most predictive power.
- Consider incorporating external temperature or precipitation data, as this can drive electricity usage.
- Flag any potential extreme weather events, as these can cause large surges in energy demand.
- Research other external variables that may provide predictive value for the ensemble model.

Conclusion

After conducting a thorough analysis, our team has examined multiple models and determined that the chosen weighted ensemble model has demonstrated superior performance. This model achieved an MAE of 146.25 megawatts and MAPE of 3.67% when applied to our validation dataset. Consequently, we highly recommend the utilization of the weighted ensemble model, which incorporates extreme weather events, as a means to predict energy usage accurately. Moving forward, we strongly advise the inclusion of supplementary weather forecasts to enhance the precision of future energy usage predictions. Additionally, we encourage the exploration of additional external variables to augment the predictive power of our models further.

Homework Report Checklist

As instructed by Dr. Egan Warren, the team member(s) responsible for checking each item should enter their initials in the field next to each question. All items should be addressed before submitting the assignment with the initialed checklist attached.

Sections & Structure

Overview

LS	Is the overview concise?
LS	Does it provide context about the business problem? <Content>
OB	Does it briefly address your team's work, quantifiable results, and recommendations? <Action>
LS	Does it offer audience-centered reasons for recommendations? <Context>

Body Sections

LS	Does the report body include information on methods, analysis, quantifiable results, and recommendations?
LS	Is content grouped into appropriate sections (*methodology*, *analysis*, *results*, *recommendations*)?

Conclusion

LS	Does the report have a conclusion?
LS	Does the conclusion sum up the report and emphasize relevant takeaways?

Structure

LS	Does each major section have a heading?
LS	Are sections, subsections, and paragraphs organized logically for easy navigation?

Visuals

Introduction, Discussion, and Captions

TY	Is each visual introduced in the text before it appears?
TY	Is each visual close to where it is introduced?
TY	Does each visual include a title with the following information: type (*table* or *figure*), number, and a descriptive caption?
TY	Is each visual discussed and interpreted in the text?
TY	Are figures and tables numbered separately?
TY	Are table captions above the table? Are figure captions below the figure?

Visual Design

TY	Do figures/tables use audience-friendly labels rather than variable names?
TY	Are the visuals easy to interpret?
TY	Are the visuals appropriately sized?
TY	Do tables appear on one page (*not split between 2 pages*)?
TY	Are legends and axis labels included for figures?
TY	Are numbers in tables right aligned?

TY	Are the visuals designed well (*ex*: *re-created in Word or Excel*, *not blurry or stretched*,...)?

Document Design

Title Page Design

TY	Does it include a descriptive title?
TY	Does it state the team name, team members' names, and the submission date?

Table of Contents Design

TY	Does it list all the major sections of the report with corresponding page numbers?
TY	Do the page numbers and sections in the Table of Contents match the report?

Document Design for Entire Report

TY	Is a standard typeface (*Calibri*, *Arial*, *etc.*) used?
TY	Is the size of the body text between 10-12 pt.?
TY	Are headings and subheadings used to organize information?
TY	Are distinctive text styles (*bold*, *italic*, *etc.*) used to distinguish between heading levels?
TY	Are text styles for headings used consistently (*ex*: *all level-one headings are bold*)?
TY	Are all paragraphs an appropriate length (*fewer than 12 lines*)?
TY	Is white space used to indicate paragraph breaks?
TY	Are bullet lists used for a series of items and numbered lists to show a hierarchy?

Writing Style and Mechanics

Spelling and Capitalization

TF	Are spelling errors located and corrected?
TF	Is spelling consistent throughout (*no switching between acceptable spellings*)?
TF	Is capitalization used appropriately (*proper nouns*, *etc.*)?
TF	Is capitalization of words consistent throughout the report?

Grammar and Punctuation

TF	Are verb tenses used appropriately?
TF	Are marks of punctuation used appropriately?
TF	Is subject-verb agreement used in every sentence?
TF	Is the grammar checker updated and are underlined grammar issues addressed?

Writing Style

TF	Are all sentences in the report easy for your audience to understand quickly?
TF	Are most sentences written in active voice?
TF	Are idioms and vague words eliminated from the report?
TF	Are acronyms introduced before being used?
TF	Are well-written topic sentences included at the beginning of each paragraph?
TF	Are lists parallel?
TF	Is the appropriate point of view used when addressing your audience or describing team actions?