Predicting Abnormal Grid Conditions

ISYE7406 Spring 2024 - Final Project Report

Will Coughlin
wcoughlin@gatech.edu

April 14, 2024

Abstract

Many organizations across the United States bear the responsibility of operating the electric grid and ensuring its reliability. Integral to the maintenance of grid reliability is the anticipation and proactive response to projected abnormal system status as a result of supply deficiencies. ISO New England, one of these grid operators, maintains public archives of power system forecasts and occurrences of abnormal grid conditions. Using this public data with common data mining techniques, we build a binary classification model that performs reasonably well in predicting the occurrences of abnormal status declarations. Support vector machine and gradient boosted tree classifiers have the ability to correctly identify 7 of 11 occurrences of abnormal conditions while still attempting to minimize false positive rate. Among all predictors, these models place high weight on projected load, anticipated cold weather outages, available generation, and projected surplus, and thus we identify these factors as most important in the prediction of power grid abnormality.

1 Introduction

1.1 Background

Planning for and mitigating interruptions to electric grid reliability is one of many crucial responsibilities for the assorted agencies who operate the grid across the United States. The non-governmental bodies in charge of grid operation are known as Regional Transmission Organizations (RTOs) or Independent System Operators (ISOs) and perform their duties under regulation by the United States government. One such RTO/ISO is ISO New England (ISO-NE), the sole grid operator for the New England region of the United States. Grid operators like ISO-NE monitor and forecast the electric power system every day, tracking factors which may threaten the stability of the grid.

1.2 Research Questions

This report addresses two key questions:

1. Can public data on power grid conditions and weather factors be used to build a binary classification model that accurately predicts the occurrence of abnormal grid conditions?
2. Which factors are the most influential in the occurrence of abnormal grid conditions?

1.3 Strategies and Outline

Historical grid forecasts and abnormal system status instances are available to the public on the ISO-NE website. Through scripted web scraping and data processing, these files were downloaded, matched, and cleaned to produce a unified dataset. This data will be discussed in greater detail in the next section, including findings from exploratory data analysis.

Then, the training and evaluation of four classes of binary classification models will be described. These are:

- Logistic Regression
- Support Vector Machine
- k-Nearest Neighbors
- Gradient Boosted Trees

Key findings from the modeling and evaluation will be presented, followed by their implications and the conclusions extracted from this data mining project. Finally, an overview of future work and lessons learned will be given to conclude this report.

2 Data Source and Description

2.1 Source and Processing

The data for this project comes from two public archives published by ISO-NE. These archives provide daily records dating back to the beginning of the year 2017. The web scraper for this project pulled all records from January 1st, 2017 to December 31, 2023. Between the two archives, there is at least one file for each day in this seven-year span.

The first data source is the *Power System Status Archive* [4], containing every instance ISO-NE has declared an abnormal system status. These records are provided as a single HTML or CSV file for each year on record. Second is the *Seven Day Capacity Forecast Archive* [5], which contains historic capacity, demand, and weather forecasts for each day, from a range of one to six days ahead. This data source also provides data downloads in HTML or CSV format.

By treating each date occurring within the two datasets as a unique key, the files were merged to produce a single dataset with one row corresponding to a single date. For each date, the columns of this unified dataset comprise of power system forecast values for that date as predicted from one to six days ahead, as well as whether certain abnormal grid conditions were declared on that date. Table 1 provides a sample of these variables. For a full listing, see Appendix A.

Table 1: A sample of attributes within the unified dataset. Note that columns ending in "_1" are the one-day-ahead values from the full dataset.

Column Name	Description
Abnormal	One or more abnormal statuses declared (Binary)
ACWO_1	Anticipated cold weather outages (MW)
CSO_1	Total capacity supply obligation (MW)
DPB_1	Dew point in Boston (°F)
HTB_1	High temperature in Boston (°F)
ITP_1	Import at time of peak (MW)
OGO_1	Other generation outages (MW)
PS_1	Projected surplus/deficiency (MW)
RRIR_1	Required reserve including replacement (MW)
TGA_1	Total generation available (MW)
TLRR_1	Total load plus required reserve (MW)

For cleaning and reduction purposes, 11 rows were dropped from the original dataset: six contained many missing values for a specific date and five were dropped to account for the lag values for the first few dates of 2017. Additionally, one power system forecast attribute was dropped due to a lack of variance. Finally, all statuses were combined into a single binary indicator of abnormal conditions. As a result, the prepared dataset contained 2545 daily rows, with 102 numeric features and one binary response.

2.2 Abnormal System Statuses

ISO-NE defines multiple levels of abnormal system declaration. These statuses are not mutually exclusive, and may occur with others. Table 2 enumerates the specific occurrences of these statuses in the dataset. See Appendix B for definitions of these statuses.

Table 2: Occurrences of Abnormal Grid Conditions in the ISO-NE service area from 2017 through 2023.

Abnormal Status(es)	Count
Only ACON	31
Only MGEN	1
ACON, OP41, and OP42	1
ACON, OP41, OP42, OP43, and OP45	1
ACON, OP41, OP42, OP43, OP44, and OP45	1

As shown in Table 2, only a total of 35 instances of abnormal status occurred in the date range. This causes a large class imbalance that introduces some challenges to modeling and evaluation. Strategies for overcoming these issues will be highlighted in the Methodology section of this report.

2.3 Correlation and Variable Dependence

High correlation was identified between several groups of variables. For example, Boston and Hartford high temperatures and dew points tended to be highly positively correlated, while the amount of generation outages and the amount of available generation showed strong negative correlation. Correlation analysis also uncovered some redundant variables within the dataset. For example, required reserve including replacement is merely required reserve plus replacement reserve requirement. In Figure 1, large dark red circles indicate the strong positive correlation for these temperature and other redundant features, while the small dark blue circles show strong negative correlation in a small subset of our features.

Figure 1: Strength of correlation between each variable in the prepared dataset.

With a combination of hierarchical feature clustering [2, 7] and knowledge of which factors are merely combinations of others, a set of features were identified for possible elimination in order to shrink the feature space. Refer to Appendix C for feature dendrograms before and after elimination.

Using the one-way ANOVA test, means were tested for each predictor between groups which are defined by occurrence or lack of occurrence of abnormal conditions. From these results, initial conjecture about important factors was deduced by the F-test statistic and P-values produced. ANOVA test results indicate that projected surplus/deficiency, projected peak load, and anticipated cold weather outages are among the factors with the most significant difference in between-group means. Figure 2 shows an example boxplot and density plot comparing the distribution of values for projected surplus/deficiency when abnormal conditions occur vs when they do not. There is a clear difference in the central tendency of each group.

Later in this report, feature importance for the best performing models will confirm some of this conjecture. It must be noted that these tests were not rigorous, as verifications were not performed related to the ANOVA normality and variance assumptions. Nonetheless, it was useful and illustrative for speculative purposes. In the following section, we will present an overview of the modeling methods used in this project.

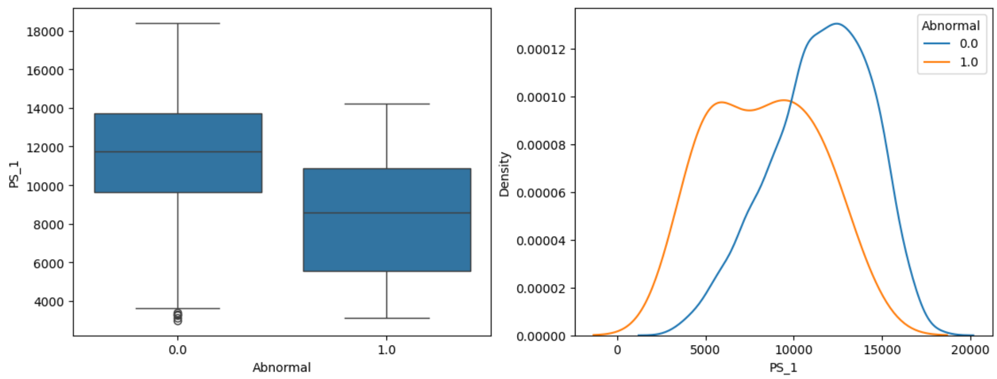

Figure 2: Distribution of Projected Surplus/Deficiency when abnormal conditions occur and when they do not occur.

3 Methodology

3.1 Minority Class Oversampling and Data Splitting

As noted in Section 2.2, only 35 out of 2545 observations in the prepared dataset are of actual dates of abnormal system status. This small 1.38% minority presented a challenge in the selection of cross validation and test evaluation metrics. For example, it would be easy to achieve high accuracy when predicting all cases to have normal status. Therefore, the Synthetic Minority Oversampling Technique (SMOTE) [3] was employed to balance the training data. First, the prepared full dataset was split with 70% reserved for training and 30% held for testing. This is a common test/train split proportion and resulted in a test dataset size of 764 observations. Additionally, splitting was performed in a stratified fashion, preserving the proportion of abnormal status observations within the test and train subsets. Then, on the training set only, SMOTE was leveraged to oversample the observations of abnormal status to balance the proportion of abnormal observations with normal ones. SMOTE oversampling resulted in a new training dataset size of 3514 observations.

3.2 Feature Selection

For this study, models using variables for one day ahead as well as up to six days ahead were explored. Any models built using only one-day-ahead variables could possibly demonstrate the ability of simpler forecasting with less computationally expensive training, providing immense value to grid operators. Usage of values as predicted up to 6 days ahead may provide more complex information that could be crucial in more accurately predicting the response.

Feature selection on the one-day-ahead factors was performed using the hierarchical distance-based clustering technique first mentioned in Section 2.3. Features at a distance threshold of less than 0.25 were identified (Appendix C). Then, using intuitive knowledge of which features are merely combinations of others, seven features were eliminated, thus reducing the number of one-day-ahead predictors for the reduced models from 17 to 10.

Finally, a feature set for a family of expanded models was produced that incorporates these narrowed features for all one through six days ahead. This expanded feature set contains 60 predictors overall. For example: an observation for January 1, 2020 contains features for high temperature in Boston as predicted one day ahead, two days ahead, and so on. Acknowledging the lack of complete rigor in this process, the full set of single-day features were not tested on all one through six days ahead. For technical purposes, the additional computational cost would have been a hindrance to this study.

3.3 Principal Components Analysis

In efforts to identify a set of engineered features that may reduce model complexity while achieving high predictive power, principal component analysis (PCA) was performed on each of the three feature sets. The specific number of principal components to use was selected as a hyper-parameter during cross-validation.

3.4 Classification Methods

A mix of parametric and non-parametric methods were utilized with each of the three feature sets. Before PCA or model fitting, all predictor values were centered and scaled. With both PCA and non-PCA variants, six total feature sets were used to fit each model family. These models and their benefits are as follows:

- Logistic Regression - Easy to interpret and fit, compatible with regularization to account for multi-collinearity or sparsity.
- Support Vector Machine (SVM) - Linear kernel is easy to interpret, alternate kernel methods to account for non-linearity.
- k-Nearest Neighbors (KNN) - Easy to interpret, can capture non-linearity.
- Gradient Boosted Trees (GBT) - Multi-tree method that fits multiple classifiers to generally perform better performance than a single model.

Except for GBT, all model types are implemented by the `sklearn` package in the Python language. GBT uses the `xgboost` Python package for its implementation.

3.5 Model Tuning and Metrics

With 10-fold cross validation, optimal hyperparameters were chosen for each of the above, then the full training set was fit on each. In cross validation, the F2 score was used to select the best model of all candidates. The F2 score is based on the harmonic mean of precision and recall, with more weight given to recall. Precision is a measure of how many predicted positives are truly positive and is important when false positives are costly. Recall is a measure of how many actual positives were identified by the model, and is important when there is high cost associated with false negatives. Therefore, using the F2 score accounts for both, while placing more weight on avoiding misclassifying an actual abnormal system status as normal.

After cross validation, the models were then refit with the full training set and evaluated on the test set. Using the F2 score, the best model of each type was identified. These results will be discussed in the next section.

4 Results

4.1 Selected Models

After retraining and predicting with each model on the test set, F2 score was used to identify the best performing model for each type. These models are described by Table 3.

Table 3: Best models of each type, according to F2 score.

Model Type	Feature Set	Selected Parameters	Precision	Recall	F2
Logistic Regression	1-6 days ahead	No regularization, $threshold = 0.22$	0.04265	0.8182	0.1765
SVM	1-6 days ahead	Linear kernel, $C = 0.1$	0.06087	0.6364	0.2201
KNN	One day ahead, reduced features	$n\_neighbors = 61$, distance weighting, Manhattan distance	0.02532	0.9091	0.1139
GBT	1-6 days ahead	*See Appendix D*	0.06306	0.6364	0.2258

The best-performing model according to F2 score was shown to be the Gradient Boosted Trees method, with an F2 score of approximately 0.23. The GBT model just edges out over the SVM, which has an F2 score of just about 0.22. Despite decent recall for all models, their precision values do appear to be quite abysmal. The confusion matrix for the GBT model (Figure 3) shows that the model predicts 104 false positives, contributing to the low precision. Confusion matrices for all models in Table 3 are available in Appendix E.

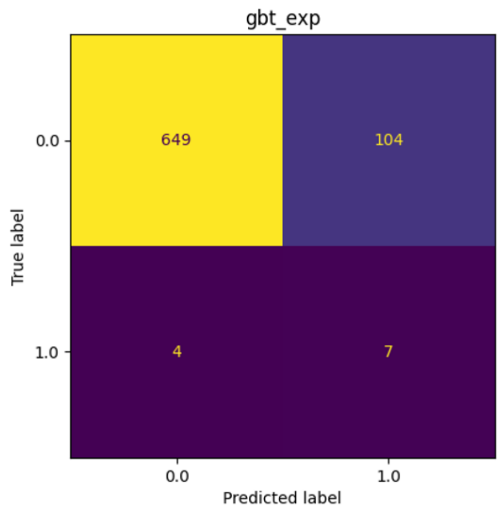

Figure 3: Confusion Matrix for test predictions by the selected GBT model.

4.2 Feature Importance

From the Logistic Regression, (Linear) SVM, and GBT models, it is easy to extract metrics for feature importance. For Logistic Regression and SVM, the absolute value of the coefficients corresponding to each feature provide a good estimate of relative feature importance. On the other hand, the metric for GBT is a measure of overall information gain provided by the feature. Table 4 lists the top five highest weighted features for these three models. KNN is excluded from the table as it does not provide an easily available metric for feature importance. A full listing of feature importance is provided in Appendix F.

Table 4: Top 5 features for Logistic Regression, SVM, and GBT models, ranked by highest importance first.

Rank	Logistic Regression Feature (`Variable Name`)	SVM Feature (`Variable Name`)	GBT Feature (`Variable Name`)
1	Predicted Peak Load - One Day Ahead (`PPL_1`)	Predicted Peak Load - One Day Ahead (`PPL_1`)	Projected Surplus/Deficiency - Three Days Ahead (`PS_3`)
2	Total Generation Available - Two Days Ahead (`TGA_2`)	Anticipated Cold Weather Outages - Five Days Ahead (`ACWO_5`)	Predicted Peak Load - Two Days Ahead (`PPL_2`)
3	Predicted Peak Load - Two Days Ahead (`PPL_2`)	Anticipated Cold Weather Outages - Six Days Ahead (`ACWO_6`)	Predicted Peak Load - One Day Ahead (`PPL_1`)
4	Predicted Peak Load - Four Days Ahead (`PPL_4`)	Total Generation Available - Two Days Ahead (`TGA_2`)	Projected Surplus/Deficiency - Six Days Ahead (`PS_6`)
5	Total Generation Available - Four Days Ahead (`TGA_4`)	Projected Surplus/Deficiency - One Day Ahead (`PS_1`)	Anticipated Cold Weather Outages - One Day Ahead (`ACWO_1`)

5 Conclusions

5.1 Research Questions

Regarding the ability to build a binary classifier that can accurately predict occurrences of abnormal power grid conditions, it is indeed possible to correctly detect real occurrences of abnormality with moderate to high success. The trade-off for higher abnormality detection comes at the cost of increased false positives. When choosing an evaluation metric, F2 was selected to ensure higher weight was placed on recall, or detection of true positives among all actual positives. The consequences of failing to prepare for a strained power grid could be very disastrous, as critical infrastructure and healthcare may be severely affected by blackouts or other interruptions to electric service. [8] While the class imbalance problem likely contributes to high prevalence of false positives in all model types, the cost of a false positive would likely be deemed to be far lower than that of a false negative.

As the GBT and SVM methods perform very closely according to F2 score, either of these classifiers trained with expanded feature set of up to six days ahead provide the best performance in detecting true positives while also minimizing false positives.

Both top-performing models assign high relative weight to four key types of data: load, generation, outages, and surplus. For both models, values from one through six days ahead appear in the top five most important features, thus indicating the utility of these extended forecast values in prediction of grid abnormality. These importance measures do also support the conjecture from Section 2.3 regarding which factors are most related to abnormality.

5.2 Future Work and Shortcomings

One unique challenge caused by severe response class imbalance is the small number of actual positive cases available to test against during both cross-validation and test set evaluation. With a larger dataset, and thus a higher count of positive occurrences, an improved model that further decreases the false positive rate while keeping true positive detection rate high may be possible to build. ISO-NE provides many more years of historic data in unstructured format. Additional effort to extract and structure this data may allow for closer approximations of true patterns to be captured from a larger dataset.

Also related to this issue is the ability to model the response as a multinomial categorical variable. As the current prepared dataset stands, most specific abnormal statuses appear fewer than ten times, and sometimes just once. This number is far too small to be able to learn any meaningful patterns in the occurrence of these specific grid statuses. A larger dataset with a sufficient number of occurrences for more specific grid conditions may allow for more accurate modeling without losing unique information that may be obscured by combining all statuses into a single response.

Given enough computing power to deal with the added complexity, variable selection could additionally be performed on a more systematic basis for all models independently, using more rigorous empirical techniques or information-based criteria. An attempt to specifically address correlation and multicollinearity was made by eliminating features which scored low on cluster distance and which were known to be combinations of other features. While additional L2 regularization and random feature selection in some of the employed models are helpful, remaining useless or highly correlated features may still muddy the waters in model building, providing lower performance than is possible with the given information.

5.3 Lessons Learned

This section provides an overview of what this project we have discussed has taught me, and the new challenges I faced in its execution. First, this project provided an interesting opportunity to scrape and mine real-world data. This experience included the programming of automatic data retrieval, inspection, and cleaning the data at a previously unexplored scale. Adequate research was required to understand what each of the candidate variables were really telling us. With such a real-world dataset, the class imbalance problem was another great learning experience. Second, while it is fortunate that there have been so few cases of power grid abnormality in the ISO-NE service area, the small minority of abnormal cases presented a problem I had not encountered. Researching and enacting oversampling with SMOTE to handle this issue was a valuable learning experience. Next, the experience of systematically tuning many different models and building processing pipelines involving standardization, dimensionality reduction, parameter grid search, and cross-validation was unlike many of the smaller scope experiences I have encountered on most homeworks and class assignments. Learning the recommended approaches to performing so many steps in the data mining process to avoid data leakage was quite interesting. Finally, I learned a lot more about classification metrics beyond accuracy and AUC. In my experience up to this project, not much attention was given to precision, recall, F1, F2, or F_β. Learning how to use these metrics to balance performance when, for example, false negatives are far more costly, was a great learning opportunity.

This project has functioned well as an exhibition of many skills and tools learned in this course. From report writing to new types of models–their strengths, weaknesses, and considerations, I found many lessons from the course and especially discussions from the Piazza forum to be useful in informing my approach.

Appendix A

Table 5: Listing of response variable and all one-day-ahead predictors in the full dataset. Note that all predictors also have values for two through six days ahead.

Column Name	Description
Abnormal	One or more abnormal statuses declared (binary)
ACWO_1	Anticipated cold weather outages (MW)
ADMO_1	Anticipated de-list offered (MW)
ADRR_1	Available demand response (MW)
CSO_1	Total capacity supply obligation (MW)
DPB_1	Dew point in Boston (°F)
DPH_1	Dew point in Hartford (°F)
HTB_1	High temperature in Boston (°F)
HTH_1	High temperature in Hartford (°F)
ITP_1	Import at time of peak (MW)
OGO_1	Other generation outages (MW)
PPL_1	Predicted peak load (MW)
PS_1	Projected surplus/deficiency (MW)
RRIR_1	Required reserve including replacement (MW)
RRR_1	Replacement reserve requirement (MW)
TAGI_1	Total available generation and imports (MW)
TGA_1	Total generation available (MW)
TLRR_1	Total load plus required reserve (MW)

Appendix B

Table 6: ISO-NE Abnormal Power System Statuses. Statuses not listed did not occur in the timeframe for this study.

Short Name/Variable Name	Long Name	ISO-NE Definition [6]
ACON	Master/Local Control Center Procedure	"A Master/Local Control Center Procedure No. 2 (M/LCC 2) alert is issued either for the entire region or for a local area to notify power system personnel and wholesale electricity market participants when abnormal conditions on the region's power system exist or are anticipated."
MGEN	Minimum Generation Warning	"ISO New England declares a Minimum Generation Emergency [. . .] for one of the following reasons: (1) ISO forecast projects a potential minimum generation condition; (2) Potential high frequency due to load and generation mismatch in ISO New England; (3) Next hour external transactions will, or possibly will, place the system in minimum generation warning or emergency condition; (4) Real-time event occurs that could cause Minimum Generation Emergency conditions if real-time only megawatts are not curtailed."
OP41	OP4 Action 1: Power Caution	"This is an advisory that available capacity resources are insufficient to meet anticipated demand plus operating reserve requirements, and ISO New England has implemented OP 4."
OP42	OP4 Action 2: Emergency Energy Alert	"ISO New England is declaring an Energy Emergency Alert (EEA) Level 1."
OP43	OP4 Action 3: Voluntary Load Curtailment	"ISO New England is requesting voluntary load curtailment of market participants' facilities in New England."
OP44	OP4 Action 4: Power Watch	"This is a notification that additional OP 4 actions may be implemented. In addition, if conditions warrant, ISO New England may issue a public appeal for voluntary electricity conservation. If the ISO asks for voluntary conservation through this action, the public can conserve electricity at home and at work by turning off unnecessary lights, equipment, and appliances, and minimizing the use of air conditioning. People who require air conditioning and other electrical appliances for health and safety reasons should not restrict their use."
OP45	OP4 Action 5: Emergency Energy Purchasing	"ISO New England is arranging to purchase available emergency capacity and energy, or energy only, (if capacity backing is not available) from market participants or neighboring regions. Control Area to Control Area transactions will normally be used as a last resort, when market-based emergency energy transactions are not available, or not available in a timely fashion."

Appendix C

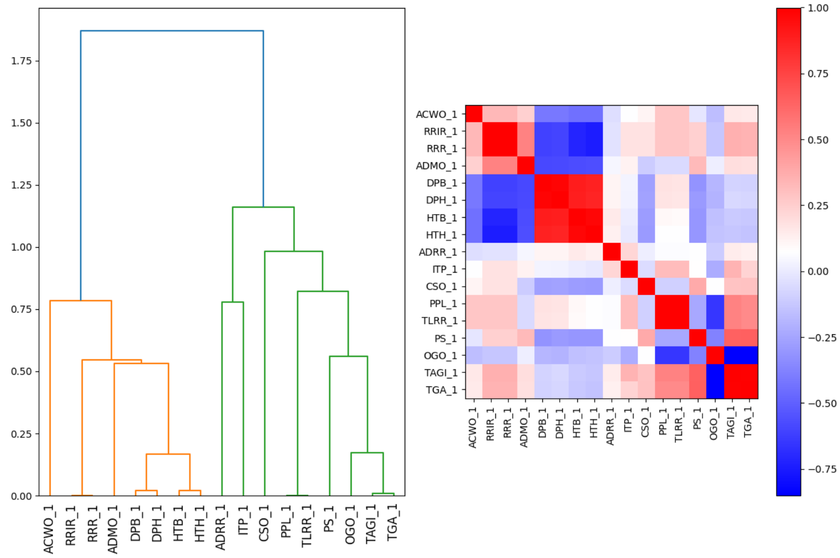

Figure 4: Hierarchical distance-based clustering of all one-day-ahead features using Ward's linkage.

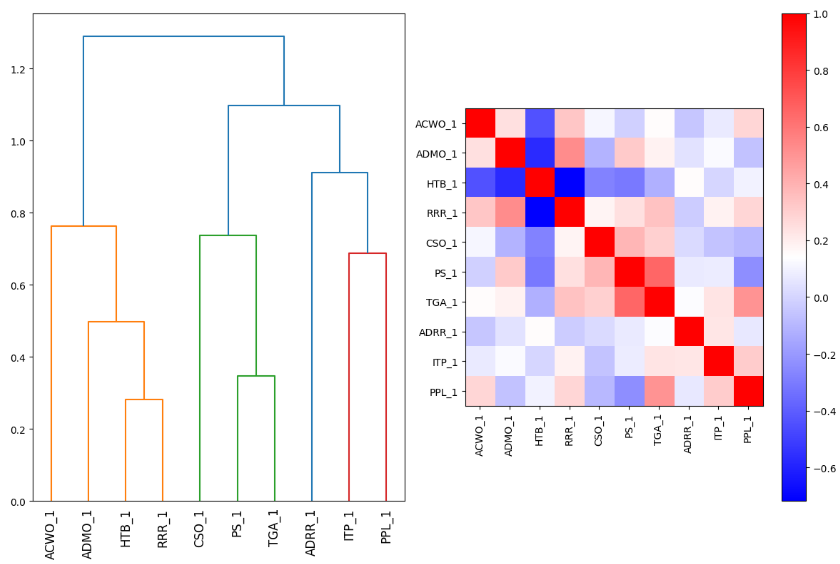

Figure 5: Hierarchical clustering of one-day-ahead features after elimination of redundant features within distance threshold of 0.25.

Appendix D

Table 7: The selected parameters for the GBT model using the `xgboost` Python language implementation.

Variable	Value	Description
`n_estimators`	1000	Number of trees in the ensemble.
`learning_rate`	0.01	
`reg_lambda`	1000	Strength of L2 regularization (λ).
`max_depth`	5	
`colsample_bynode`	0.5	Proportion of features sampled at each node of a tree.
`subsample`	0.5	Proportion of observations sampled during each boosting round.
`max_delta_step`	3	Useful for class imbalance. See reference [1].
`gamma`	10	Minimum loss reduction required for leaf node partition.

Appendix E

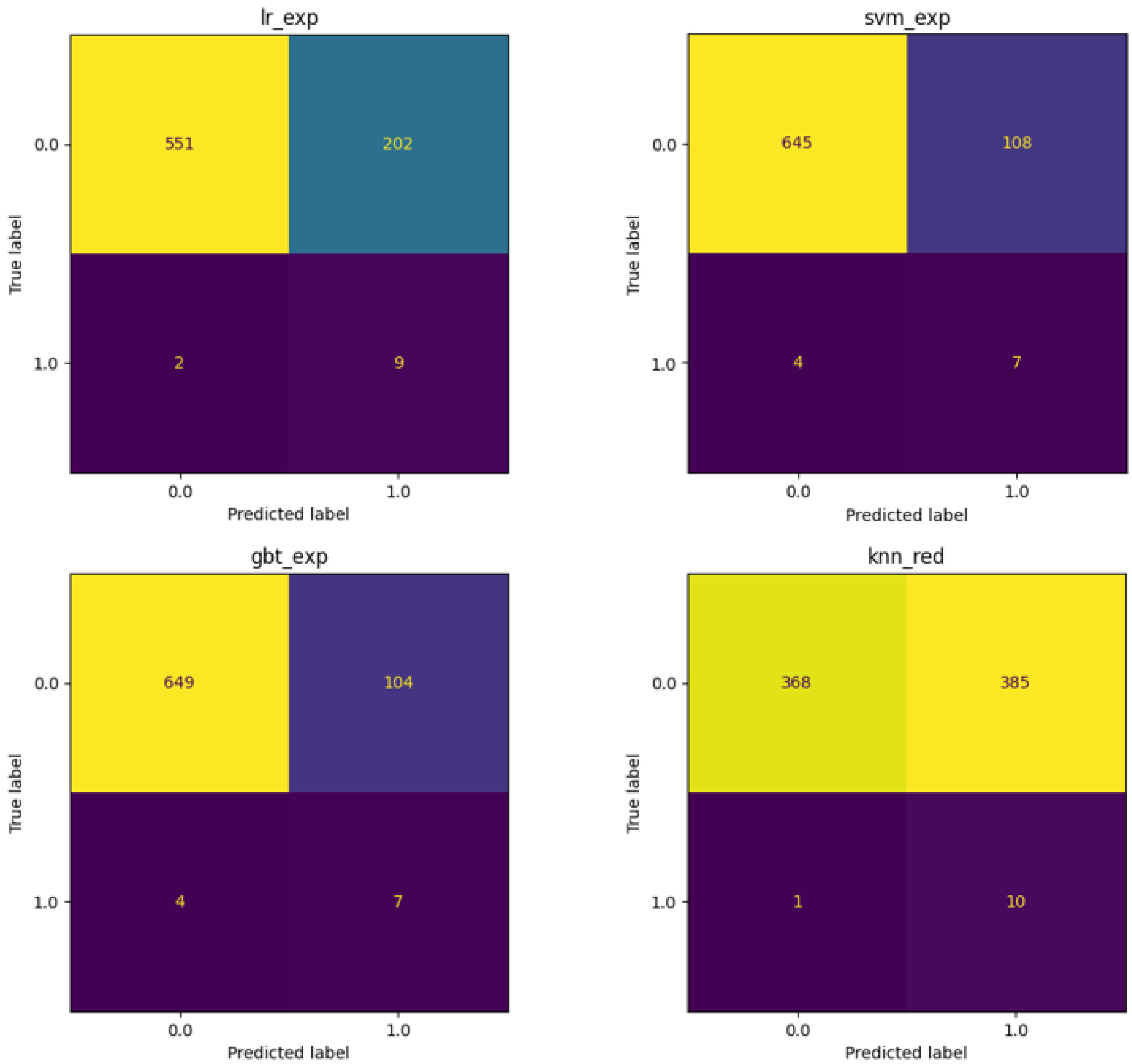

Figure 6: Confusion matrices for test set predictions by the best model of each type. Clockwise, from top left: Logistic Regression (`lr_exp`), SVM (`svm_exp`), KNN (`knn_red`), GBT (`gbt_exp`).

Appendix F

Table 8: Full listing of features and their importance measures for Logistic Regression, SVM, and GBT on the extended one to six-day-ahead feature set.

Rank	Logistic Regression Variable Name	Abs. Val. of LR Coefficient	SVM Variable Name	Abs. Val. of SVM Weight	GBT Variable Name	GBT Accuracy Gain
1	PPL_1	3.960496	PPL_1	1.647307	PS_3	0.061411
2	TGA_2	3.121723	ACWO_5	1.10353	PPL_2	0.054629
3	PPL_2	2.85314	ACWO_6	1.069217	PPL_1	0.04634
4	PPL_4	2.812669	TGA_2	1.018218	PS_6	0.030904
5	TGA_4	2.693831	PS_1	0.998177	ACWO_1	0.028123
6	HTB_5	2.274561	ACWO_1	0.993617	PS_2	0.027129
7	TGA_6	2.146682	PPL_2	0.900244	PS_1	0.026288
8	PPL_6	2.049483	PPL_4	0.889669	ACWO_6	0.025393
9	TGA_1	1.935402	HTB_6	0.880761	ACWO_5	0.024003
10	ACWO_5	1.908907	PPL_6	0.798596	ACWO_3	0.023457
11	HTB_2	1.843698	ADRR_5	0.788796	ACWO_2	0.023178
12	PS_1	1.81146	HTB_2	0.745746	CSO_1	0.022818
13	ITP_4	1.707533	TGA_4	0.717037	ACWO_4	0.021986
14	HTB_6	1.662613	HTB_5	0.705498	PPL_3	0.020365
15	ITP_3	1.629982	TGA_6	0.697166	CSO_2	0.019764
16	ACWO_1	1.610988	ITP_4	0.659119	RRR_2	0.019275
17	ACWO_2	1.316799	ITP_3	0.60047	CSO_6	0.019054
18	ACWO_6	1.29697	ACWO_3	0.599508	CSO_3	0.018018
19	TGA_3	1.220718	HTB_1	0.551044	PPL_6	0.017569
20	ADMO_2	1.199897	TGA_1	0.533337	RRR_1	0.017253
21	HTB_1	1.178018	ADMO_3	0.527023	PS_4	0.016895
22	ADRR_5	1.08827	ADRR_1	0.522572	RRR_3	0.016784
23	ACWO_4	1.056699	ADMO_2	0.470708	CSO_4	0.015723
24	ADMO_3	1.013792	PPL_3	0.412721	PPL_4	0.015367
25	ITP_5	0.919848	ADMO_6	0.397257	RRR_4	0.015118
26	ITP_2	0.861366	TGA_3	0.389044	PPL_5	0.014442
27	ACWO_3	0.759526	RRR_6	0.378629	ADMO_4	0.014166
28	PPL_5	0.753129	PS_5	0.370499	ADMO_3	0.014052
29	RRR_6	0.737458	PPL_5	0.342381	CSO_5	0.013687
30	ADRR_1	0.618764	ACWO_2	0.312487	RRR_6	0.012863
31	PS_3	0.570175	ITP_2	0.234433	HTB_6	0.012718
32	ITP_1	0.524755	ADMO_4	0.21003	RRR_5	0.012597
33	ADMO_4	0.506284	ITP_1	0.2099	PS_5	0.012511
34	PPL_3	0.434115	ITP_5	0.204972	ADRR_2	0.012249
35	PS_6	0.408404	ADRR_4	0.199881	ADRR_6	0.012143
36	ADMO_6	0.374987	HTB_4	0.153684	HTB_5	0.012137
37	TGA_5	0.364786	ADRR_6	0.152063	ADMO_2	0.011849
38	HTB_4	0.355267	RRR_3	0.144386	HTB_4	0.011836
39	CSO_6	0.344141	RRR_4	0.144338	ADRR_5	0.0118
40	RRR_4	0.259574	RRR_5	0.144291	TGA_6	0.011782
41	RRR_1	0.245847	ADRR_3	0.126556	TGA_1	0.011184
42	RRR_2	0.245847	ADMO_5	0.094372	ITP_6	0.010948
43	RRR_5	0.241444	RRR_1	0.090137	ADRR_1	0.010754
44	RRR_3	0.226371	RRR_2	0.090137	ITP_1	0.010689

Rank	Logistic Regression Variable Name	Abs. Val. of LR Coefficient	SVM Variable Name	Abs. Val. of SVM Weight	GBT Variable Name	GBT Accuracy Gain
45	PS_5	0.220903	TGA_5	0.08524	TGA_2	0.010652
46	ITP_6	0.175724	PS_3	0.079964	ADMO_1	0.010599
47	PS_4	0.16821	PS_2	0.078129	ITP_5	0.010314
48	ADRR_4	0.154637	ACWO_4	0.071077	TGA_3	0.010188
49	CSO_1	0.154496	PS_4	0.064612	HTB_3	0.010156
50	CSO_3	0.154327	CSO_3	0.059027	HTB_2	0.010031
51	ADRR_6	0.143781	CSO_4	0.058511	ITP_4	0.010027
52	ADMO_1	0.142532	HTB_3	0.054682	TGA_4	0.00999
53	PS_2	0.140397	CSO_6	0.051084	ITP_2	0.009954
54	CSO_4	0.113968	ITP_6	0.044044	ADMO_5	0.009797
55	ADRR_2	0.108312	CSO_5	0.034459	HTB_1	0.009748
56	ADRR_3	0.098819	ADRR_2	0.027439	ADRR_4	0.009152
57	ADMO_5	0.061675	CSO_2	0.023568	TGA_5	0.009138
58	CSO_2	0.05729	ADMO_1	0.022659	ITP_3	0.009004
59	HTB_3	0.049072	CSO_1	0.017714	ADMO_6	0
60	CSO_5	0.015176	PS_6	0.011723	ADRR_3	0
60	CSO_5	0.015176	PS_6	0.011723	ADRR_3	0

Appendix G

Source code for this project is available on GitHub at https://github.com/willcoughlin/dmsl-project.

References

[1] *XGBoost parameters*. Retrieved from https://xgboost.readthedocs.io/en/stable/parameter.html

[2] *Permutation importance with multicollinear or correlated features*. Retrieved from https://scikit-learn.org/stable/auto_examples/inspection/plot_permutation_importance_multicollinear.html

[3] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. 2002. SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research* 16, (June 2002), 321–357. https://doi.org/10.1613/jair.953

[4] ISO New England. *Power system status archive*. Retrieved from https://www.iso-ne.com/markets-operations/system-forecast-status/current-system-status/power-system-status-list

[5] ISO New England. *Seven day capacity forecast archive*. Retrieved from https://www.iso-ne.com/markets-operations/system-forecast-status/seven-day-capacity-forecast-list

[6] ISO New England. *Power system status descriptions*. Retrieved from https://www.iso-ne.com/markets-operations/system-forecast-status/current-system-status/alert-descriptions

[7] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2021. An introduction to statistical learning: With applications in r. In (Second). Springer. Retrieved from https://www.statlearning.com/

[8] The University of Texas at Austin Energy Institute. *The timeline and events of the february 2021 texas electric grid blackouts*. Retrieved from https://energy.utexas.edu/research/ercot-blackout-2021