Prediction of Hour of Coincident Daily Peak Load

Jinxiang Liu, Laura E. Brown
Department of Computer Science
Michigan Technological University
Houghton, MI
jliu15@mtu.edu, lebrown@mtu.edu

***Abstract*—To manage the power system well, it is important to keep the balance between the electricity supply and its demand. For many utilities, consumers need to pay additional money for the electricity during the peak time on days with the highest demands – coincident peak pricing. Therefore, if the consumers are informed the daily load peak a day or hours in advance, such surcharges may be avoided. The accurate forecasting of load peak time not only helps to provide a reliable electricity supply, but also be useful to reduce the cost of electricity for consumers. In this paper, we study the historical data and then build the classification models for winter and summer to predict the time of daily peaks 24 hours ahead. Several classification algorithms, including Naïve Bayes, SVM, Random Forest, AdaBoost, CNN, LSTM, Stacked Autoencoder, are applied to solve this problem. Finally, the performance of these methods have been examined and compared with LSTM having the best overall accuracy, precision, and recall.**

***Index Terms*—forecasting, load prediction, coincident peak**

I. Introduction

Knowledge of peak demands are critical to ensure proper management of electric grids. Ensuring sufficient power supply to meet the peak demands, may require additional infrastructure, purchase of power, integration of dispatchable power resources, or customer demand response programs. Many different local utilities, Independent System Operators (ISOs), and Regional Transmission Organizations (RTOs) now implement Coincident Peak Pricing (CPP) programs. These program add a surcharge for consumers (mainly, for industrial and large power consumption companies) during the coincident peak hour, that is, the hour when the power consumption is at its highest from a utility requesting it from a wholesale supplier or across an ISO or RTO geography.

The Independent Electricity System Operator (IESO) in Ontario, Canada, which is a power system corporation, encourages its consumers to reduce the electricity usage during the peak time by applying a surcharge for power. The consumers pay higher price for the electricity during the highest five daily peak hours every fiscal year (5 Coincident Peak - 5CP). Since only at the end of every fiscal year these five peak hours can be known, it can be hard for the consumers to avoid such surcharges. Hence, the forecasting of daily peak time is a important process for the consumers. Note, IESO itself sends warnings and supplies information on peak tracking (top 10 peaks in the fiscal year) for their customers because of the importance of this problem.

In the technique presented by Bon Ryu et al. [1], predicting five daily peak hour with highest power demands (5CP) has been transformed into a classification problem. Using historical power and weather data, their models, Naïve Bayes, classifies the following three tasks: *daily peak* - predicting weather a 5CP occurs on a day, *3-hour peak* - for a daily peak, predict the three hours mostly to have the 5CP, and *1-hour peak* - for a daily peak, predict the exact hour when the 5CP occurs. Although the prediction of peak day and 3-hour peaks have the precision up to 0.71 and recall up to 1.0, the performance for 1-hour peak prediction still need to be improved. In this study, we also use classification models to solve the prediction problem. However, we focus on the forecasting of daily peak hour for each day 24 hours ahead.

We studied a dataset of power demands and weather variables in Ontario from the fiscal year 2003 to 2008. Two separate models have been built for two main seasons: summer and winter. Based on the previous hourly data, these models can decide if the daily load peak for tomorrow happens at the same hour or not. The classification methods that we considered in this study include Naïve Bayes, Support Vector Machine (SVM), Random Forest, AdaBoost, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Stacked Autoencoder. Their performance also have been evaluated with precision, recall, and accuracy. Overall, the prediction for winter is better than it for summer. The technique of LSTM shows the best performance to address this problem, but different methods have advantages and disadvantages in their usage and results. In general, classification models are effective to locate the load peak hour for next day.

II. Related Work

Extensive approaches have been proposed to predict power demands. Such solutions often use the methods containing autoregressive integrated moving average (ARIMA) [2], K-nearest neighbors (KNN) [3], artificial neural networks (ANN) [2] [4] [5], autocorrelation with weather data [6], and so on. However, less work has been published about predicting the time of daily load peak and coincident peaks. Moreover, peak prediction is often posed as a classification [7] problem rather than a regression formulation (although that is not always the case [8]).

Livik et al. [9] propose a statistical method for predicting annual coincident peak. This problem differs from ours in that the daily peak is not the target of the prediction. Demand response operations in computer data centers is another application where accurate peak prediction is needed [10], [11]. Other

978-1-5386-8232-6/19/$31.00

related work focuses on peak reduction in individual homes or smart cities; [12] uses reinforcement learning for decision making on peak reduction in a home energy management system. [13] proposes a two-stage scheme to predict the hourly power demand in Polish power system.

Several papers address the peak prediction problem using the IESO data. As mentioned previously, [1] address the peak prediction problem with a classification model based on Naïve Bayes, which predicts the five peak hours with highest power demands in a fiscal year. In addition, they did the forecasting of 3-hour peak and peak day. The hourly data over 21 years have been used for testing. One of advantages for their approach is it does not rely on any heuristic thresholds. However, they use simple Naïve Bayes models and do not include historic data or features in their predictions.

Jiang et al. [14] proposed a new algorithm to predict if tomorrow is the day with one of top five daily load peaks on IESO data. Their models included 14-day ahead forecasts of load demand from IESO. We do not explore inclusion of the forecast data because this information is not now available for the historical records. In their method, probability theory has been used to calculate the possible of being one of five highest daily load peaks for the predicted load peak in next day. If this probability is greater than the threshold, the next day will be labeled as peak day. A data-driven approach is applied to set the threshold.

Xu and Cercone [15] also use a Naïve Bayes model for predicting the 5CPs. The authors identify 6 informative features from the historical data to inform their peak prediction, e.g., the forecast temperature is greater than 32 degrees Celsius, if the hour is within historical peak hour zones, etc.

Instead of focusing on top five daily load peaks in a fiscal year, our study works on the prediction of daily peak hour in Ontario. We also use classification models to solve the problem. But more attributes are taken into consideration, such as the hour of sunrise and sunset, relative humidity. Additionally, our models are based on the previous data to decide the daily peak hour 24 hours ahead. The classification techniques that applied in this study contains not only classical methods such as Naïve Bayes, SVMs and Random Forests, but also deep learning techniques including CNN, Stacked Autoencoder, and LSTM.

III. Dataset

The data of total power demand for Ontario, from May 1st, 2003 to April 30th, 2008, is provided by the IESO [16]. We obtained the corresponding hourly weather data from Canadian government climate website [17] for Toronto Buttonville Airport. Both the power demand and weather data have the sampling rate of 1 hour. The information about sunrise and sunset at Toronto is from the website of *Time and Date* [18]. The function of *isHoliday* in R package of *timeDate* has been used to check if a day is a holiday in Toronto.

For the weather data, we considered four attributes, which are temperature, humidex, windchill and relative humidity. According to the formulas given by Canadian government climate website [17], we calculated the values of humidex and windchill, by using the hourly data of dew point temperature, air temperature and wind speed from the website. Moreover, linear interpolation has been used to replace the missing values in the weather data. Because windchill does not exist when temperature is above zero Celsius or wind speed is equal to zero, we assigned such empty values of windchill as 1.

For the dataset of sunrise and sunset, a binary number has been used to indicate if corresponding hour has the sunrise or sunset. Since the time data of sunrise or sunset has the precision of a minute, and if it is more than 45 minutes past the hour, then we counted the sunrise or sunset to the next hour.

IV. Methodology

A. Summer and Winter

In this study, a year has been separated into two parts, summer and winter. The time from April 1st to September 30th is set as summer, while the time from October 1st to March 31st belongs to winter. Figure 1 and Figure 2 demonstrate why to consider two models for the different seasons; the two figures are based on the real-time hourly power demands from the fiscal year 2003 to 2008.

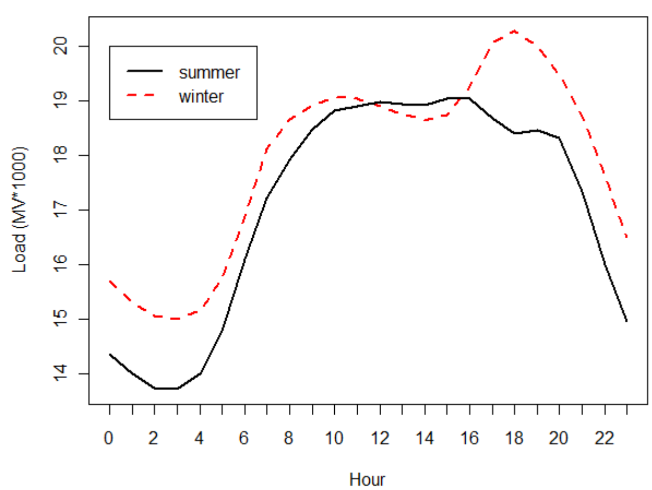

Fig. 1: Canadian fiscal year 2003-2008 average daily power demands on winter and summer

Figure 1 displays averages of daily load for summer and winter. Although the general trend of electricity consumption for summer has similarities to the trend for winter, the changes in the period of 12 p.m. to 6 p.m. are different, for example, from 4 p.m. to 6 p.m., the averages of electricity consumption increased in the winter, but it decreased in the summer. Also, such changes for summer are not so dramatic like those for winter.

Figure 2 shows an estimate of the probability density function of daily peak hour for summer and winter. It is easy to find that during the winter, the daily load peak is very likely to happen from 3 p.m. to 8 p.m.. However, during the summer, the daily peak hour has larger range, from 5 a.m. to 12 a.m..

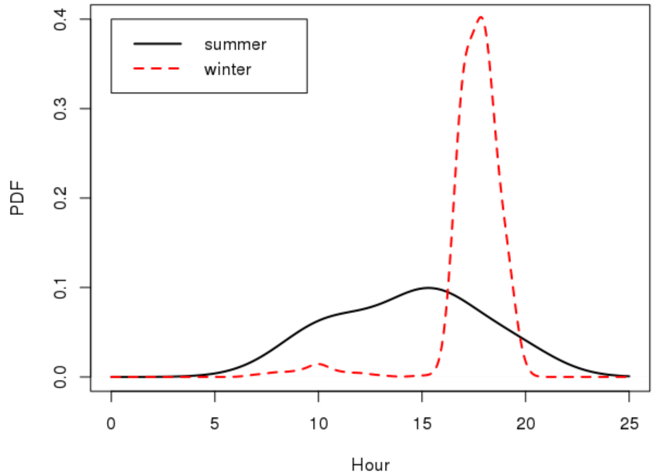

Fig. 2: Canadian fiscal year 2003-2008 daily peak time distribution for winter v.s. summer

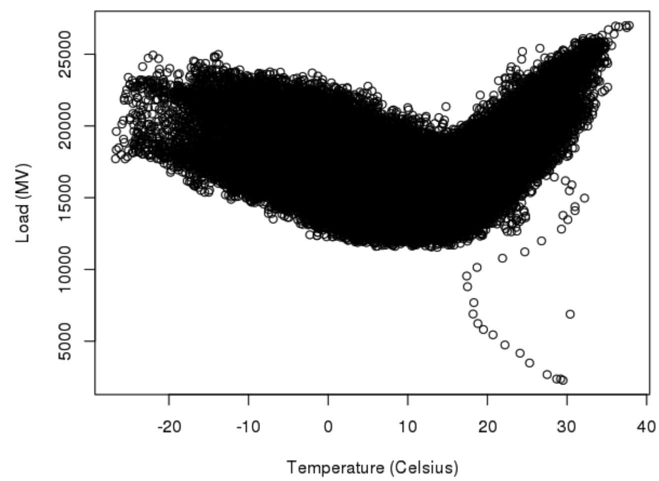

Fig. 3: Canadian fiscal year 2003-2008 power demand with temperature

Figure 1 and Figure 2 also indicate that the daily peak hour during the summer should be harder to predict than those during the winter, because of the range of peak hour and the changes of electricity demands.

B. Attributes

Many works [19], [20], [21] illustrate the great influence from weather on electricity consumption. The features, like temperature, relative humidity, windchill and humidex, are often used in models for load forecasting.

A strong correlation between load usage and temperature have been disclosed [19] [20]. To examine the relationship between temperature and load demands in Ontario, Figure 3 displays hourly electricity demands with corresponding value of temperature from the fiscal year of 2003 to 2008. In this figure, when the temperature lower than approximately 13 Celsius, the power consumption tends to increase with the decreasing of the temperature. But when the temperature higher than that value, the power demand is likely to grow with the increasing of the temperature. Since the values of humidex, windchill and relative humidity change with temperature, and they are important weather indicators, we also took them into account to predict the load peak time.

Table I and Table II present the attributes that are used for modelling. Table I lists all the continuous variables that have been used, while Table II outlines the categorical variables. In additional to the original hourly data we obtained, we also added power demands in the same hour over previous 2 weeks and the averages of load and weather data over 3-hour intervals (each 3 hours in the past 24) as attributes to train and test the model.

C. Training and Testing

To predict the time of daily peaks of power demand 24 hours ahead, we applied several classification techniques to classify the cases as peak or non-peak. The methods that have been used in our study include Naïve Bayes, SVM, Random Forest, AdaBoost, CNN, LSTM, and Stacked Autoencoder.

TABLE I: Continuous attributes for forecasting models

Continuous Attributes	**Description or Formula**
Hourly power demand	L_h
Hourly temperature	T_h
Hourly humidex	H_h
Hourly windchill	W_h
Hourly relative humidity	RH_h
Load in the same hour over previous 2 weeks	$L_{h-i*24}, i = 1, ..., 14$
Ave. of power demands over 3-hour intervals	$mean\{L_{h+2-i}, L_{h+1-i}, L_{h-i}\}$, $i = 3, 6, 9, 12, 15, 18, 21, 24$
Ave. of temperature over 3-hour intervals	$mean\{T_{h+2-i}, T_{h+1-i}, T_{h-i}\}$, $i = 3, 6, 9, 12, 15, 18, 21, 24$
Ave. of humidex over 3-hour intervals	$mean\{H_{h+2-i}, H_{h+1-i}, H_{h-i}\}$, $i = 3, 6, 9, 12, 15, 18, 21, 24$
Ave. of relative humidity over 3-hour intervals	$mean\{RH_{h+2-i}, RH_{h+1-i}, RH_{h-i}\}$, $i = 3, 6, 9, 12, 15, 18, 21, 24$

When training the model, the categorical variables have been converted into binary code through the technique of one-hot encoding. According to the date, the attributes and targets have been divided into two datasets for summer and winter. Two separate models were built on corresponding datasets. The training dataset consists of 5-year data which is from fiscal year 2003 to 2007, while the data for fiscal year 2008 has been used for the testing. The testing dataset for both summer and winter contains 183 days with 4392 samples.

As classical supervised learning techniques, Naïve Bayes, Random Forest, AdaBoost and SVM are very popular to be used to solve classification problem. Naïve Bayes uses Bayesian theorem to estimate possibilities and then decide the class for each sample [22]. Random Forest, an ensemble method, consists of a great number of decision trees, and its output is the mode of categories decided by each tree [23]. AdaBoost constructs a strong classifier using an iteratively generated set of weak classifiers [24]. SVM classifier searches separating hyperplanes to find the max-margin hyperplane to

TABLE II: Discrete attributes for forecasting models

Discrete Attributes	Description or Formula
Weekday	Working Monday, working Tuesday, working Wednesday, working Thursday, working Friday, holiday or weekend
Sunrise	1: sunrise happened; 0: sunrise did not happen
Sunset	1: sunset happened; 0: sunset did not happen
Hour of the day	Morning: 7 a.m. to 12 p.m.; Afternoon: 1 p.m. to 6 p.m.; Evening: 7 p.m. to 12 a.m.; Night: 1 a.m. to 6 a.m.

separate and label the samples [25]. This study works on SVM with Radial Basis Function (RBF) kernel.

CNNs are a powerful method for classification [26]. Through convolution layers and pooling layers, it learns the hidden features to classify the data. In this study, we choose the CNN that has 3 hidden layers with 64, 64, 128 filters respectively. The sequences of attributes have been transferred into matrices as the inputs. Since the hourly data used in this study can also be treated as time series, LSTM has been applied to solve the prediction problem [27]. LSTM is a special type of recurrent neural network that generally works very well on time series data, as LSTM learns from the previous long-term information to make the decision. The LSTM model here has 6 hidden layers with 24 neurons each. We set its lookback length as 24 cases, which is for one day. As a kind of neural network, Stacked Autoencoder (SAE) is very useful for extracting representations for datasets [28]. It can transform the data with high dimension to the data with low dimension. The Stacked Autoencoder we built for this study has 15 hidden layers with 128 units each hidden layers.

Because the rate of number of peaks to the number of non-peaks is 1:23, the model is likely to classify the daily peak hour into non-daily-peak hour, which can cause the probability of being peak hour be very low. To address it while keep the same order and range of the probabilities, we used the equation:

$$P_{new} = 0.1 \times \sqrt{P \times 100} \tag{1}$$

to re-calculate the possibilities, where P represent the probability from the classifiers. Currently, we set 0.5 as the threshold of possibilities here, but some heuristic methods can be used to determine it. Then, if the value of probability is larger than 0.5, the load peak will be at the same hour next day. Otherwise, the load peak hour for tomorrow should be another time.

V. Results

The performance of each model was evaluated by precision, recall and accuracy, which are based on the number of true positives (TP), the number of false positives (FP), number of false negatives (FN), and number of true negatives (TN). Precision is computed on TP and FP, which also be named as positive predictive value,

$$Precision = \frac{TP}{TP + FP}. \tag{2}$$

Recall, also referred to as sensitivity, is

$$Recall = \frac{TP}{TP + FN}. \tag{3}$$

In the context of this paper, the sum of TP and FN should equal to the number of actual daily peaks. Therefore, for the testing dataset, $TP + FN = 183$. Accuracy is also a measure that used to evaluate the classification methods and is calculated as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{4}$$

The sum of TP, TN, FP and FN is the number of total samples; thus, for the testing dataset, $TP+TN+FP+FN = 4392$. For each model we report the TP, FP, precision, recall, accuracy in Table III and Table IV.

TABLE III: Classification results for winter

Methods	# TP	# FP	Precision	Recall	Accuracy
Naïve Bayes	164	707	0.19	0.90	0.83
SVM	146	104	0.58	0.80	0.97
Random Forest	133	84	0.61	0.73	0.97
AdaBoost	177	786	0.18	0.97	0.82
CNN	148	100	0.60	0.81	0.97
LSTM	160	76	0.68	0.87	0.98
SAE	127	65	0.66	0.69	0.97

TABLE IV: Classification results for summer

Methods	# TP	# FP	Precision	Recall	Accuracy
Naïve Bayes	145	1434	0.09	0.79	0.66
SVM	43	87	0.33	0.23	0.95
Random Forest	53	120	0.31	0.29	0.94
AdaBoost	177	1641	0.01	0.97	0.63
CNN	92	240	0.28	0.50	0.92
LSTM	81	111	0.42	0.44	0.95
SAE	73	142	0.34	0.40	0.94

From Table III, the precision ranges from 0.18 to 0.68, the recall has the range of 0.69 to 0.97, while accuracy is from 0.82 to 0.98. The method of AdaBoost has the highest recall with the value of 0.97, but its precision is the lowest with the value of 0.18, which means the number of FP is very high. It has five methods, SVM, Random Forest, CNN, LSTM and Stacked Autoencoder, whose accuracy is larger than 0.90. Comparing three indicators, precision, recall and accuracy, the performance of LSTM is better than other methods.

From Table IV, as expected, the values of three indicators are significantly less than those for the winter, which means the result for winter is much better than the result for summer. The highest precision is 0.42, the value of recall can up to 0.97, while accuracy has the largest value of 0.95. Similarly, the highest recall is from AdaBoost, but it has low values of precision and accuracy.

Each method has different advantages or disadvantages. For example, Naïve Bayes has the best performance in terms of time complexity, while the number of TP for Adaboost is larger than others. LSTM has a longer training time, but boast top overall performance. In general, using classification models to solve the prediction of daily load peak hour is effective.

VI. Conclusions

In this study, we explored the use of classification methods to solve the problem of daily load peak hour prediction in Ontario, Canada. Separate models have been built for summer and winter, due to different distributions of attributes. A five-year data has been used to train the models, while the testing was on one-year data. The experiments exhibit the promising results that the time of daily load peak can be located effectively.

In the future, we will use more historical data and attributes for model training and testing. We are going to study how to decrease the number of false positives and then how to improve the overall performance of models, especially for the summer models. Finally, we intend to extend these methods to examine the problem of identifying the days and hours of the five coincident peaks in each fiscal year.

References

[1] B. Ryu, T. Makanju, A. Lasek, X. An, and N. Cercone, "A naive bayesian classification model for determining peak energy demand in ontario," in *Smart City 360*, 2016, pp. 517–529.

[2] A. El Desouky and M. El Kateb, "Hybrid adaptive techniques for electric-load forecast using ann and arima," *IEE Proceedings-Generation, Transmission and Distribution*, vol. 147, no. 4, pp. 213–217, 2000.

[3] E. El-Attar, J. Goulermas, and Q. Wu, "Forecasting electric daily peak load based on local prediction," in *Power & Energy Society General Meeting, 2009. PES'09. IEEE*. IEEE, 2009, pp. 1–6.

[4] L. M. Saini, "Peak load forecasting using bayesian regularization, resilient and adaptive backpropagation learning based artificial neural networks," *Electric Power Systems Research*, vol. 78, no. 7, pp. 1302–1310, 2008.

[5] J. V. Ringwood, D. Bofelli, and F. T. Murray, "Forecasting electricity demand on short, medium and long time scales using neural networks," *Journal of Intelligent and Robotic Systems*, vol. 31, no. 1-3, pp. 129–147, 2001.

[6] P. E. McSharry, S. Bouwman, and G. Bloemhof, "Probabilistic forecasts of the magnitude and timing of peak electricity demand," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1166–1172, 2005.

[7] M. Goodwin and A. Yazidi, "A pattern recognition approach for peak prediction of electrical consumption," in *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, 2014, pp. 265–275.

[8] C. Sigauke and D. Chikobvu, "Prediction of daily peak electricity demand in south africa using volatility forecasting models," *Energy Economics*, vol. 33, no. 5, pp. 882 – 888, 2011. [Online]. Available: https://doi.org/10.1016/j.eneco.2011.02.013

[9] K. Livik, N. Feilberg, and J. A. Foosnaes, "Estimation of annual coincident peak demand and load curves based on statistical analysis and typical load data," in *12th International Conference on Electricity Distribution, 1993. CIRED*, May 1993, pp. 6.20/1–6.20/6 vol.6.

[10] Z. Liu, A. Wierman, Y. Chen, B. Razon, and N. Chen, "Data center demand response: Avoiding the coincident peak via workload shifting and local generation," *Performance Evaluation*, vol. 70, no. 10, pp. 770 – 791, 2013. [Online]. Available: https://doi.org/10.1016/j.peva.2013.08.014

[11] M. Uddin, M. F. Romlie, M. F. Abdullah, S. A. Halim, A. H. A. Bakar, and T. C. Kwang, "A review on peak load shaving strategies," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 3323 – 3332, 2018. [Online]. Available: https://doi.org/10.1016/j.rser.2017.10.056

[12] C. Mahapatra, A. K. Moharana, and V. C. M. Leung, "Energy management in smart cities based on internet of things: Peak demand reduction and energy savings," *Sensors*, vol. 17, no. 12, 2017. [Online]. Available: http://www.mdpi.com/1424-8220/17/12/2812

[13] K. Gajowniczek and T. Zabkowski, "Two-stage electricity demand modeling using machine learning algorithms," *Energies*, vol. 10, no. 10, 2017. [Online]. Available: http://www.mdpi.com/1996-1073/10/10/1547

[14] Y. H. Jiang, R. Levman, L. Golab, and J. Nathwani, "Predicting peak-demand days in the ontario peak reduction program for large consumers," in *Proceedings of the 5th international conference on Future energy systems*. ACM, 2014, pp. 221–222.

[15] X. An and N. Cercone, "A statistical model for predicting power demand peaks in power systems," in *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, Aug 2015, pp. 1022–1026.

[16] "IESO data directory," 2018, accessed: Sept. 14, 2018. [Online]. Available: http://www.ieso.ca/power-data/data-directory

[17] "Environment canada: Climate data - environment canada," 2018, accessed: Sept. 14, 2018. [Online]. Available: http://climate.weather.gc.ca/

[18] "Time and date," 2018, accessed: Sept. 14, 2018. [Online]. Available: https://www.timeanddate.com/worldclock/canada/toronto

[19] K. X. Perez, K. Cetin, M. Baldea, and T. F. Edgar, "Development and analysis of residential change-point models from smart meter data," *Energy and Buildings*, vol. 139, pp. 351–359, 2017.

[20] M. T. Paulus, D. E. Claridge, and C. Culp, "Algorithm for automating the selection of a temperature dependent change point model," *Energy and Buildings*, vol. 87, pp. 95–104, 2015.

[21] J. Xie, Y. Chen, T. Hong, and T. D. Laing, "Relative humidity for load forecasting models," *IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 191–198, 2018.

[22] D. J. Hand and K. Yu, "Idiot's bayes: Not so stupid after all?" *International Statistical Review / Revue Internationale de Statistique*, vol. 69, no. 3, pp. 385–398, 2001. [Online]. Available: http://www.jstor.org/stable/1403452

[23] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct 2001. [Online]. Available: https://doi.org/10.1023/A:1010933404324

[24] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*, ser. ICML'96. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1996, pp. 148–156. [Online]. Available: http://dl.acm.org/citation.cfm?id=3091696.3091715

[25] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, ser. COLT '92. New York, NY, USA: ACM, 1992, pp. 144–152. [Online]. Available: http://doi.acm.org/10.1145/130385.130401

[26] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov 1998.

[27] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997. [Online]. Available: http://dx.doi.org/10.1162/neco.1997.9.8.1735

[28] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *Journal of Machine Learning Research*, vol. 11, pp. 3371–3408, Dec. 2010. [Online]. Available: http://www.jmlr.org/papers/v11/vincent10a.html