Machine Learning Techniques for Short-Term Electric Load Forecasting

Praxitelis - Nikolaos Kouroupetroglou, MSc Student
Supervisor: Tsoumakas Grigorios, Assistant Professor
Master Studies in: Knowledge, Data and Software Technologies
School of Informatics, Aristotle University of Thessaloniki

Abstract

- This presentation shows the results of master thesis: applying Machine Learning Models for the Short-term Load forecasting in the Greek Electric Network.

Outline

1. Introduction
2. Theoretical Background and Data Science Applications Development Platforms
3. Review from Scientific Literature and Commercial Software
4. Implementation - ML Models for STLF in Greek Electric Grid
5. Conclusions - Remarks
6. Future Work
7. Papers

Introduction

What is the problem and Why it is important

- Nowadays the economic and social development depends on the abundance of electric power.
- The electric load forecasting provides various benefits for an electrical company and its business needs [1, 2, 3]:
    a. For purchasing and producing electric power.
    b. For transmitting, transferring and distributing electric power.
    c. For managing and maintaining the electric power sources.
    d. For managing the daily electric load demand.
    e. For economic and marketing planning.
- *electric load forecasting is defined as the methodology for the load forecasting for a specific duration/horizon* [1, 2, 3].

Electric Load Forecasting per business needs [2]

Business Needs / Forecasting Factors	Minimum updating cycle	Max forecasting horizon
purchasing and producing electric power	1 hour	10 years and above
transmitting, transferring and distributing electric power.	1 day	30 years
managing and maintaining the electric power sources.	15 minutes	2 weeks
managing the daily electric load demand	15 minutes	10 years and above
financial and marketing	1 month	10 year and above

Types of Electric Load Forecasting [1, 2]

- Very Short-term Load Forecasting
  - small forecast, some hours only, usually utilizes previous hourly loads and climate factors partially
- Short-term Load Forecasting
  - forecast for 24-hours to 2 weeks forecast, climate factors is used at most
- Medium-term Load Forecasting
  - forecast for 1 month to 3 years, climate factors are inaccurate, economic factor and land use.
- Long-term Load Forecasting
  - forecast for 3 years to 10 years, simulations of climate factors and economic transactions are used.

Availability of factors in load forecasting [2]

Factors / forecasting accuracy	Accurate	Inaccurate	Unreliable
Climate Factors	1 day	2 weeks	> 2 weeks
Economics	1 month	3 years	> 3 years
Land Use	1 year	5 years	> 5 years

Classification of Load Forecasting [1, 2]

Business Needs / Forecasting Factors	Climate Factors	Economics	Land Use	Updating Cycle	Horizon
Very Short-term Load Forecasting	Optional	Optional	Optional	<= 1 Hour	1 Day
Short-term Load Forecasting	Required	Optional	Optional	1 Day	2 weeks
Medium-term Load Forecasting	Simulated	Required	Optional	1 Month	3 years
Long-term Load Forecasting	Simulated	Simulated	Required	3 Years	30 years

Applications of different types of electric load forecasting in Business Needs [2]

Business Needs / types of electric Load forecasting	Very Short-term Load Forecasting	Short-term Load Forecasting	Medium-term Load Forecasting	Long-term Load Forecasting
purchasing and producing electric power	X	X	X	X
transmitting, transferring and distributing electric power.	X	X	X	X
managing and maintaining the electric power sources.	X	X		
managing the daily electric load demand	X	X	X	X
financial and marketing			X	X

Theoretical Background, Data Science Applications Development Platforms

What is Machine Learning

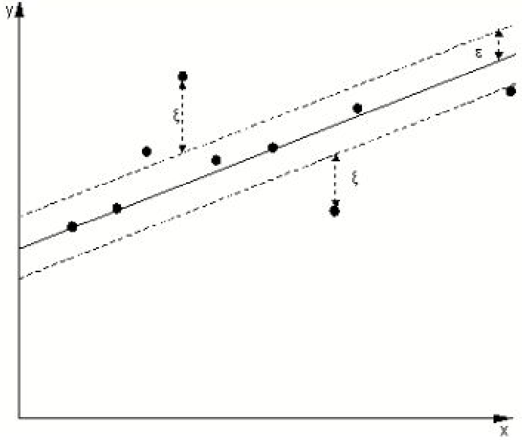
- Is the science which enables computers to get into a mode of self-learning of a dataset without being explicitly programmed.
- *“Machine Learning is the process which, a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.” - Tom Mitchell - “machine Learning” book*

Types of Machine Learning

- Machine learning tasks are typically classified into four broad categories based on the nature of learning
  a. Supervised Learning
  b. Unsupervised Learning
  c. Semisupervised Learning
  d. Reinforcement Learning
- Based on learning task
  a. Regression
  b. Classification
  c. Clustering

Theoretical Background (1) - SVM

- SVM can be applied not only to classification problems but also to the case of regression.
- In the same way as with classification approach there is motivation to seek and optimize the generalization bounds given for regression.



Theoretical Background (2) - Random Forest

- Random forest combine the predictions of multiple decision trees.
- Multiple decision trees are being constructed, by randomizing the combination and order of features used.
- Aggregated result from these forest of trees would form an ensemble, known as a random forest.

Theoretical Background (3) - k-Nearest Neighbors

- predict the numerical target based on a similarity measure (e.g., distance functions).
- A simple implementation of KNN regression is to calculate the average of the numerical target of the K nearest neighbors.

Distance functions

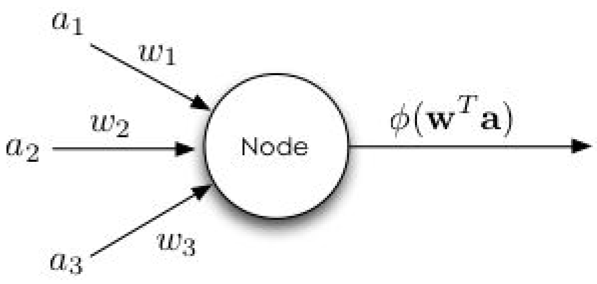
Euclidean $$\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$$

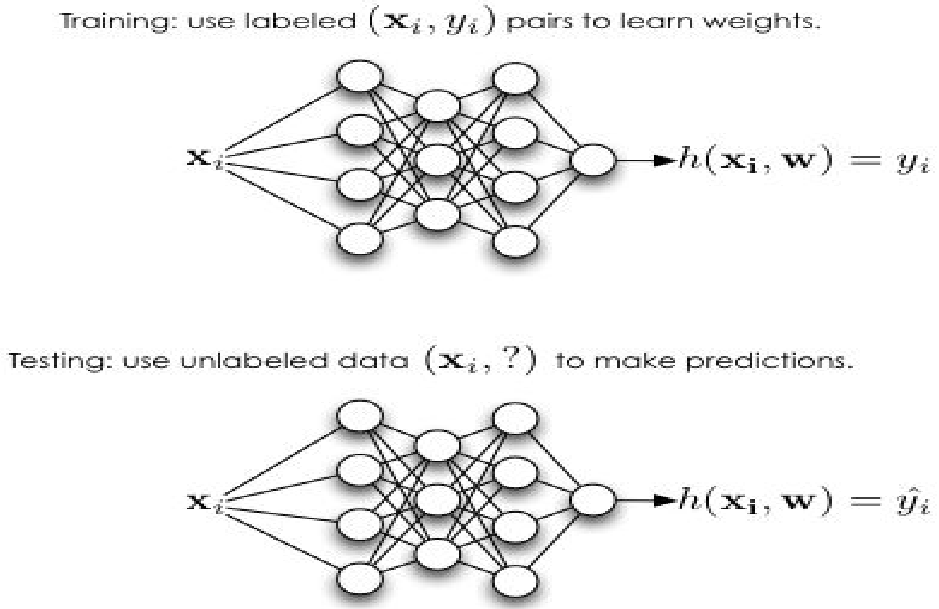
Manhattan $$\sum_{i=1}^{k}|x_i - y_i|$$

Minkowski $$\left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{1/q}$$

Theoretical Background (4) - Neural Networks

- Artificial neural networks (ANNs) were originally devised in the mid-20th century as a computational model of the human brain.



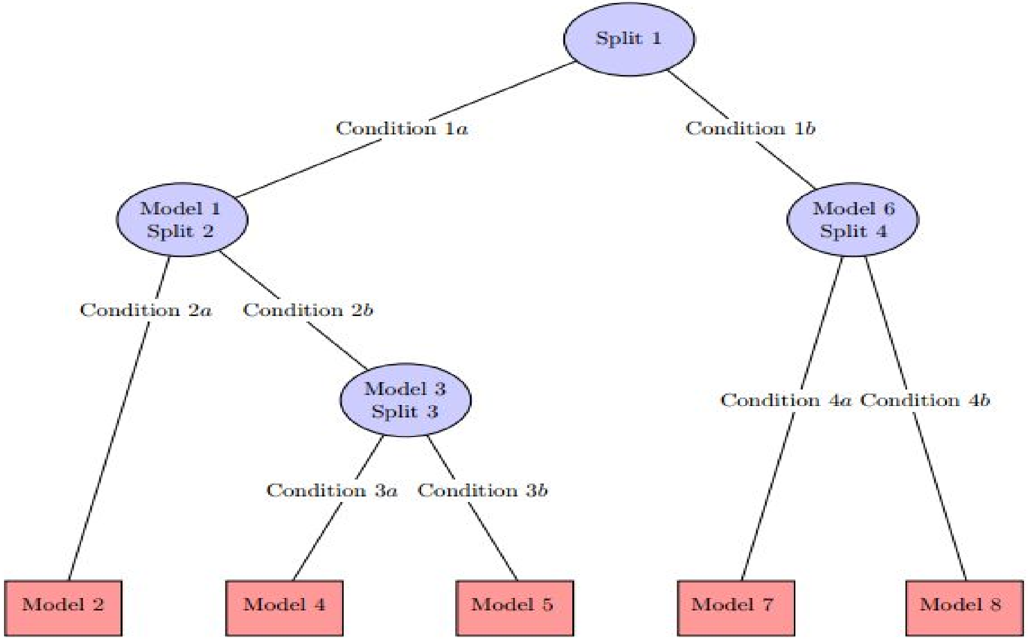


Theoretical Background (5) - XGBoost

- "Extreme Gradient Boosting", is an open-source software library which provides the gradient boosting framework.
- XGBoost is focused on computational speed and model performance. It supports the gradient boosting forms of:
    a. Gradient Boosting algorithm also called including the learning rate.
    b. Stochastic Gradient Boosting
    c. Regularized Gradient Boosting with both L1 and L2 regularization.

Theoretical Background (6) - Model Trees

- Very similar to regression trees except that the terminal nodes - leaves, predict the outcome using a linear model.



Data Analysis and Machine Learning Development Platforms

- R
- Python
  a. Scikit Learn
  b. Anaconda
- Weka
- Knime
- Rapidminer
- Orange
- SPSS

Review from Scientific Literature and Commercial Software

Scientific Review (1)

- Load Forecasting - Statistical Approaches
  a. Regression Analysis [4, 5]
  b. Time Series Analysis
     - ARIMA [6, 8]
     - Time Series Decomposition - Exponential Smoothing [7, 9]

Scientific Review (2)

- Load Forecasting Artificial Intelligence / Machine Learning Approaches
  a. Neural Networks [10 - 12]
  b. Deep Learning [15 - 16]
  c. SVM [13 - 14]
  d. Random Forests [17, 18]
  e. Model / Regression Trees [19, 20]
  f. k-Nearest Neighbors [21, 22]
  g. XGBoost [23, 24]

Commercial Software

- ETAP's load forecasting software
- Aiolos Forecasting Studio
- GMDH Shell's Electric Load Forecasting
- Escoware® demand-forecasting
- SAS® Energy Forecasting
- Statgraphics

Implementation - ML Models for STLF in Greek Electrical Grid

Goals - Development Platform - Software Environment

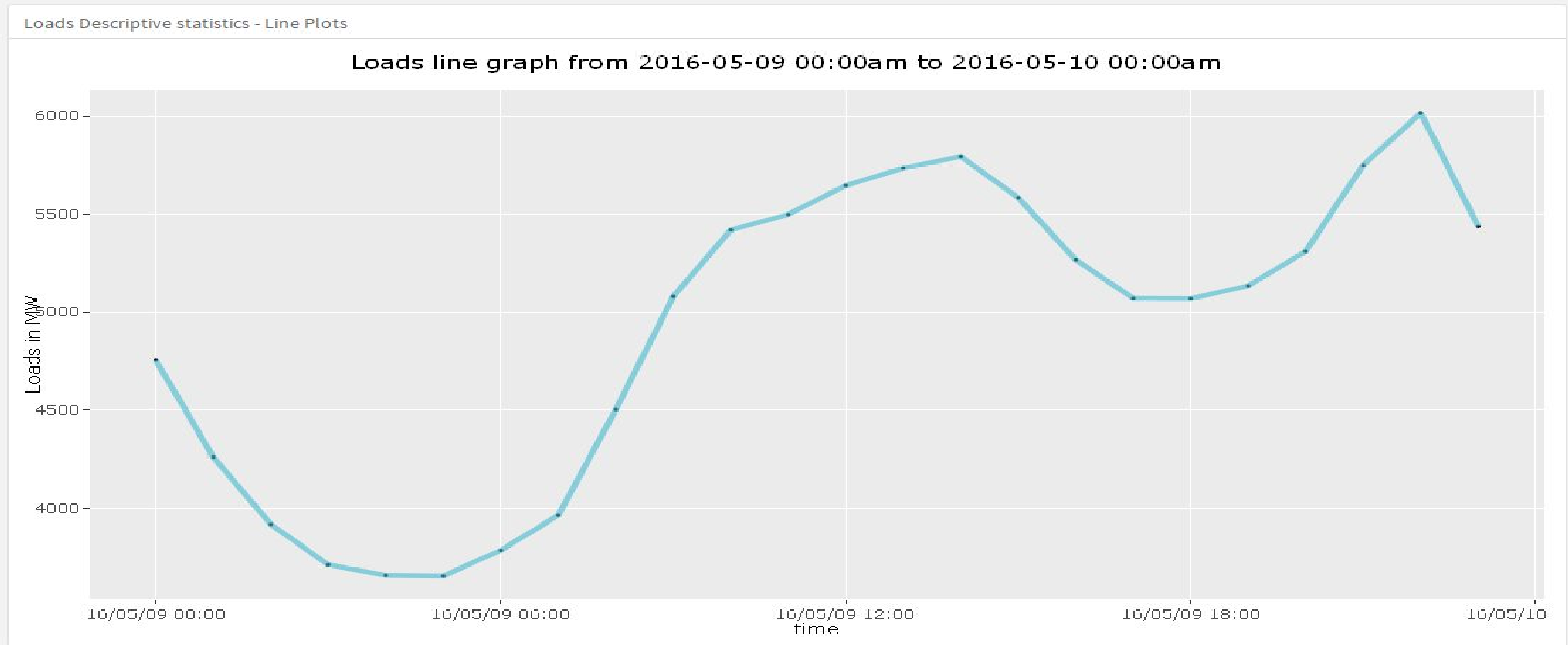
- An R application was developed
    a. R 3.4.1
    b. R-Studio
    c. R-markdown
    d. R-Shiny
    e. FlexDashboard
    f. Shinyapps (ipto-ml)
- Goal: Accurate ML models with small prediction error.

Datasets

- The Datasets come from 3 origins.
    a. IPTO Loads (Independent Power Transmission Operator in Greece - ΑΔΜΗΕ in greek)
    b. OoEM predictions (Operator of Electricity Market in Greece - ΛΑΓΗΕ in greek, existed predictions)
    c. Meteorological Data (extracted from Dark Sky API)
        - Meteorological Data from Athens and Thessaloniki was extracted [3].
- The Datasets span from 2010-10-01 to 2017-04-30

Electricity Load Demand Plot

- A typical daily Load line graph:



Meteorological Features

- Meteorological Features taken from DarkSky API:
    a. Summary
    b. Icon
    c. Temperature
    d. Dew Point
    e. Humidity
    f. Wind Speed
    g. Wind Bearing
    h. Visibility
    i. Cloud cover
    j. UV index

Preprocessing - Cleaning the Data

- Filling the missing values
    a. Filling NA values from features with the mean from adjacent hour, day(s).
- Convert the hourly dataset to daily

Calendar Features

- Calendar features were constructed from the dates of the meteorological dataset.
- The reason behind is that specific hours and specific days affect the human behavior beyond meteorological factors.
- Hence the following features were constructed:
    - isHoliday
    - isRushHour
    - isWeekend

Joining the Datasets and Constructing the Cases (1)

- Inner Join between IPTO and meteorological/calendar datasets with “date-time” as common field.
- 1484 features for 2394 cases for daily basis data was created from hourly IPTO and DarkSky hourly datasets.

Joining the Datasets and Constructing the Cases (2)

Case and its Features	
1	Weather forecast in Athens for the forecasted day(d).
2	Weather forecast in Thessaloniki for the forecasted day(d).
3	Calendar Features for the forecasted day(d).
4	Weather data in Athens for preceding day(d-1) of the forecasted day(d).
5	Weather data in Thessaloniki for preceding day(d-1) of the forecasted day(d).
6	Calendar Features for preceding day(d-1)
7	hourly load measurements of the second preceding day (d-2) of the forecasted day(d).
8	hourly load measurements of the second preceding day (d-3) of the forecasted day(d).
9	Current Loads for the forecasted day (d)

Adding Noise

- Dark Sky API provides meteorological data for a location / city with the aid of nearby weather or from its own ML models.
- Trying to be as precise and close to the real weather conditions for a location we added a moderate small noise (+/- 0.5 from uniform distribution) to the extracted meteorological data taken from DarkSky API.

Feature Selection

- A single case consists of 1484 features for 2394 cases, each case contains load.x target variables.
- Feature selection per target variable was applied to reduce the number of features.
- Feature selection R package: "Boruta"
- Use of Random Forest importance measure to find the most relevant features per target variable.

Machine Learning - Experiments (1)

- 6 Machine Learning algorithms was applied
  a. SVM - e1071 package
  b. k-Nearest Neighbors - FNN package
  c. Random Forest - randomForest package
  d. Neural Networks - RSNNS package
  e. XGBoost - xgboost package
  f. Model Trees - cubist package
- 24 models per target variable load.x for each Machine Learning algorithm were developed

Machine Learning - Experiments (2)

- Dataset partition:
- The Datasets span from 2010-10-01 to 2017-04-30.
  - Train Set: spans from *2010-01-05 to 2015-04-30*.
  - Validation Set: spans from *2015-05-01 to 2016-04-30*.
  - Test Set: spans from *2016-05-01 to 2017-04-30*.
- Evaluation Metrics: *Mean absolute Percent Error, MAPE*
- Parameters: *Default ML Parameters, Grid Search, Full Features, Feature Selection.*
- Ensemble models per target variable (if applicable)

Machine Learning - Experiments (3)

- Machine Learning Experiments:
  - *Experiment #1*: training and testing ML models with full features and default ML parameters.
  - *Experiment #2*: training and testing ML models with feature selection and default ML parameters.
  - *Experiment #3*: training and testing ML models with full features and grid search.
  - *Experiment #4*: training and testing ML models with feature selection and grid search.

Machine Learning - Experiments - Results (1)

- *Experiment #1*: training and testing ML models with full features and default ML parameters - results:

classifier selection from target variable ensemble, mape evaluation and performance evaluation for experiment #1	
Model Trees mean mape:	2.65%
OoEM existed model mape:	2.53%
Performance evaluation	-4.74%

Machine Learning - Experiments - Results (2)

- *Experiment #2*: training and testing ML models with feature selection and default ML parameters - results:

classifier selection from target variable ensemble, mape evaluation and performance evaluation for experiment #2	
SVM, Model Trees mean mape:	2.55%
OoEM existed model mape:	2.53%
Performance evaluation	-0.78%

Machine Learning - Experiments - Results (3)

- *Experiment #3*: training and testing ML models with full features and Grid Search - results:

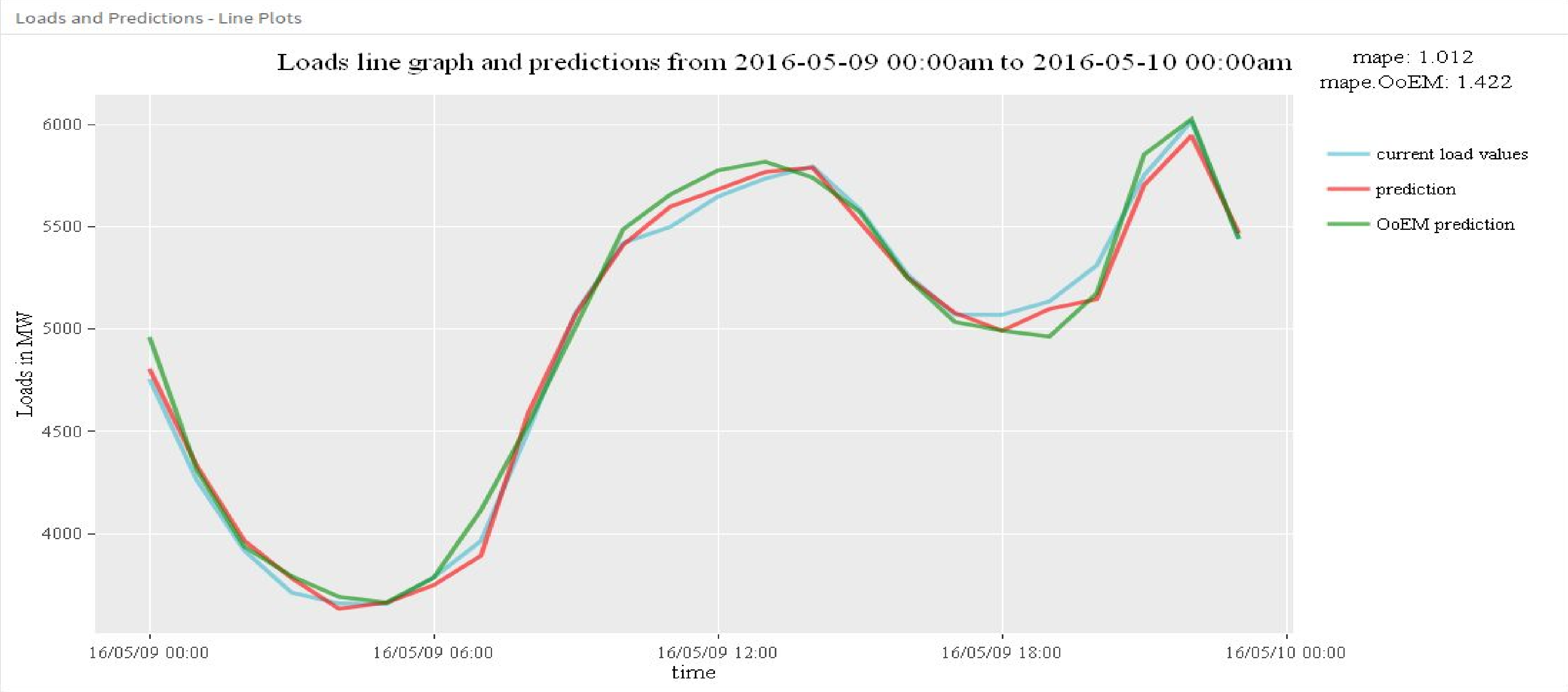
classifier selection from target variable ensemble, mape evaluation and performance evaluation for experiment #3	
XGBoost, Model Trees mean mape:	2.58%
OoEM existed model mape:	2.53%
Performance evaluation	-1.97%

Machine Learning - Experiments - Results (4)

- *Experiment #4*: training and testing ML models with feature selection and Grid Search - results:

classifier selection from target variable ensemble, mape evaluation and performance evaluation for experiment #4	
SVM, XGBoost, Model Trees mean mape:	2.41%
OoEM existed model mape:	2.53%
Performance evaluation	+4.74%

Machine Learning - Experiments - Results (5)



Conclusions - Remarks

Conclusions - Final Remarks

- Electric load forecasting is a vital process, due to the fact that the increased needs for electrical power from consumers.
- Electrical companies should take into account this process and constantly improve their models for more accurate forecasts.
- 4 experiments with 6 machine learning algorithms were developed with best prediction error: *2.41%*
- Classifier selection via model ensemble per target variable
- A shiny application was developed: ipto-ml

Future Work

Future Work

- Adding more Greek cities to the cases
- On the fly extract all datasets
- Short-term load forecasting for every future date
- A shiny application for daily use, an energy analytics application.

Papers

Papers (1)

[1]: Electrical load forecasting : modeling and model construction / Soliman Abdel-hady Soliman (S.A. Soliman), Ahmad M. Al-Kandari.
[2]: Short Term Electric Load Forecasting by Tao Hong, A dissertation submitted to the Graduate Faculty of North Carolina State University
[3]: A. Apostolos, "Short-Term Load Forecasting using a Cluster of Neural Networks for the Greek Energy Market", Electrical & Computer Engineer Public Power Corporation Greece
[4]: A. D. Papalexopoulos and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," IEEE Transactions on Power Systems.
[5]: T. Haida and S. Muto, "Regression based peak load forecasting using a transformation technique," IEEE Transactions on Power Systems.
[6]: B. Krogh, E. S. de Llinas, and D. Lesser, "Design and Implementation of An on-Line Load Forecasting Algorithm," IEEE Transactions on Power Apparatus and Systems.

Papers (2)

[7]: M. T. Hagan and S. M. Behr, "The Time Series Approach to Short Term Load Forecasting," IEEE Transactions on Power Systems.
[8]: N. Amjady, "Short-term hourly load forecasting using time-series modeling with peak load estimation capability," IEEE Transactions on Power Systems.
[9]: D. C. Park, M. A. El-Sharkawi, R. J. Marks, II, L. E. Atlas, and M. J. Damborg, "Electric load forecasting using an artificial neural network," IEEE Transactions on Power Systems
[10]: D. K. Ranaweera, N. F. Hubele, and A. D. Papalexopoulos, "Application of radial basis function neural network model for short-term load forecasting," IEEE Proceedings - Generation, Transmission and Distribution
[11]: Sp.Kiartzis, Artificial Intelligence Applications in Short-term Load Forecasting, PhD Dissertation, AUTH
[12]: G.Bakirtzis, V.Petridis, S.J.Kiartzis, M.C.Alexiadis, A.H.Maissis, A neural network short term load forecasting model for the Greek power system

Papers (3)

[13]: N. Sapankevych and R. Sankar, "Time Series Prediction Using Support Vector Machines: A Survey," IEEE Computational Intelligence Magazine, vol. 4, pp. 24-38, 2009

[14]: Electrical Load Forecasting using Support Vector Machines, Belgin Emre, Dilara Demren, Instanbul Technical University, Turkey.

[15]: Ensemble Deep Learning for Regression and Time Series Forecasting, Xueheng Qiu, Le Zhang, Ye Ren and P. N. Suganthan, School of Electrical and Electronic Engineering, Nanyang Technological Univeristy, Singapore, Gehan Amaratunga, Department of Engineering, University of Cambridge, UK.

[16]: Deep Neural Network Based Demand Side Short Term Load Forecasting, Seunghyoung Ryu, Department of Electronic Engineering, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul 121-742, Korea

Papers (4)

[17]: Short Term Electrical Load Forecasting Using Mutual Information Based Feature Selection with Generalized Minimum-Redundancy and Maximum-Relevance Criteria, Nantian Huang, Zhiqiang Hu, Guowei Cai and Dongfeng Yang, School of Electrical Engineering, Northeast Dianli University

[18]: Short-Term Load Forecasting using Random Forests, Grzegorz Dudek, Department of Electrical Engineering, Czestochowa University of Technology, Czestochowa, Poland

[19]: Rule-Based Prediction of Short-term electric Load, Petr Berka, Prof, PhD, University of Economics Prague \& University of Finance and Administration Prague, Czech Republic

[20]: Formulation and analysis of a rule-based short-term load forecasting algorithm, S. Rahman, Dept of Electr. Eng., Virginia Polytech. Inst. & State Univ. Blacksburg, VA, USA

Papers (5)

[21]: A Novel Hybrid Model Based on Extreme Learning Machine, k-Nearest Neighbor Regression and Wavelet Denoising Applied to Short-Term Electric Load Forecasting, Weide Li, Demeng Kong and Jinran Wu, School of Mathematics and Statistics, Lanzhou University, Gansu, China
[22]: A composite k-nearest neighbor model for day-ahead load forecasting with limited temperature forecasts, Rui Zhang, Yan Xu, Zhao Yang Dong, Weicong Kong, Kit Po Wong, School of Electrical and Information Engineering, University of Sydney, NSW 2006, Australia
[23]: Short-Term Electricity Load Forecasting Based on the XGBoost Algorithm, Guangye Li, Wei Li, Xiaolei Tian, Yifeng Che, State Grid Liaoning Electric Power Co., Ltd., Shenyang Liaoning
[24]:A gradient boosting approach to the Kaggle load forecasting competition, Souhaib Ben Taieb, Department of Computer Science, Department of Econometrics and Business Statistics, Monash University, Australia

Machine Learning Techniques for Short-Term Electric Load Forecasting

Thank you for listening and for your time

Praxitelis - Nikolaos Kouroupetroglou