Introduction to Forecasting in the Utilities (202)

Description

Continued population growth, socioeconomic improvements, and technological advancements in the past few decades have caused a significant rise in the consumption of energy and materials. Many utilities find themselves concerned -- the volatility of wind and solar power generation, the uncertainty of rooftop solar adoption, and rising gas and electricity prices pose serious challenges. The modern consumer-centric paradigm of transactive energy has changed the traditional load forecasting methodologies, as it evolves and reshapes utility strategies.

This training intends to provide a comprehensive introduction to forecasting methods and present enough information about each method for participants to use them sensibly. Examples and applications from the utility industry, including forecasting with AMI data, are included.

An outline for the training, *Introduction to Forecasting in the Utilities*, has been included with this document.

Audience

This training is intended for the following audiences:

1. Analytics professionals who are interested in learning forecasting methods with applications in utilities.
2. Utility professionals who find themselves doing forecasting without prior formal training.
3. *Utility Analytics 101* completers who want to continue advancing their analytics knowledge in the utility setting.
4. Positions include, but are not limited to, Data Scientists, Forecasting Analysts, Energy Analysts, and Research Analysts.

Prerequisites

College- or university-level statistics and algebra or equivalent experience. Some exposure to statistical programming (for example, in Python or R language) is helpful but not required.

Objectives

Upon completion of this training, students will be able to:

1. Understand select applications of time series forecasting within the utility sector.
2. Use statistical and graphical approaches to exploratory data analysis with time series data.
3. Use software and/or programming languages (e.g., Python or R) to create statistical forecasts.
4. Develop load, price, wind power, and/or solar power forecasts.

Detailed Outline

Prework: Learn to use Google collab for Python Scripting

Day One (Introducing Time Series and Python for Time Series)

Introducing Time Series

1. What is a time series?
    a. Types of time series
    b. Main areas of application for time series analysis
2. Data-Generating Process (DGP)
    a. Generating synthetic time series
    b. Stationary and non-stationary time series
3. What do we forecast?
4. Forecasting terminology

Acquiring and Processing Time Series Data

1. Understanding the time series data set
    a. Preparing a data model
2. Smart Meter Data
3. Python Refresher: pandas, datetime operations, indexing, and slicing
    a. Converting the date columns into pd.Timestamp/DatetimeInde
    b. Slicing and Indexing
    c. Creating date sequences and managing date offsets
4. Handling missing data
    a. Compact, expanded, and wide forms of data
    b. Regular intervals in time series
    c. Imputing missing data
    d. Seasonal interpolation

Day Two (Visualizing Time Series and Forecasting)

Analyzing and Visualizing Time Series Data

1. Components of a time series
    a. The trend component
    b. The seasonal component
    c. The cyclical component
    d. The irregular component
2. Visualizing time series data
    a. Line charts, seasonal plots
    b. Seasonal box plots
    c. Calendar heatmaps
    d. Autocorrelation plot
3. Decomposing a time series
    a. Detrending
    b. Deseasonalizing

    c. Implementations
4. Detecting and treating outliners
    a. Standard Deviation
    b. Interquartile range
    c. Isolation Forest
    d. Extreme studentized deviate (ESD) and seasonal ESD (S-ESD)
    e. Treating outliners

Setting a Strong Baseline Forecast

1. Setting up a test harness
    a. Creating holdout (test) and validation datasets
    b. Choosing an evaluation metric
2. Generating strong baseline forecasts
    a. Naïve forecast
    b. Moving average forecast
    c. Seasonal naïve forecast
    d. Exponential smoothing (ETS)
    e. ARIMA
    f. Theta Forecast
    g. Fast Fourier Transform forecast
    h. Evaluating the baseline forecasts
3. Assessing the forecastability of a time series
    a. Coefficient of Variation (CoV)
    b. Residual variability (RV)
    c. Entropy-based measures
    d. Kaboudan Metric

Time Series Forecasting as Regression

1. Understanding the basics of machine learning
    a. Supervised machine learning tasks
    b. Overfitting and underfitting
    c. Hyperparameters and validation sets
2. Time series forecasting as regression
    a. Time delay embedding
    b. Temporal embedding
3. Introducing Global Forecasting Models
    a. Why Global Forecasting Models
    b. Sample Size
    c. Cross-learning

Day 3 (Machine Learning with Time Series)

Feature Engineering for Time Series Forecasting

1. Feature Engineering
    a. Avoiding data leakage
    b. Setting a forecast horizon

2. Time Delay Embedding
    a. Lags or backshift
    b. Rolling window aggregations
    c. Seasonal rolling window aggregations
    d. Exponentially weighted moving averages (EWMA)
3. Temporal embedding
    a. Calendar features
    b. Time elapsed
    c. Fourier terms

Target Transformations for Time Series Forecasting

1. Handling non-stationary in time series
2. Detecting and correcting for unit roots
    a. Unit roots
    b. The Augmented Dickey-Fuller (ADF) test
    c. Differencing transform
3. Detecting and correcting for trends
    a. Deterministic and stochastic trends
    b. Kendall's Tau
    c. Mann-Kendall test (M-K test)
    d. Detrending transform
4. Detecting and correcting for seasonality
    a. Detecting seasonality
    b. Deseasonalizing transform
5. Detecting and correcting for heteroscedasticity
    a. Detecting heteroscedasticity
    b. Log transform
    c. Box-Cox transform

Forecasting Time Series with Machine Learning Models

1. Training and predicting with machine learning models
2. Generating single-step forecast baselines
3. Standardized code to train and evaluate machine learning models
    a. FeatureConfig
    b. MissingValueConfir
    c. ModelConfig
    d. MLForecast
    e. Helper functions for evaluating models
    f. Linear regression
    g. Regularized linear regression
    h. Decision trees
    i. Random forest
    j. Gradient boosting decision trees