Time Series Modeling

- Hunter

Quiz

- What time series models do you know?
- What time series evaluation metrics do you know?

Topics

- Time Series Forecasting
- Time Series Evaluation

1.

Time Series Forecasting

ARIMA

ARIMA stands for AutoRegressive Integrated Moving Average

- AR (Autoregression): This emphasizes the dependent relationship between an observation and its preceding or ‘lagged’ observations.
- I (Integrated): To achieve a stationary time series, one that doesn’t exhibit trend or seasonality, differencing is applied. It typically involves subtracting an observation from its preceding observation.
- MA (Moving Average): This component zeroes in on the relationship between an observation and the residual error from a moving average model based on lagged observations.
- The parameters of the ARIMA model are defined as follows:
  - p: The lag order, representing the number of lag observations incorporated in the model.
  - d: Degree of differencing, denoting the number of times raw observations undergo differencing.
  - q: Order of moving average, indicating the size of the moving average window.

SARIMA

Seasonal Autoregressive Integrated Moving Average, SARIMA or Seasonal ARIMA, is an extension of ARIMA that explicitly supports univariate time series data with a seasonal component.

- There are four seasonal elements that are not part of ARIMA that must be configured; they are:
  - P: Seasonal autoregressive order.
  - D: Seasonal difference order.
  - Q: Seasonal moving average order.
  - m: The number of time steps for a single seasonal period.

Exponential Smoothing

Exponential smoothing is a time series forecasting method for univariate data that can be extended to support data with a systematic trend or seasonal component.

- Hyperparameters:
  - Alpha: Smoothing factor for the level.
  - Beta: Smoothing factor for the trend.
  - Gamma: Smoothing factor for the seasonality.
  - Trend Type: Additive or multiplicative.
  - Dampen Type: Additive or multiplicative.
  - Phi: Damping coefficient.
  - Seasonality Type: Additive or multiplicative.
  - Period: Time steps in seasonal period.

Tree-based Models

For boosting models, integrating a linear component into the tree-building process can allow the model to capture both linear and non-linear patterns, offering a way to include trend extrapolation capabilities.

- ✗ Convert time features
- ✗ Convert previous observations to features.

2.

Time Series Evaluations

MSE, RMSE, MAE

- MSE

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

- RMSE

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

- MAE

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

MAPE & sMAPE

- MAPE

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

  - MAPE is asymmetric and reports higher errors if the forecast value is higher than the actual value and reports lower errors when the forecast value is less than the actual value.

Month Year	Actual Spend	Forecasted Spend	Absolute Percentage Error
Jan-22	500	600	20.00
Feb-22	600	500	16.67

- sMAPE (Symmetric Mean Absolute Percentage Error)

$$sMAPE = \frac{100\%}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)}$$

Month Year	Actual Spend	Forecasted Spend	Absolute Percentage Error	Symmetric Absolute Percentage Error
Jan-22	100	90	10.00	10.53
Feb-22	90	100	11.11	10.53
Mar-22	150	100	33.33	40.00
Apr-22	100	150	50.00	40.00
May-22	80	100	25.00	22.22
Total	**520**	**540**	**3.85**	3.77

MASE

- MASE (mean absolute scaled error)

$$\mathrm{MASE} = \mathrm{mean}\left(\frac{|e_j|}{\frac{1}{T-1}\sum_{t=2}^{T}|Y_t - Y_{t-1}|}\right) = \frac{\frac{1}{J}\sum_j |e_j|}{\frac{1}{T-1}\sum_{t=2}^{T}|Y_t - Y_{t-1}|}$$

Recommended Readings

https://medium.com/@wainaina.pierre/the-complete-guide-to-time-series-forecasting-models-ef9c8cd40037

https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/

https://medium.com/@mubarakdaha/understanding-time-series-forecasting-with-arima-59cd7140d6c3

https://medium.com/@tirthamutha/time-series-forecasting-using-sarima-in-python-8b75cd3366f2

https://towardsdatascience.com/time-series-in-python-part-2-dealing-with-seasonal-data-397a65b74051

https://towardsdatascience.com/time-series-in-python-exponential-smoothing-and-arima-processes-2c67f2a52788

https://medium.com/@simon.peter.mueller/overcoming-the-limitations-of-tree-based-models-in-time-series-forecasting-c2c5bd71a8f1

https://medium.com/@jonatasv/metrics-evaluation-mse-rmse-mae-and-mape-317cab85a26b

https://medium.com/@vinitkothari.24/time-series-evaluation-metrics-mape-vs-wmape-vs-smape-which-one-to-use-why-and-when-part1-32d3852b4779

https://mlpills.dev/cheatsheets/evaluation-metrics-for-time-series-forecasting/#aioseo-mean-absolute-scaled-error-mase