International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

Long-Term vs. Short-Term Forecasting Data Scientist's Strategies for Different Horizons

Sowmya Ramesh Kumar

Seattle, WA
Email: *rsowmyash[at]gmail.com*

Abstract: ***Forecasting is a critical aspect of data science, and the choice between long - term and short - term forecasting strategies depends on the nature of the data, the objectives of the analysis, and the specific needs of the application. In this discussion, we will explore the distinctions between long - term and short - term forecasting, the strategies employed by expert data scientists for each horizon, and the considerations that influence their decision - making.***

Keywords: Long Term Forecasting, Short term forecasting, seasonality, decomposition, forecasting, modeling

1. Introduction

In forecasting space of data science, there are two terminologies: [a] grain [b] horizon. [a]. grain: what level of granularity that forecasts are expected. For e. g. Grain could be at Yearly, Monthly, weekly and daily for time. There can be different hierarchies for Product and/or Location. This is determined based on consumption and application of these forecasts. [b]. horizon. How far out the forecasts are expected it to be. There are two types of forecasting w. r. t horizon.
[I] Long - Term Forecasting:
[II] Short - Term Forecasting

[I] Long - term forecasting [2]:

It involves predicting future values of a variable over an extended period, often spanning several years or even decades.
There are different data science strategies for accurate development of forecasts.

a) Trend Analysis [3]: Expert data scientists leverage trend analysis methods to identify and model long - term patterns in the data. This may include linear or nonlinear trends that provide insights into the overall direction of the variable over an extended horizon.
b) Seasonal Decomposition: Decomposing the time series into its components, such as trend, seasonality, and residual, helps in understanding the long - term behavior of the series. This allows for more accurate forecasting by capturing underlying patterns.

$$yt=Tt+St+Et$$

where y_t is the observed value at time t, T_t is the trend component at time t, S_t is the seasonal component at time t, and E_t is the error component at time t 4

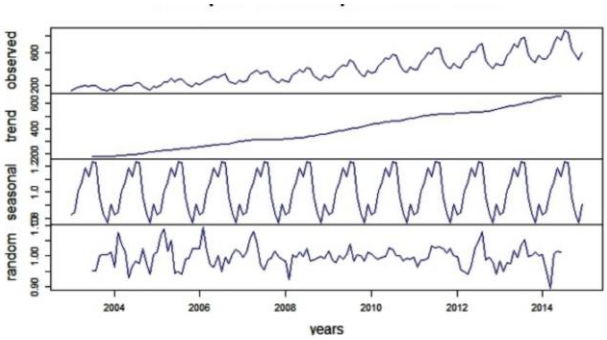

Figure 1: Decomposed data with seasonality, trend and noise.

c) Econometric Models [1]: For economic and financial data, data scientists often turn to econometric models that incorporate various factors influencing the long - term trends. These models may include variables such as interest rates, inflation, and macroeconomic indicators. These models are as accurate as inputs into these models. Success usage of these models help in prediction of long-term forecasts with current and predicted market trends. Potential challenges of this methodology to consider is productionalization of these new data sources (interest rates, inflation etc.) once the models are tested to be producing high quality forecasts.
d) Machine Learning Models: Complex machine learning models, such as deep learning architectures, can capture intricate patterns in long - term data. These models excel at learning from historical data to make predictions for extended time horizons.

[II] Short - Term Forecasting:

Short - term forecasting focuses on predicting near - future values of a variable, typically within days, weeks, or a few months.

Strategies:

a) Time Series Decomposition [5]: Like long - term forecasting, short - term strategies involve decomposing the time series to capture trends and seasonality. However, the emphasis is on capturing shorter - term fluctuations.
b) ARIMA Models [5]: Autoregressive Integrated Moving Average (ARIMA) models are effective for short - term forecasting. They consider the autocorrelation and moving

averages of the time series, making them suitable for capturing short - term dependencies.

This approach involves considering value Y at time point t and adding/subtracting based on the Y values at previous time points (e. g., t - 1, t - 2, etc.), and adding/subtracting error terms from previous time points.
The formula itself looks like this:
$Yt=c+\phi 1ydt-1+\phi pydt-p+...+\theta 1et-1+\theta qet-q+et$
Where “e” is an error term and “c” is a constant.

c) Exponential Smoothing [5]: Exponential smoothing methods, such as Holt - Winters, are popular for short - term forecasting. These methods assign exponentially decreasing weights to past observations, giving more importance to recent data.
d) Machine Learning Ensembles [4]: Ensemble models, such as Random Forests and Gradient Boosting, are powerful for short - term predictions. They can capture complex relationships and dependencies within the data over shorter time intervals.

There are several considerations a data scientist needs to keep in mind for choosing these strategies:

- *Data Characteristics*: The nature of the time series data, including its volatility, seasonality, and trend, influences the choice between long - term and short - term strategies. In many business cases, there are ‘n’ number of time - series to forecast, looking at patterns of each group of time - series is important. Sometimes, if time - series are different from each other, then we have a need to apply clustering to different time - series. And apply different model methods to different group of time - series.
- *Application Requirements:* Understanding the specific needs of the application is crucial. Some scenarios demand accurate short - term predictions for immediate decision - making, while others require a broader outlook provided by long - term forecasting.
- *Computational Resources*: Long - term forecasting often involves more complex models that may require significant computational resources. Short - term strategies might be more computationally efficient, allowing for real - time or near - real - time predictions. As we apply sophisticated models like deep learning, there are needs of GPU - enabled capabilities to run models. This is another important consideration that involves team alignment and budget.

2. Conclusion

In conclusion, the choice between long - term and short - term forecasting strategies depends on a nuanced understanding of the data and the objectives of the analysis. Expert data scientists carefully consider the characteristics of the time series, the application requirements, and the available computational resources when selecting the most appropriate forecasting approach. Both long - term and short - term strategies play vital roles in providing valuable insights for decision - makers, contributing to the efficacy of data - driven decision - making processes.

References

[1] AA Grasa – 2013, Econometric model selection: A new approach
[2] JS Armstrong, Long - Range Forecasting, 2nd - Available at SSRN 666990, 2010 - papers. ssrn. com
[3] Clive W. J. Granger, Yongil Jeon, Long - term forecasting and evaluation, International Journal of Forecasting, Volume 23, Issue 4, 2007, Pages 539 - 551, ISSN 0169 - 2070, https: //doi. org/10.1016/j. ijforecast.2007.07.002.
[4] H. Sangrody, N. Zhou, S. Tutun, B. Khorramdel, M. Motalleb and M. Sarailoo, "Long term forecasting using machine learning methods, " *2018 IEEE Power and Energy Conference at Illinois (PECI)*, Champaign, IL, USA, 2018, pp.1 - 5, doi: 10.1109/PECI.2018.8334980.
[5] Iram Naim, Tripti Mahara, Ashraf Rahman Idrisi, Effective Short - Term Forecasting for Daily Time Series with Complex Seasonal Patterns, Procedia Computer Science, Volume 132, 2018, Pages 1832 - 1841, ISSN 1877 - 0509, https: //doi. org/10.1016/j. procs.2018.05.136.