☰   🔍   (http://bit.ly/8ipv)   SUBSCRIBE

PUBLISHED ON NOVEMBER 1, 2021 IN DEVELOPERS CORNER (HTTPS://ANALYTICSINDIAMAG.COM/CATEGORY/DEVELOPERS_CORNER/)

A Guide to Different Evaluation Metrics for Time Series Forecasting Models

Measuring the performance of any machine learning model is very important, not only from the technical point of view but also from the business perspective.

BY VIJAYSINH LENDAVE(HTTPS://ANALYTICSINDIAMAG.COM/AUTHOR/VIJAYSINH-LENDAVEANALYTICSINDIAMAG-COM/)

Measuring the performance of any machine learning (https://analyticsindiamag.com/how-machine-learning-can-be-used-with-blockchain-technology/) model is very important, not only from the technical point of view but also from the business perspective. Especially when the business decisions are dependent on the insights generated from the forecasting models, knowing its accuracy becomes vital. There are different types of evaluation metrics used in machine learning depending on the model used and the results generated. In the same context, there are different evaluation metrics used to measure the performance of a time-series forecasting model. In this post, we will discuss different evaluation metrics used for measuring the performance of a time series model with their importance and applicability. The major points to be covered in this article are listed below.

Table of Contents

1. Measuring Time Series Forecasting Performance
2. Evaluation Metrics to Measure Performance

    1. R-Squared

    2. Mean Absolute Error

    3. Mean Absolute Percentage Error

    4. Mean Squared Error

    5. Root Mean Squared Error

    6. Normalized Root Mean Squared Error

    7. Weighted Absolute Percentage Error

    8. Weighted Mean Absolute Percentage Error

3. Summary

Let's start the discussion by understanding why measuring the performance of a time series forecasting model is necessary.
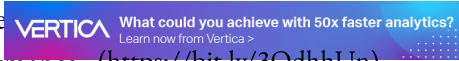
Measuring Time Series Forecasting Performance

The fact that the future is wholly unknown and can only be predicted from what has already occurred is a significant distinction in forecasting. The ability of a time series forecasting model to predict the future is defined by its performance. This is frequently

Time series prediction performance measurements provide a summary of the forecast model's skill and capability in making the forecasts. There are numerous performance metrics from which to pick. Knowing which metric to use and how to interpret the data might be difficult.

Moving further, we will see different performance measures that can be applied to evaluate the forecasting model under different circumstances.

Evaluation Metrics to Measure Performance

Now, let us have a look at the popular evaluation metrics used to measure the performance of a time-series forecasting model.

R-Squared

The stationary R-squared is used in time series forecasting as a measure that compares the stationary part of the model to a simple mean model. It is defined as,

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$$

Where SSres denotes the sum of squared residuals from expected values and SStot denotes the sum of squared deviations from the dependent variable's sample mean. It denotes the proportion of the dependent variable's variance that may be explained by the
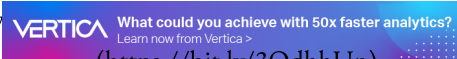
independe... ...variance. A high R^2
model's va... ...true values... ...a low
R^2 value suggests that the two values are not strongly related.

The most important thing to remember about R-squared is that it does not indicate whether or not the model is capable of making accurate future predictions. It shows whether or not the model is a good fit for the observed values, as well as how good of a fit it is. A high R^2 indicates that the observed and anticipated values have a strong association.

Mean Absolute Error (MAE)

The MAE is defined as the average of the absolute difference between forecasted and true values. Where y_i is the expected value and x_i is the actual value (shown below formula). The letter n represents the total number of values in the test set.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

The MAE shows us how much inaccuracy we should expect from the forecast on average. MAE = 0 means that the anticipated values are correct, and the error statistics are in the original units of the forecasted values.

The lower the MAE value, the better the model; a value of zero indicates that the forecast is error-free. In other words, the model with the lowest MAE is deemed superior when comparing many models.

errors. It can be combined with other measures to see if the errors
are higher (see Root Mean Square Error below). Furthermore, MAE
might obscure issues related to low data volume; for more
information, check the last two metrics in this article.

Mean Absolute Percentage Error (MAPE)

MAPE is the proportion of the average absolute difference between
projected and true values divided by the true value. The anticipated
value is Ft, and the true value is At. The number n refers to the total
number of values in the test set.

$$\text{MAPE} = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

It works better with data that is free of zeros and extreme values
because of the in-denominator. The MAPE value also takes an
extreme value if this value is exceedingly tiny or huge.

The model is better if the MAPE is low. Remember that MAPE
works best with data that is devoid of zeros and extreme values.
MAPE, like MAE, understates the impact of big but rare errors
caused by extreme values.

Mean Square Error can be utilized to address this issue. This statistic
may obscure issues related to low data volume; for more
information, check the last two metrics in this article.

Mean Squared Error (MSE)
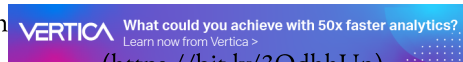
MSE is defined as average of the error squares. It is known as the metric that evaluates the quality of a forecasting model or predictor. MSE also takes into account variance (the difference between anticipated values) and bias (the distance of predicted value from its true value).

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

Where y' denotes the predicted value and y denotes the actual value. The number n refers to the total number of values in the test set. MSE is almost always positive, and lower values are preferable. This measure penalizes large errors or outliers more than minor errors due to the square term (as seen in the formula above).

The closer MSE is to zero, the better. While it overcomes MAE and MAPE extreme value and zero problems, it may be harmful in some instances. When dealing with low data volume, this statistic may ignore issues; to address this, see Weighted Absolute Percentage Error and Weighted Mean Absolute Percentage Error.

Root Mean Squared Error(RMSE)

This measure is defined as the square root of mean square error and is an extension of MSE. Where y' denotes the predicted value and y denotes the actual value. The number n refers to the total number of values in the test set. This statistic, like MSE, penalizes greater errors more.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}$$

This statist... always positive, wi... higher performance. Anyhow the RMSE number is in the same symbol as the projected value, which is an advantage of this technique. In comparison to MSE, this makes it easier to comprehend.

The RMSE can also be compared to the MAE to see whether there are any substantial but uncommon inaccuracies in the forecast. The wider the gap between RMSE and MAE, the more erratic the error size. This statistic can mask issues with low data volume.

Normalized Root Mean Squared Error (NRMSE)

The normalized RMSE is used to calculate NRMSE, which is an extension of RMSE. The mean or the range of actual values are the two most used methods for standardizing RMSE (difference of minimum and maximum values). The maximum true value is y_{max}, while the smallest true value is y_{min}.

$$\text{NRMSD} = \frac{\text{RMSD}}{y_{\max} - y_{\min}} \text{ or } \text{NRMSD} = \frac{\text{RMSD}}{\bar{y}}$$

NRMSE is frequently used to compare datasets or forecasting models with varying sizes (units and gross revenue, for example). The smaller the value, the better the model's performance. When working with little amounts of data, this metric can be misleading. However, Weighted Absolute Percentage Error and Weighted Mean Absolute Percentage Error can help.

Weighted Mean Absolute Percentage Error (WMAPE)

WMAPE (sometimes called wMAPE) is an abbreviation for Weighted Mean Absolute Percentage Error. It is a measure of a forecasting method's prediction accuracy. It is a MAPE version in

which error is created by real values (e.
forecasting errors are weighted by sales volume).

$$\text{WMAPE} = \frac{\sum_{t=1}^{n} |A_t - F_t|}{\sum_{t=1}^{n} |A_t|}$$

where A is the current data vector and F is the forecast This metric has an advantage over MAPE in that it avoids the 'infinite error' problem.

The higher the model's performance, the lower the WMAPE number. When evaluating forecasting models, this metric is useful for low volume data where each observation has a varied priority. The weight value of observations with a higher priority is higher. The WMAPE number increases as the error in high-priority forecast values grows.

Summary

Let's have a quick summary of all the above-mentioned measures.

- When the relation between the forecasted and actual value is to be known then R^2 is used.
- When absolute error must be measured, MAE is useful. It is simple to understand, but in the case of data with extreme values, it is inefficient. MAPE is also simple to understand and is used to compare different forecast models or datasets because it is a percentage value. MAPE has the same problem as MAE in that it is inefficient when data contains extreme values.
- MSE is beneficial when the spread of prediction values is significant and larger values must be punished. However,
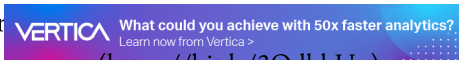
- When the spread is important and bigger values need to be penalized, RMSE (NRMSE) is also useful. When compared to MSE, RMSE is easier to interpret because the RMSE number is on the same scale as the projected values.
- When dealing with low-volume data, WMAPE is also useful. WMAPE uses the weight (priority value) of each observation to help incorporate the priority.

Conclusion

Through this post, we have seen different performance evaluation metrics used in time series forecasting in different scenarios. Most of all above-mentioned measures can directly be utilized from sklearn.metrics class or can be directly implemented from scratch with NumPy and math modules.

References

- A Comparative Study of Performance Estimation Methods for Time Series Forecasting (https://ieeexplore.ieee.org/abstract/document/8259815)
- How to Evaluate Your Forecast (https://www.jedox.com/en/blog/error-metrics-how-to-evaluate-forecasts/)

More Great AIM Stories

Using GANs For Generating High-Quality Faces From Degraded Images (https://analyticsindiamag.com/using-gans-for-generating-high-quality-

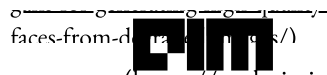How Uber is Leveraging Apache Kafka For More Than 300 Micro Services (https://analyticsindiamag.com/how-

Guide to Different Padding Methods
for CNN Models
(https://analyticsindiamag.com/guide-
to-different-padding-methods-for-
cnn-models/)
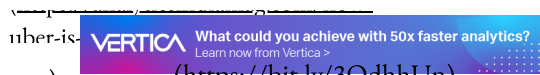
Tour to All The Useful Processing
Steps on Image Data
(https://analyticsindiamag.com/tour-
to-all-the-useful-processing-steps-on-
image-data/)

What's New With IBM's Mayflower
Autonomous Ship
(https://analyticsindiamag.com/whats
new-with-ibms-mayflower-
autonomous-ship/)

Inside Neurograins, A New
Technology To Record Brain
Impulses
(https://analyticsindiamag.com/inside
neurograins-a-new-technology-to-
record-brain-impulses/)

(https://analyticsindiamag.com/author/vijaysinh-
lendave/analyticsindiamag.com/)

Vijaysinh is an enthusiast in machine learning and deep learning. He is skilled in ML
algorithms, data manipulation, handling and visualization, model building.

3 Ways to Join our Community

Discord
Server

Telegram Channel

Discover special offers, top stories,
upcoming events, and more.

Subscribe to
our newsletter

MOST POPULAR

3D Reconstruction Models Make Metaverse More Possible (https://analyticsindiamag.com/3d-reconstruction-models-make-metaverse-more-possible/)

Mohit Pandey

RawNeRF's noise reduction method when combined with scene gives a high-resolution output which is seamless while transitioning between angle and positions.

Contrary to Popular Belief, Hackathons Do Get Deployed for Real-World Problems (https://analyticsindiamag.com/contrary-to-popular-belief-hackathons-do-get-deployed-for-real-world-problems/)

Tasmia Ansari

Tech giants such as Microsoft and Facebook have rolled out features generated from hackathons.

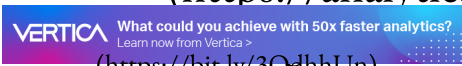### Building a Career in Web3 as a Developer (https://analyticsindiamag.com/building-a-career-in-web3-as-a-developer/)

Pritam Bordoloi

Text-to-Image is All the Rage. Why Aren't We Talking About Imagen? (https://analyticsindiamag.com/...)

e-rage-so-wh
aren't-we-talking-about-imag

Shraddha Goled

OpenAI has also allowed full usage rights to commercial
that they create with DALL.E – including the right to re
merchandise.

My Journey to Data Science: Mishra of VMware (https://analyticsindiamag.co journey-to-data-science-pragy mishra-of-vmware/)

Bhuvana Kamath

Pragya Mishra, business analyst (data sciences and analyt
speaks about the ideal inclusive workplace and the challe
business analyst in the data analytics space

Council Post: ML observabil monitoring: The tactical/stra paradox (https://analyticsindiamag.co post-ml-observability-vs-ml-n the-tactical-strategic-paradox

Vinay Kumar Sankarapu

A good ML Observability tool can provide a common fr
stakeholders to understand, debug, monitor and deliver
framework for AI Governance.

5G is Hackable, But it Can b Saved (https://analyticsindiamag.co is-hackable-but-it-can-be-save

Zinnia Banerjee

Security Research Labs uncovered new hacking frontiers
up despite improvements in 5G standards. The team wa:
into the network multiple times, thereby getting hold of

data or disrupting operations due to poorly configured c
technology

Behind Indian Government AI & Robotics Innovation F: (https://analyticsindiamag.co indian-government-supporte robotics-innovation-firm/)

Pritam Bordoloi

If we can get self-driving cars to work on Silk Board in Bc
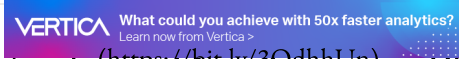in the US, it will work in Europe, it will work in Elon M

PIM

Google Just Stepped Up the Text-to-Image AI (https://analyticsindiamag.co just-stepped-up-the-game-for image-ai/)

Mohit Pandey

DreamBooth can understand the subject of the given im the existing context in the image, and then synthesise it i context with high fidelity.

## Data Centres, Please Stop Dr My Water (https://analyticsindiamag.co centres-please-stop-drinking-water/)

Lokesh Choudhary

Data centres globally are using too much water to cool d storing data is important, we need to have a robust data address the issue
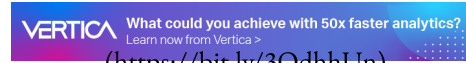
Our Mission Is To Bring About Better-Informed And More Conscious Decisions About Technology Through Authoritative, Influential, And Trustworthy Journalism.

SHAPE THE FUTURE OF TECH

CONTACT US ⟶
(HTTPS://ANALYTICSINDIAMAG.COM/CONTACT-US/)

(https://analyticsindiamag.com)

(https://www.linkedin.com/company/analytics-india-magazine)
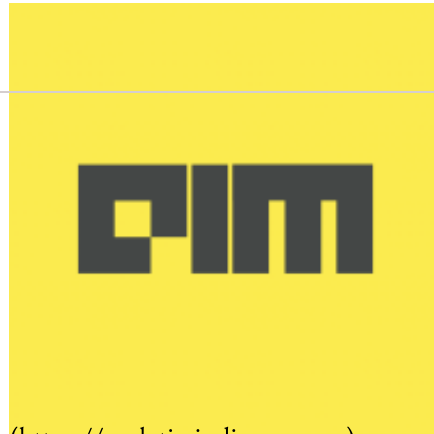
About Us

Advertise

Weekly Newsletter

Write for us

Careers

Contact Us

RANKINGS & LISTS

Academic Rankings

Best Firms To Work For

Top Analytics Providers

Top AI Startups

OUR CONFERENCES

Cypher

The MachineCon

Machine Learning Developers Summit

The Rising

Data Engineering Summit
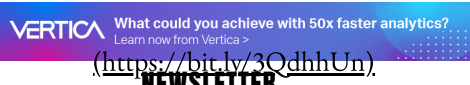
OUR BRANDS

MachineHack

AIM Recruits

AIM Leaders Council

Best Firm Certification

AIM Research

VIDEOS

Documentary – The Transition Cost

Web Series – The Dating Scientists

Podcasts – Simulated Reality

Analytics India Guru

The Pretentious Geek

Deeper Insights with Leaders

Curiosum – AI Storytelling

GET INVOLVED

AIM Campus Ambassador

AIM Mentoring Circle

Python Libraries Repository

PIM

(https://analyticsindiamag.com)

(https://bit.ly/3QdhhUn)

AWARDS

NEWSLETTER

Stay up to date with our latest news, receive exclusive deals, and more.

Data Science Excellence Awards

AI100

40 under 40 Data Scientists

Women in AI Leadership

Enter Your Email Address

EVENTS

SUBSCRIBE ⟶

Our Events

AIM Custom Events

AIM Virtual

FOR ML DEVELOPERS

Hackathons

Discussion Forum

Job Portal

Mock Assessments

Practice ML

Free AI Courses

Terms of use   Privacy Policy Copyright

© Analytics India Magazine Pvt Ltd 2022

(https://analyticsindiamag.com/terms-https://analyticsindiamag.com/privacy-(https://analyticsindiamag.com/

use/)   policy/)   trademarks/)