Metrics to Calculate Performance of Machine Learning Model

After implementing a machine learning algorithm, the next step we move towards is to find how effective our model is based on some metrics. This is the most essential part of any project as different performance metrics are used to evaluate different Machine Learning algorithms.

When we train our model, the model generalizes on unseen data and then we need to know whether it actually works. Thus we use some evaluation techniques.

These performance metrics are categorized based on the type of Machine Learning problem. It means we have different evaluation techniques for respective Regression problems.

Regression Problem

- Mean Absolute Error
- Mean Squared Error
- Root Mean Squared Error
- R squared
- Adjusted R squared

Mean Absolute Error (MAE)

Measures average/mean squared error of our predictions.

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

Gives less weight to the outliers, when you are sure that they are outliers prefer MAE to MSE.

Mean Absolute Percentage Error (MAPE)

Also known as mean absolute percentage deviation (MAPD), it measures the size of the error in percentage tems. It is calculated by taking the absolute deviation and dividing it by the data to get the error percentage.

$$\mathrm{M} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|,$$

where At is the actual value and Ft is the forecast value. The MAPE is also sometimes reported as a percentage, which is the above equation multiplied by 100. The difference between At and Ft is divided by the actual value At again. The absolute value in this calculation is summed for every forecasted point in time and divided by the number of fitted points n. Multiplying by 100% makes it a percentage error.

Lower numbers are better than large ones.

The main issues with MAPE are that near zero and zero predicted figures will yield infinite or non valid results due to the y_hat denominator.

Mean Squared Error (MSE)

Incorporates both the variance and the bias of the predictor.

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

When you have unexpected values that you should take into account use MSE instead of MAE.

Root Mean Square Error (RMSE)

RMSE is an absolute measure of fit. Lower values of RMSE are indicative of a better fit

RMSE can be interpreted as the standard deviation of the unexplained variance.

RMSE is in the same units as the predicted variable.

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

By squaring the errors we can get more accurate results as the negative and positive errors don't cancel each other.

R Squared

The r2_score or commonly known as the R^2 (R-squared) is the proportion of the variance in the dependent variable that is predictable from the independent variable(s). It is known as the coefficient of determination. It is a statistical measure of how close the data are to the fitted regression line Or indicates the goodness of fit of a set of predictions to the actual values. The value of R^2 lies between 0 and 1 where 0 means no-fit and 1 means perfectly-fit.

The formula to find R^2 is as follows:

$$\mathbf{R^2 = 1 - SSE/SST}$$

Where SSE is the Sum of Square of Residuals. Here residual is the difference between the predicted value and the actual value.

And SST is the Total Sum of Squares.

Adjusted R squared

R-squared explains the degree to which our input variables explain the variation of our output/predicted variable. So, the higher the R squared, the more variation is explained by our input variables and hence better is our model. The Adjusted R-squared value is similar to the R-squared value, but it accounts for the number of variables that is, R-squared will either stay the same or increase with the addition of more variables, even if they do not have any relationship with the output variables. This is where "Adjusted R square" comes to help. Adjusted R-square penalizes for adding variables that are not useful for predicting the target.

Hence, if you are building Linear regression on multiple variables, it is always suggested that we use Adjusted R-squared to judge the goodness of the model. In case you only have one input variable, R-squared and Adjusted R squared would be exactly the same.

Note:

If the R^2 increases by a significant value, then the adjusted r-squared would increase.

If there is no significant change in R^2, then the adjusted r2 would decrease.