Overview

Model evaluation is very important in data science. It helps you to understand the performance of your model and makes it easy to present your model to other people.

There are many different evaluation metrics out there but only some of them are suitable to be used for regression.

Well, unlike classification, accuracy in regression model is slightly harder to illustrate. It is merely impossible for you to predict the exact value but rather how close your prediction is against the real value.

Types of Error metrics

There are 3 main metrics for model evaluation in regression:

R Square/Adjusted R Square

Mean Square Error(MSE)/Root Mean Square Error(RMSE)

Mean Absolute Error(MAE)

R Square/Adjusted R Square

R Square measures how much of variability in dependent variable can be explained by the model. It is square of Correlation Coefficient(R) and that is why it is called R Square.

$$R^2 = 1 - \frac{SS_{Regression}}{SS_{Total}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

R Square/Adjusted R Square

R Square is calculated by the sum of squared of prediction error divided by the total sum of square which replace the calculated prediction with mean.

R Square value is between 0 to 1 and bigger value indicates a better fit between prediction and actual value.

R Square/Adjusted R Square

R Square is a good measure to determine how well the model fits the dependent variables.

- **Then why do we need Adjusted R square?**

  However, it does not take into consideration of over-fitting problem.

If your regression model has many independent variables, because the model is too complicated, it may fit very well to the training data but performs badly for testing data. That is why Adjusted R Square is introduced because **it will penalize additional independent variables added to the model and adjust the metric to prevent over-fitting issue.**

In Python, you can calculate R Square using,

- **sklearn.metrics.r2_score()**

Mean Square Error(MSE)

While R Square is a relative measure of how well the model fits dependent variables, ***Mean Square Error is an absolute measure of the goodness for the fit***.

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

Mean Square Error(MSE)

MSE is calculated by the sum of square of prediction error which is real output minus predicted output and then divide by the number of data points.

It gives you an absolute number on how much your predicted results deviate from the actual number.

You cannot interpret much insights from one single result but it gives you an real number to compare against other model results and help you select the best regression model.

Root Mean Square Error(RMSE)

Why RMSE ??

- Root Mean Square Error(RMSE) is the square root of MSE.
- It is used more commonly than MSE because ***firstly sometimes MSE value can be too big to compare easily***.
- Secondly, MSE is calculated by the square of error, and thus square root brings it back to the same level of prediction error and ***make it easier for interpretation.***

Mean Absolute Error(MAE)

Mean Absolute Error(MAE) is similar to Mean Square Error(MSE). However, instead of the sum of square of error in MSE, MAE is taking the sum of absolute value of error.

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

Mean Absolute Error(MAE)

Compare to MSE or RMSE, MAE is a ***more direct representation of sum of error terms.***

MSE gives larger penalization to big prediction error by square it **while MAE treats all errors the same.**

Conclusion

Square/Adjusted R Square are better used to explain the model to other people ***because you can explain the number as a percentage of the output variability***.

MSE, RMSE or MAE are better to be used ***to compare performance between different regression models***.

However, it makes total sense to use ***MSE if value is not too big*** and **MAE** if you do not want ***to penalize large prediction error.***

Adjusted R square is the only metric here that considers ***overfitting problem***.