Bias & variance:-

In Machine Learning, a model's performance is based on its predictions and how well it generalizes towards unseen, independent data. One way to measure a model's accuracy is by keeping account of the bias and variance in the model. In this article, we will learn how bias-variance plays an important role in determining the authenticity of the model.

Errors:-

Any model in Machine Learning is assessed based on the prediction error on a new independent, unseen data set. Error is nothing but the difference between the actual output and the predicted output. To calculate the error, we do the summation of reducible and irreducible error a.k.a bias-variance decomposition.

There are 2 types of errors in ML:-
1)Irreduceble Error
2)Reduceble Error

1)Irreduceble Error:-(kadime maadabahudadanta error)

Irreversible error is nothing but those errors that cannot be reduced irrespective of any algorithm that you use in the model. It is caused by unusual variables that have a direct influence on the output variable.(idu direct aagi o/p alle depend aagirodarinda naavu idanna calculate maadoke aagalla,yaake andre adu o/p aadagle maatra gothagodu)

2)Reduceble Error:-(kadime maadoke aagadantaha error):-

So in order to make your model efficient, we are left with the reducible error(idonde maarga namage model na efficient maadoke option irodu) that we need to optimize at all costs.
It has 2 types:-
A reducible error has two components –
1)Bias
2)Variance
presence of bias and variance influence the model's accuracy in several ways like overfitting, underfitting, etc.

1)Bias:-(illi kadime dataset kagondre jaasti bais iratte jaasti dataset tagondre kadime bias iratte)

Bias is basically how far we have predicted the value from the actual value. We say the bias is too high if the average predictions are far off from the actual values.

A high bias will cause the algorithm to miss a dominant pattern or relationship between the input and output variables. When the bias is too high, it is assumed that the model is quite simple and does not form the complexity of the data set to determine the relationship and thus, causing underfitting.

2)Variance:-(idu suppose naavu tagonda dataset nalli unwanted darasets idre eege aagatte)

When a model does not perform ,as well as,but it does with the trained data set, there is a possibility that the model has a variance. It basically tells how

scattered the predicted values are from the actual values.

A high variance in a data set means that the model has trained with a lot of noise and irrelevant data. Thus causing overfitting in the model. When a model has high variance, it becomes very flexible and makes wrong predictions for new data points. Because it has tuned itself to the data points of the training set.

How Does It Affect The Machine Learning Model?

We can put the relationship between bias-variance in four categories listed below:

1)High Variance-High Bias – The model is inconsistent and also inaccurate on average.
2)Low Variance-High Bias – Models are consistent but low on average.(Overfitting)
3)High Variance-Low Bias – Somewhat accurate but inconsistent on averages(Underfitting)
4)Low Variance-Low Bias – It is the ideal scenario, the model is consistent and accurate on average.(Not Posible)

Bias-Variance Trade-Off:-
Finding the right balance between the bias and variance of the model is called the Bias-Variance trade-off. It is basically a way to make sure the model is neither overfitted or underfitted in any case.

*If the model is too simple and has very few parameters, it will suffer from high bias and low variance.
*model has a large number of parameters, it will have high variance and low bias.
* This trade-off should result in a perfectly balanced relationship between the two. Ideally, low bias and low variance is the target for any Machine Learning model.

Total Error:-

In any Machine Learning model, a good balance between the bias and variance serves as a perfect scenario in terms of predictive accuracy and avoiding overfitting, underfitting altogether. An optimal balance between the bias and variance, in terms of algorithm complexity, will ensure that the model is never overfitted or underfitted at all.
*The mean squared error in a statistical model is considered as the sum of squared bias and variance and variance of error. All this can be put inside a total error where we have bias, variance and irreducible error in a model.

$$err(x) = Bias2 + Variance + irreducible\ error$$