MSE, MAE, RMSE:

MSE: Mean Squared Error

MAE: Mean Absolute Error

RMSE: Root Mean Squared Error

We have already seen R squared and adjusted R squared. They are used to check how to model is performing.

But if we really need to focus on the error with respect to each and every data point, we can use these three.

Suppose if we have a dataset with two features: Experience and Salary

Our main aim is to reduce the error and for that we use different cost functions.

First cost function is Mean Squared Error.

MSE =

are the predicted data points

n is the number of data points we have in the dataset

Our main aim is to reduce this cost function

We can also use this as a parameter to see how our model is performing

Advantages:

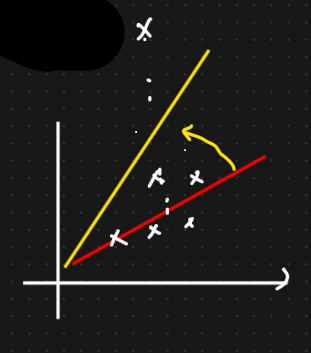
1. is a quadratic equation, that gives a bell curve or gradient descent when plotted, hence we will be able to reach a global minima. Whenever we have such a gradient descent, it is differentiable at all points. Hence at every point we will be able to calculate a slope.

2. It has one local or global minima. In other types of curves, we may have one global minima and multiple local minima. If we have some more local minimas, then the convergence will get stuck over there because there the slope will be 0 as well.

3. It converges faster.

Disadvantages:

1. Not robust to outliers. If we have an outlier, the distance would be large and it would get even more amplified when squaring. Thus this process penalizes outliers. Error will increase in the presence of outliers.



Suppose, for the initial data points, the red line represents the best fit line.

Later when we added an outlier, the MSE increases, to compensate that, the best fit line moves towards it to optimize the cost function, which is incorrect.

Thus because of an outlier, the best fit line moves by a huge difference.

2. When we do , the units get changed. Suppose if we have salary in thousands, then the new unit would be . With respect to MSE, whenever we try to compute , the unit changes compared to the output variable. So it is no longer in the same unit. It makes it difficult to understand the error when comparing it with the truth value.

Mean Absolute Error:

MAE =

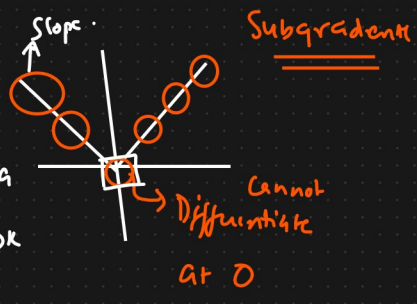
Advantage:

1. Robust to outliers. Since the formula is not squaring the error, it does not penalize the outlier. Hence if an outlier is added, MAE will increase but not as much when compared to MSE. Thus the best fit line would not deviate that much.

2. It will be in the same unit. Understanding is easier.

Disadvantage:

1. We will get a curve like this that passes through the origin and we cannot differentiate any value at 0. To solve that, we take subgradients or part of the curve and then try to compute slope for each subgradient. Thus the convergence takes more time as Optimization is a complex task.



Root Mean Squared Error:

RMSE =

Advantage:

1. Same unit because of square root.

2. Differentiable

Disadvantage:

1. Not robust to outliers.