RMSE - Root Mean Squared Root;

⇒ RMSE is the standard deviation of the errors which occur when a prediction is made on a dataset.

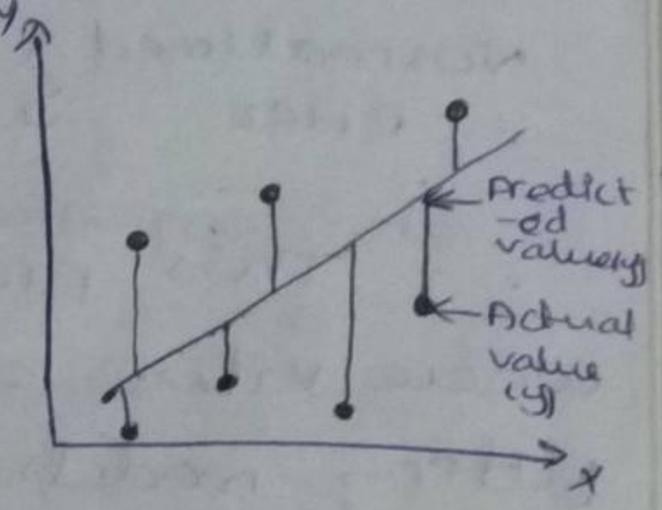
⇒ It is the standard deviation of the residuals. (prediction errors).

⇒ Residuals are a measure of how far from the regression line data points are; RMSE is a measure of how spread out these residuals are.

⇒ In other words, it tells you how concentrated the data is around the line of best fit.

⇒ one of the most commonly used measure for evaluating the quality of predictions.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)^2}{N}}$$



⇒ In RMSE, the erros are squared before they are averaged.

⇒ This basically implies that RMSE assigns a higher weight to large errors.

⇒ This indicates that RMSE is much more useful when large errors are present.

⇒ It avoids taking the absolute value of the error and this trait is useful in many mathematical calculations.

⇒ one of the ~~most~~ commonly used measures for evaluating the quality of predictions.

⇒ It shows how far predictions fall from measured true values using Euclidean distance.

What is a good RMSE value?

⇒ It depends.

⇒ the lower the RMSE, the better a given model is able to "fit" a dataset.

⇒ However the range of the dataset we're working w/d is important in determining whether or not a given RMSE value is "low" not.

Normalizing the RMSE value:

One way to gain a better understanding of whether a certain RMSE value is "good" is to normalize it using the following formula:

$$\left.\begin{matrix}\text{Normalized} \\ \text{RMSE}\end{matrix}\right\} = \frac{\text{RMSE}}{(\text{max value} - \text{min value})}$$

This produces a value b/w 0 and 1, where values closer to 0 represent better fitting models.

MAE : Mean Absolute Error:

⇒ MAE evaluates the absolute distance of the observations (the entries of dataset) to the predictions on a regression, taking the average over all observations.

⇒ We use the absolute value of the distance so that negative errors are accounted properly.

⇒ It refers to the magnitude of difference between the prediction of an

observation and the true value of that observation.

⇒ These metrics tell us how accurate our predictions are and, what is the amount of deviation from the actual values.

⇒ But it fails to punish large errors in prediction.

$$MAE = \frac{\sum_{i=1}^{N} (Predicted_i - Actual_i)}{N}$$

Batch:

⇒ The batch size is a hyperparameter that defines the no. of samples to work through before updating the internal model parameters.

⇒ A training dataset can be ~~dea~~ divided into one or more batches.

Epoch:

⇒ The no. of epochs is a hyperparameter that defines the number times that the learning algorithm will work through the entire training dataset.

⇒ One epoch means that each sample in the training dataset has had an opportunity to update the internal model parameters.

⇒ An epoch is comprised of one or more batches.

⇒ The no. of epochs is traditionally large, often 100 or 1000, allowing the learning algorithm to run until the error from the model has been sufficiently minimized.