Loss Function is also known as Cost Function or Objective Function. It is a crucial component used to quantify the error between the predicted output of a model and the actual truth.

MSE - Mean Squared Error: $1/N\sum(y_i - \hat{y}_i)^2$.

MSE is used for regression problems, where the goal is to predict continuous values. It is sensitive to outliers because of squaring.

MAE - Mean Absolute Error: $1/N\sum|y_i - \hat{y}_i|$

MAE is used for regression problems. It is less sensitive to outliers compared to MSE. It tends to be more robust when dealing with noisy data or outliers.

Huber Loss: It is used in robust regression, that is less sensitive to outliers in data than the squared error loss.

BCE - Binary Cross Entropy: $-1/N \sum(y_i.\log(\hat{y}_i) + (1-y_i).\log(1-\hat{Y}_i))$

It is used in binary classification problems as the output probability is between 0 and 1. The log nature of this loss function heavily penalizes predictions that are confidently wrong. It measures the dissimilarity between the true binary label and the predicted probability distribution.

CCE - Categorical Cross Entropy: $-1/N\sum\sum y_{i,j}.\log(\hat{y}_{i,j})$

CCE is used in multi -class classification tasks , where each example belongs to one class out of the several classes.
The output should be in the form of one-hot encoded vectors or probability distributions over classes.

Sparse Categorical Cross Entropy: This is a variant of categorical cross entropy loss used when the true labels are integers rather than one-hot encoded vectors. It converts integer labels to one-hot encoded vectors and then calculates categorical cross entropy.

Hinge Loss: It is often used in binary classification tasks, particularly with SVM. It is used for maximum margin classification, aiming to maximize the margin between classes. It is less sensitive to outliers compared to squared loss function.