Machine Learning - The Jargon Terms

Here is another set of machine learning terms.

1. Weights
2. Bias
3. Ensemble
4. Embedding
5. Tuning
6. Gradient
7. Convergence
8. Parameter
9. Hyper-Parameter
10. Inference
11. Logits
12. Optimizer
13. Overfitting
14. Underfitting
15. Regularization
16. Baseline
17. Supervised
18. Unsupervised

1.WEIGHTS

A coefficient for a feature in a linear model, or edge in a deep network.

Weight in a way represents the importance or the contribution of a feature in a model or of a neuron in a neural network towards the final output.

If weight is 0 then feature has no contribution.

2.BIAS

Bias in terms of ethics can refer to favoritism or leaning towards a certain prejudice or viewpoint that leads to erroneous results.

In terms of mathematics, bias refers to the intercept in the linear equation model.

3.ENSEMBLE

Merging multiple models with varying properties to get the final prediction outputs.

Ensemble of models can be done on various criteria like:

different initializations, hyperparameters, overall structure, etc.

4.EMBEDDING

Generally, embedding is a low-dimensional representation of a high-dimensional input.

Like representing an image into a 128-dimensional or N-dimensional vector is known as image embedding.

5.PARAMETER

Parameters are the variables that are updated during the learning process.

Weights and biases in the model can be referred to as parameters.

6.HYPERPARAMETER

These are different from parameters, parameters contribute directly to the final output calculation, while hyperparameters are required to estimate the parameters.

Hyperparameters decide how efficient the training is.

Like, learning rate, dropout rate, etc.

7.GRADIENT

The partial derivatives with respect to all of the independent variables.

In terms of the error function, the gradient is calculated and used to decide what should be the updates in the parameters of the model to make the solution more accurate.

8.CONVERGENCE

It is the state reached during the training where additional training will not significantly improve the model, and the loss is changing very little or not at all with the number of epochs.

9.TUNING

Tuning is essentially selecting the best values of hyperparameters for an algorithm to optimize its performance given a working environment.

Tuning can be done manually as well as through an automated process.

10.INFERENCE

In machine learning, often refers to the process of making predictions by applying the trained model to unlabeled examples.

Making a prediction on a real-time observation, like marking an incoming mail as spam, can be termed as inferencing.

11.LOGITS

The raw predictions that are generated by a classification model, which is ordinarily then passed to a normalization function.

If the model is solving a multi-class classification problem, logits typically become an input to the softmax function.

12.OPTIMIZER

An optimizer is the algorithm or rules to be followed for the updation of the parameters during the training process.

These are generally variations of Gradient Descent algorithm.

Like, Stochastic gradient descent, Adam, RMSprop, etc.

13.OVERFITTING

Overfitting refers to the situation where the model fits very well on the seen data but fails to make good predictions on the unseen data.

It's like being able to solve well only those problems that you have solved before but the unseen ones not that well.

14.UNDERFITTING

Underfitting occurs when a model is unable to make good predictions because it has not learned or captured the complexity of the training data sufficiently.

This could happen dues to multiple reasons, not enough parameters in the model, very few iterations, etc.

15.REGULARIZATION

Regularization is a penalizing step to prevent the model from overfitting problem.

There are several techniques that help in avoiding the overfitting problem.

Like, L1 L2 regularization, dropout, etc.

16.BASELINE

Baseline generally refers to a performance of a model that is used to compare with the performance of more complex models to have a reference of how much we have improved.

It also helps put a minimum performance benchmark.

Like, using Linear regression vs Neural Net

17.SUPERVISED

It is a type of learning where we have the inputs as well as their correct answers (labels).

It is like learning to solve problems from a book that also has the respective answers so we can check or update our learning.

18.UNSUPERVISED

Unlike supervised learning we don't have the actual labels, the data available is unlabeled.

We typically use unsupervised learning algorithms to find out the underlying patterns in the data, reducing dimensions of the data, etc. Like, clustering the data points.