



Week 8 Glossary

Agile model development

The process of increasing a model's complexity in separate efforts.

Algorithmic accountability

A concept where model owners and developers are accountable for the decisions that their machine learning systems make.

Allocative harm

A discriminatory system that withholds certain opportunities, freedoms, or resources from specific groups.

Bias

Model bias expresses a type of error a model makes (how different the prediction is from the training data). High bias means that the model is too simple and failed to capture the relationship between the features and labels; it is a sign that the model is underfitting. This happens, for example, when you make the wrong modeling assumptions, such as training a model on data for which it is not suited.

Bias-variance tradeoff

The property of a model where the variance of the parameter estimated across samples can be reduced by increasing the bias in the estimated parameters.

Class imbalance

A problem in machine learning where the total number of a class of positive data is far less than the total number of another class of negative data.

Concept drift

Changes in the statistical properties of data over time.

Ethical AI

The subfield of AI that studies model fairness and accountability.

Execution bottleneck

A data preparation or modeling process not terminating in a reasonable amount of time. This



common and frustrating issue is often caused by a misalignment among the data, tools, and hardware you use to process the data.

Fairness

The various attempts at correcting algorithmic bias in automated decision processes based on machine learning models.

Feature importance

Techniques that assign a score to input features based on how useful they are at predicting a target variable.

Feature leakage

Data leakage that occurs in information that is used in the model training process that would not be available at prediction time.

Features

Input variables which are predictive data elements of a machine learning problem. They are the data contained in the columns of a data matrix; one feature value is contained in one column.

Learning curve

A plot that shows the relationship between data size and model prediction performance.

Learning curve analysis

A method to empirically measure data size efficiency.

Logistic regression

A linear classification method that is trained by iteratively tuning a set of weights to minimize the log loss.

MapReduce

A programming framework that performs distributed and parallel processing on large data sets in a distributed environment.

Model performance failure

When evaluating a test, a performance failure describes a model that may have poor performance. This can lead to overconfidence in a system and poor generalization performance when the model is applied to live data.



Model variance

Expresses how consistent the predictions of a model are if it is trained on different sections of the training data set. High variance is a sign that the model is overfitting to the particular data set on which it is trained.

Representational harm

A type of harm where a system reinforces negative stereotypes along the lines of identity and protected class.

Reproducibility

The ability to duplicate a model exactly such that given the same raw data as input, both models return the same output.

Societal failures

Failures which happen when a machine learning model produces unintended discrimination or disparate impact and lacks accountability. It is where failure exists between the model's owner and the social context in which the model operates.

Stratified sampling

A sampling method that reduces the sampling error in cases where the population can be partitioned into subgroups.

Unit testing

A software development process in which the smallest testable parts of an application, called units, are individually and independently tested for proper operation.

Upsampling

A strategy of taking 100 percent of the negative classes and sampling the positive class cases with replacement until you get equal sizes for both. Upsampling is a preferred strategy when you have limited data to begin with and can't afford to discard any.

