Artificial Intelligence Vocabulary List

February 2025

General Terms

Artificial Intelligence: the capability of computer systems or algorithms to imitate intelligent human behaviour

Generative AI: artificial intelligence that is capable of generating new content in response to a submitted prompt by learning from a large reference database of examples

Machine Learning: a subfield of artificial intelligence that enables a computer to learn to perform tasks by analyzing a large dataset without being explicitly programmed

Deep Learning: a form of machine learning in which the computer network rapidly teaches itself to understand a concept without human intervention by performing a large number of iterative calculations on an extremely large dataset

Supervised machine learning: requires labelled data to adjust the parameters of the model during training

Unsupervised machine learning: uses machine learning algorithms to analyze and cluster unlabeled datasets

Reinforcement machine learning: trains machines through trial and error to take the best action by establishing a reward system

Neural network: a computer architecture in which a number of processors are interconnected in a manner suggestive of the connections between neurons in a human brain and which is able to learn by a process of trial and error

Model: programs that detect specific patterns using a collection of data sets. It is an illustration of a system that can receive data inputs and draw conclusions or conduct actions depending on those conclusions.

- The number of layers in the model refers to the depth of the neural network. Adjusting the number of layers is essential because a model with too few layers may not be able to capture the complexity of the data, while a model with too

many layers may overfit the training data and not generalize well to new, unseen data.

- The number of neurons in each layer refers to the width of each layer in the neural network. Adjusting the number of neurons is important as too few neurons may not be able to capture the underlying patterns of the data, while too many neurons may lead to overfitting and increased computational cost.

Regression Model: a model that estimates values within a continuous range (such as price and revenue), finds correlations between dependent and independent variables. These models are used to help predict continuous variables such as house prices, market trends, weather patterns, oil and gas prices, etc.

Classification Model: a model that divides the dataset into classes based on various parameters. When using a Classification algorithm, a computer program gets taught the categories on the training dataset and then categorizes the data into various categories depending on what it learned.

Large Language Model (LLM): a language model that utilizes deep methods on an extremely large data set as a basis for predicting and constructing natural-sounding text

Natural Language Processing (NLP): knowledge that is crucial for preprocessing text data, selecting relevant features, and understanding the linguistic nuances that the AI model needs to capture

Chatbot: a bot algorithm with a predefined dataset that is designed to converse with human beings

Algorithm: a procedure for solving a mathematical problem in a finite number of steps that frequently involves repetition of an operation

Dataset: a set a data used to train and test algorithms and models

Big data: an accumulation of data that is too large and complex for processing by traditional database management tools. Today, big data can refer to large data sets or to systems and solutions developed to manage such large accumulations of data, as well as for the branch of computing devoted to this development.

Bias: systematic error introduced into sampling or testing by selecting or encouraging one outcome or answer over others

Parameter: an arbitrary constant whose value characterizes a member of a system

Learning rate: The learning rate refers to a hyperparameter (setting when configuring a learning model) that determines the size of the steps that the model takes during the training process. A too small learning rate makes the model learn very slowly, while a large learning rate may make the model oscillate or overshoot the minimum.

Artificial Intelligence Performance Metrics

Accuracy: the proportion of all classifications that were correct, whether positive or negative, showing how often a classification model is correct overall. A perfect model would have zero false positives and zero false negatives and therefore an accuracy of 1.0, or 100%.

Precision: shows how often an model is correct when predicting the target or goal state, showing the proportion of all the model's positive classifications that are actually positive. A hypothetical perfect model would have zero false positives and therefore a precision of 1.0.

Recall: the ratio of the number of true positive predictions, i.e. the number of positive instances correctly identified by the model to the total number of actual positive instances in the dataset. False negatives are actual positives that were misclassified as negatives. A hypothetical perfect model would have zero false negatives and therefore a recall (TPR) of 1.0, which is to say, a 100% detection rate.

Source:
https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall