Supervised learning

Unsupervised learning

Semi-supervised learning

Self-supervised learning

Reinforcement learning

Meta-learning

Online learning

Batch learning

Curriculum learning

Rule-based learning

Neuro-symbolic AI

Neuromorphic engineering

Quantum machine learning

Classification

Generative modeling

Regression

Clustering

Dimensionality reduction

Density estimation

Anomaly detection

Data cleaning

AutoML

Association rules

Semantic analysis

Structured prediction

Feature engineering

Feature learning

Learning to rank

Grammar induction

Ontology learning

Multimodal learning

Apprenticeship learning

Decision trees

Ensembles Bagging Boosting Random forest

Bagging

Boosting

Random forest

k -NN

Linear regression

Naive Bayes

Artificial neural networks

Logistic regression

Perceptron

Relevance vector machine (RVM)

Support vector machine (SVM)

BIRCH

CURE

Hierarchical

k -means

Fuzzy

Expectation–maximization (EM)

DBSCAN

OPTICS

Mean shift

Factor analysis

CCA

ICA

LDA

NMF

PCA

PGD

t-SNE

SDL

Graphical models Bayes net Conditional random field Hidden Markov

Bayes net

Conditional random field

Hidden Markov

RANSAC

k -NN

Local outlier factor

Isolation forest

Autoencoder

Deep learning

Feedforward neural network

Recurrent neural network LSTM GRU ESN reservoir computing

LSTM

GRU

ESN

reservoir computing

Boltzmann machine Restricted

Restricted

GAN

Diffusion model

SOM

Convolutional neural network U-Net LeNet AlexNet DeepDream

U-Net

LeNet

AlexNet

DeepDream

Neural field Neural radiance field Physics-informed neural networks

Neural radiance field

Physics-informed neural networks

Transformer Vision

Vision

Mamba

Spiking neural network

Memtransistor

Electrochemical RAM (ECRAM)

Q-learning

Policy gradient

SARSA

Temporal difference (TD)

Multi-agent Self-play

Self-play

Active learning

Crowdsourcing

Human-in-the-loop

Curriculum learning is a technique in machine learning in which a model is trained on examples of increasing difficulty, where the definition of "difficulty" may be provided externally or discovered as part of the training process. This is intended to attain good performance more quickly, or to converge to a better local optimum if the global optimum is not found. [1] [2]

Approach

Most generally, curriculum learning is the technique of successively increasing the difficulty of examples in the training set that is presented to a model over multiple training iterations. This can produce better results than exposing the model to the full training set immediately under some circumstances; most typically, when the model is able to learn general principles from easier examples, and then gradually incorporate more complex and nuanced information as harder

examples are introduced, such as edge cases . This has been shown to work in many domains, most likely as a form of regularization . [3]

There are several major variations in how the technique is applied:

A concept of "difficulty" must be defined. This may come from human annotation [4] [5] or an external heuristic ; for example in language modeling , shorter sentences might be classified as easier than longer ones. [6] Another approach is to use the performance of another model, with examples accurately predicted by that model being classified as easier (providing a connection to boosting).

Difficulty can be increased steadily [7] or in distinct epochs, [8] and in a deterministic schedule or according to a probability distribution . This may also be moderated by a requirement for diversity at each stage, in cases where easier examples are likely to be disproportionately similar to each other. [9]

Applications must also decide the schedule for increasing the difficulty. Simple approaches may use a fixed schedule, such as training on easy examples for half of the available iterations and then all examples for the second half. [3] Other approaches use self-paced learning to increase the difficulty in proportion to the performance of the model on the current set. [10]

Since curriculum learning only concerns the selection and ordering of training data, it can be combined with many other techniques in machine learning. The success of the method assumes that a model trained for an easier version of the problem can generalize to harder versions, so it can be seen as a form of transfer learning . Some authors also consider curriculum learning to include other forms of progressively increasing complexity, such as increasing the number of model parameters. [11] It is frequently combined with reinforcement learning , such as learning a simplified version of a game first. [12]

Some domains have shown success with anti-curriculum learning : training on the most difficult examples first. One example is the ACCAN method for speech recognition , which trains on the examples with the lowest signal-to-noise ratio first. [13]

History

The term "curriculum learning" was introduced by Yoshua Bengio et al in 2009, [14] with reference to the psychological technique of shaping in animals and structured education for humans: beginning with the simplest concepts and then building on them. The authors also note that the application of this technique in machine learning has its roots in the early study of neural networks such as Jeffrey Elman 's 1993 paper Learning and development in neural networks: the importance of starting small . [15] Bengio et al showed good results for problems in image classification , such as identifying geometric shapes with progressively more complex forms, and language modeling , such as training with a gradually expanding vocabulary . They conclude that, for curriculum strategies, "their beneficial effect is most pronounced on the test

set", suggesting good generalization.

The technique has since been applied to many other domains:

Natural language processing : Part-of-speech tagging [16] Intent detection [17] Sentiment analysis [18] Machine translation [19] [20] Speech recognition [21] Language model pre-training [22]

Part-of-speech tagging [16]

Intent detection [17]

Sentiment analysis [18]

Machine translation [19] [20]

Speech recognition [21]

Language model pre-training [22]

Image recognition : Facial recognition [23] Object detection [24]

Facial recognition [23]

Object detection [24]

Reinforcement learning : Game-playing [25]

Game-playing [25]

Graph learning [26] [27]

Matrix factorization [28]

References

Further reading

Curriculum Learning: A Survey

A Survey on Curriculum Learning

Curriculum Learning for Reinforcement Learning Domains: A Framework and Survey

Curriculum learning at IEEE Xplore