Module 4 Glossary: Unsupervised Learning and Generative Models in Keras

Welcome! This alphabetized glossary contains many terms used in this course. Understanding these terms is essential when working in the industry, participating in user groups, and participating in other certificate programs.

Term	Definition
Anomaly detection	A type of unsupervised learning used to identify unusual data points that do not fit the general pattern in a dataset.
Adam Optimizer	An optimization algorithm that can be used instead of the classical stochastic gradient descent procedure to update network weights iteratively based on training data.
Adversarial training	A training method in GANs where the generator and discriminator networks are trained simultaneously, with the generator aiming to fool the discriminator, and the discriminator trying to accurately classify real and fake data.
Autoencoders	A type of neural network used to learn efficient representations of data, often for dimensionality reduction or feature learning.
Binary crossentropy loss	A loss function used in binary classification tasks, often utilized in training neural networks to measure the difference between predicted and actual outputs.
Bottleneck	The central, most compressed layer in an autoencoder that contains the most critical features of the input data.
Clustering	A method in unsupervised learning that involves grouping data points into clusters, where data points in the same cluster are more similar to each other.
Convolutional autoencoders	A type of autoencoder that uses convolutional layers, making it particularly effective for tasks involving image data.
Convolutional neural network (CNN)	A type of deep neural network commonly used in image processing tasks, known for its ability to capture spatial hierarchies in images.
Data augmentation	A technique used to increase the diversity of data available for training models by generating new synthetic data, often by applying transformations to existing data.
Decoder	The part of an autoencoder that reconstructs the input data from the compressed latent space representation.
Denoising	The process of removing noise from data, such as images, to improve their quality.
Diffusion	A physical process where particles spread from regions of high concentration to low concentration; in diffusion models, this concept is simulated to generate or enhance data.
Diffusion model	A type of probabilistic generative model that iteratively refines noisy data to produce high-quality samples, often used in image generation.
Dimensionality reduction	A process in unsupervised learning that reduces the number of random variables under consideration by obtaining a set of principal variables.
Discriminator network	In a generative adversarial network (GAN), this network evaluates the authenticity of the generated data, distinguishing between real and fake data.
Encoder	The part of an autoencoder that compresses the input data into a latent-space representation.
Epochs	In machine learning, an epoch refers to one complete pass of the training dataset through the learning algorithm.
Feature learning	A set of techniques that allow a machine to automatically discover the representations needed for feature detection or classification from raw data.
Forward process	In diffusion models, the process of gradually adding noise to data over a series of steps.
Functional API	A way to build neural networks in Keras that allows for more flexible model architectures than the Sequential API.
Generative adversarial networks (GANs)	A class of neural networks where two networks, the generator and the discriminator, compete against each other, leading to the generation of realistic data.
Generator	In GANs, the neural network creates synthetic data from random noise, aiming to produce data that closely resembles real data.
Image-to-image translation	A task in computer vision where an image from one domain is transformed into an image in another domain, such as converting a sketch into a photo.
K-means algorithm	A popular clustering technique that partitions a dataset into distinct groups based on the features of the data points.
Keras	An open-source software library that provides a Python interface for artificial neural networks and is used to create deep learning models.
Latent-space representation	The compressed version of input data generated by the encoder in an autoencoder.
MNIST dataset	A large database of handwritten digits that is commonly used for training image processing systems and machine learning models.
Modified National Institute of Standards and Technology (MNIST) Dataset	A large database of handwritten digits commonly used for training various image processing systems.
Neural network architecture	The structured layout of a neural network, including its layers, connections, and the flow of information within it.
Normalization	The process of scaling input features so they have a mean of zero and a standard deviation of one, often used to improve the performance of neural networks.
Principal component analysis (PCA)	A dimensionality reduction technique that transforms data into a set of linearly uncorrelated variables called principal components.
Probabilistic model	A model that incorporates randomness and uncertainty, often used to predict distributions or simulate processes that have inherent variability.
Reverse process	In diffusion models, the process of removing noise step by step to reconstruct the original data from a noisy sample.
Stochastic gradient descent (SGD)	An optimization method that adjusts weights iteratively based on a subset of training data, used in training neural networks.
Supervised learning	A type of machine learning where the algorithm is trained on labeled data, meaning the outcome or target variable is known during training.
t-distributed stochastic neighbor embedding (t-SNE)	A dimensionality reduction technique used for visualizing high-dimensional data by giving each datapoint a location in a two or three-dimensional map.
TensorFlow	An open-source machine learning library developed by Google, widely used for building and training machine learning models.
Text-to-image synthesis	A process where a model generates an image based on a textual description provided as input.
Training epoch	A single pass through the entire training dataset during the training process of a machine learning model.
Unsupervised learning	A type of machine learning that finds patterns in data without any labels or predefined outcomes.
Variational autoencoders (VAEs)	A type of autoencoder that introduces probabilistic elements to generate new data samples, often used in generative models.
Zero-Sum Game	A situation in competitive contexts where gain or loss of participants is exactly balanced by the losses or gains of another participant.