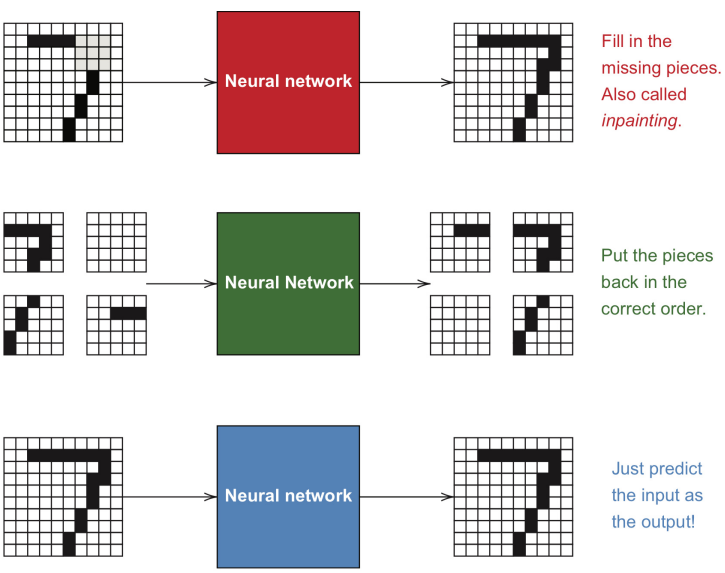
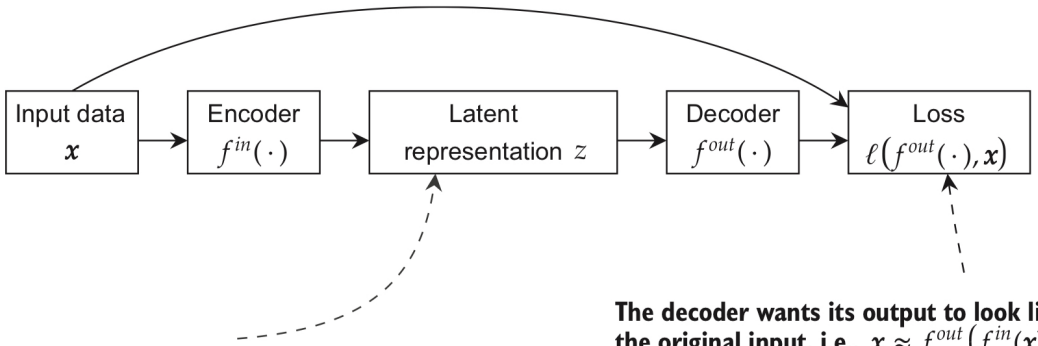
Properties	Unsupervised learning	Supervised learning
Definition	Unsupervised learning is the type of machine learning that happens without human supervision. A machine tries to find any patterns in data by itself.	Supervised learning is the type of machine learning that happens under human supervision, meaning people label input data with answer keys showing a machine the desired outputs.
Input data	Unlabeled	Labeled
Use of data	A model is given only input variables (X) and no corresponding output data.	A model is given input variables (X), output variables (Y), and an algorithm to learn the function from input to output.
When to use	You don't know what you're looking for in data.	You know what you're looking for in data.
Applicable in	Clustering and association problems	Classification and regression problems
Accuracy of the results	May provide less accurate results	Provides more accurate results
Algorithms	• K-Means • Gaussian Mixture Models • Frequent Pattern (FP) Growth • Principal Component Analysis	• Support vector machines • Decision trees • Random forest • Naive Bayes
Use cases	• Recommender systems • Anomaly detection • Customer segmentation • Preparing data for supervised learning	• Spam filters • Demand forecasting • Price prediction • Image recognition

altexsoft

Self-supervised learning = use a regression or classification loss function where the label is the data itself



Autoencoder



Useful for visualization (2D or 3D scatter plot), clustering (fed to your favorite algorithm like k-means or HDBSCAN), information retrieval/search (k-nearest neighbor searching), or as the input to another classifier (e.g, SVMs, random forests, etc.)

The decoder wants its output to look like the original input, i.e., $x \approx f^{out}\left(f^{in}(x)\right)$. Once done with training, cases where $x \not\approx f^{out}\left(f^{in}(x)\right)$ have a good chance of being outliers or otherwise interesting datapoints worth investigation (e.g., was it mislabeled? Is it a special case?).