anomaly detection ---- is a ---- application

autoencoder ---- can do ---- anomaly detection

autoencoder ---- can do ---- dimensionality reduction

autoencoder ---- can do ---- feature extraction

autoencoder ---- can do ---- feature learning

autoencoder ---- can do ---- fortune telling

autoencoder ---- can do ---- image compression

autoencoder ---- can do ---- machine translation

autoencoder ---- can do ---- noise reduction

autoencoder ---- can do ---- speech recognition

autoencoder ---- description ---- "An autoencoder is a type of artificial neural network used to learn efficient data codings in an unsupervised manner. - Wikipedia; It learns good representations in the hidden layer, capturing the essence of the input"

autoencoder ---- has advantage ---- "Accomplishes roughly the same thing as PCA, but with nonlinearity"

autoencoder ---- has advantage ---- deep autoencoders are much better than PCA at extracting non-linear features

autoencoder ---- has disadvantage ---- requires training

autoencoder ---- is a ---- method

autoencoder ---- link ---- "https://en.wikipedia.org/wiki/Autoencoder"

autoencoder ---- subclass of ---- unsupervised learning

autoregressive models ---- can do ---- feature extraction

autoregressive models ---- can do ---- image generation

autoregressive models ---- description ---- "Used to describe certain time-varying processes. Specifies that the output variable depends linearly on its own previous values and on a

stochastic term (an imperfectly predictable term). -Wikipedia; Given examples from a distribution, generate samples belonging to the same distribution. -CSE598"

autoregressive models ---- has advantage ---- "can explicitly estimate likelihood p(x)"

autoregressive models ---- has advantage ---- "pretty good results for generation"

autoregressive models ---- has advantage ---- "simple and stable training process"

autoregressive models ---- has disadvantage ---- "cannot provide low-dimensional latent space"

autoregressive models ---- has disadvantage ---- "inefficient at generation because it is a sequential process, which cannot be parallelized"

autoregressive models ---- is a ---- method

autoregressive models ---- link ---- "https://en.wikipedia.org/wiki/Autoregressive_model"

bayesian network ---- can do ---- anomaly detection

bayesian network ---- can do ---- inference

bayesian network ---- can do ---- spam filtering

bayesian network ---- can do ---- time series forecasting

bayesian network ---- description ---- "A probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graph (DAG). Ideal for taking an event that occurred and predicting the likelihood that any one of several possible known causes was the contributing factor. -Wikipedia"

bayesian network ---- has advantage ---- "can encode conditional independencies"

bayesian network ---- has disadvantage ---- "cannot automatically learn the structure of data; must be encoded manually"

bayesian network ---- has disadvantage ---- "must assume prior distributions of variables"

bayesian network ---- is a ---- method

bayesian network ---- link ---- "https://en.wikipedia.org/wiki/Bayesian_network"

classification ---- is a ---- application

convolutional neural network ---- can do ---- anomaly detection

convolutional neural network ---- can do ---- image classification

convolutional neural network ---- can do ---- natural language processing

convolutional neural network ---- can do ---- noise reduction

convolutional neural network ---- can do ---- recommendation

convolutional neural network ---- can do ---- video recognition

convolutional neural network ---- description ---- "A type of deep neural network that takes advantage of the hierarchical pattern in data and assembles patterns of increasing complexity using smaller and simpler patterns embossed in their filters. -Wikipedia"

convolutional neural network ---- has advantage ---- "independent from prior human knowledge in feature extraction"

convolutional neural network ---- has advantage ---- "learns optimal filters automatically"

convolutional neural network ---- has disadvantage ---- "there are many hyperparameters to optimize"

convolutional neural network ---- has disadvantage ---- not interpretable

convolutional neural network ---- is a ---- method

convolutional neural network ---- link ---- "https://en.wikipedia.org/wiki/Convolutional_neural_network"

data analysis ---- is a ---- application

data clustering ---- is a ---- application

decision tree ---- can do ---- capture interactions in data

decision tree ---- can do ---- classification

decision tree ---- can do ---- data visualization

decision tree ---- can do ---- feature importance

decision tree ---- can do ---- recommendation

decision tree ---- can do ---- regression

decision tree ---- description ---- "A decision tree is a flowchart-like structure in which each internal node represents a test on an attribute (e.g. whether a coin flip comes up heads or tails), each branch represents the outcome of the test, and each leaf node represents a class label (decision taken after computing all attributes). The paths from root to leaf represent classification rules. -Wikiedia"

decision tree ---- example ---- 001-DecisionTree

decision tree ---- has advantage ---- "can deal with data where features interact with each other"

decision tree ---- has advantage ---- "can deal with large, complicated, differently-scaled data"

decision tree ---- has advantage ---- "can deal with non-linear data"

decision tree ---- has advantage ---- "creates good explanations"

decision tree ---- has advantage ---- interpretable

decision tree ---- has disadvantage ---- "fails to deal with linear relationships without adding splits, which is inefficient"

decision tree ---- has disadvantage ---- "lack of smoothness; slight changes in an input feature can have a big impact on the predicted outcome"

decision tree ---- has disadvantage ---- "only easily interpretable when short"

decision tree ---- has disadvantage ---- unstable

decision tree ---- is a ---- method

decision tree ---- link ---- "https://en.wikipedia.org/wiki/Decision_tree"

dimensionality reduction ---- is a ---- application

feature extraction ---- is a ---- application

feature importance ---- is a ---- application

feature learning ---- is a ---- application

generative adversarial network ---- can do ---- image generation

generative adversarial network ---- description ---- "Consists of two neural networks that compete with each other: The generator, which tries to generate realistic images from noise, and the discriminator, which tries to tell whether a given image is real or fake. There are thousands of papers on GANs in the literature and hundreds of applications -ASU CSE598"

generative adversarial network ---- has advantage ---- "very realistic images"

generative adversarial network ---- has disadvantage ---- "can be notoriously hard to optimize"

generative adversarial network ---- has disadvantage ---- "does not provide the inference p(x) - the probability of the input image x"

generative adversarial network ---- is a ---- method

generative adversarial network ---- link ---- "https://en.wikipedia.org/wiki/Generative_adversarial_network"

generative adversarial network ---- link ---- "https://www.coursera.org/learn/cse598deeplearninginvisualcomputing/lecture/gAijA/generative-adversarial-networks-gans"

graph clustering ---- is a ---- application

image classification ---- is a ---- application

image compression ---- is a ---- application

image generation ---- is a ---- application

inference ---- is a ---- application

k-means clustering ---- can do ---- data clustering

k-means clustering ---- can do ---- feature learning

k-means clustering ---- can do ---- recommendation

k-means clustering ---- description ---- "an unsupervised clustering algorithm that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean (or centroid), serving as a prototype of the cluster. -Wikipedia, ASU CSE 575"

k-means clustering ---- has advantage ---- "modifications such as k-means++ compensate for initialization sensitivity"

k-means clustering ---- has advantage ---- very simple

k-means clustering ---- has disadvantage ---- "finding the optimal solution is NP-hard"

k-means clustering ---- has disadvantage ---- sensitive to initial parameters

k-means clustering ---- is a ---- method

k-means clustering ---- link ---- "https://en.wikipedia.org/wiki/K-means_clustering"

k-means clustering ---- subclass of ---- unsupervised learning

k-nearest-neighbors ---- can do ---- anomaly detection

k-nearest-neighbors ---- can do ---- classification

k-nearest-neighbors ---- can do ---- recommendation

k-nearest-neighbors ---- description ---- "classifies a data point based on the classes of the nearest k neighbors in the feature space; assumes that similar things are near each other"

k-nearest-neighbors ---- has advantage ---- strong consistency results

k-nearest-neighbors ---- has advantage ---- very simple

k-nearest-neighbors ---- has disadvantage ---- "sensitive to the local structure of the data -Wikipedia"

k-nearest-neighbors ---- has disadvantage ---- computationally expensive for large data sets

k-nearest-neighbors ---- is a ---- method

k-nearest-neighbors ---- link ---- "https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm"

k-nearest-neighbors ---- subclass of ---- supervised learning

knowledge representation and reasoning ---- subclass of ---- artificial intelligence

lasso ---- can do ---- feature selection

lasso ---- description ---- "Introduces sparsity into a linear regression model that has lots of features; performs feature selection and regularization of the selected feature weights; - from Interpretable Machine Learning by Christoph Molnar."

lasso ---- is a ---- method

lasso ---- link ---- "https://en.wikipedia.org/wiki/Lasso_(statistics)"

linear regression ---- can do ---- feature importance

linear regression ---- can do ---- price prediction

linear regression ---- can do ---- regression

linear regression ---- can do ---- time series forecasting

linear regression ---- description ---- "Models the relationship between a scalar response and one or more explanatory variables. A fitted linear regression model can be used to identify the relationship between a single predictor variable xj and the response variable y when all the other predictor variables in the model are held fixed."

linear regression ---- has advantage ---- interpretable

linear regression ---- has advantage ---- very simple

linear regression ---- has advantage ---- widely used

linear regression ---- has disadvantage ---- "Assumes constant variance (homoscedasticity)"

linear regression ---- has disadvantage ---- "Assumes independence of errors"

linear regression ---- has disadvantage ---- "loses interpretability when there are more than a few features. Lasso can help with this."

linear regression ---- has disadvantage ---- "non-linearity has to be hand-crafted"

linear regression ---- is a ---- method

linear regression ---- link ---- "https://en.wikipedia.org/wiki/Linear_regression"

linear regression ---- subclass of ---- supervised learning

logistic regression ---- can do ---- classification

logistic regression ---- can do ---- image classification

logistic regression ---- description ---- " The logistic model (or logit model) is used to model the probability of a certain class or event existing such as pass/fail, win/lose, etc. When a cutoff probability is specified, the model acts as a binary classifier. Can also be extended into a multi-class classifier."

logistic regression ---- has advantage ---- interpretable

logistic regression ---- has advantage ---- very simple

logistic regression ---- has advantage ---- widely used

logistic regression ---- has disadvantage ---- "can suffer from Complete Separation - when a feature perfectly separates two classes, the model cannot be trained. In this case, a simple rule would work better anyway. - from Interpretable Machine Learning by Christoph Molnar"

logistic regression ---- has disadvantage ---- "interactions must be added manually"

logistic regression ---- has disadvantage ---- "other models may have better predictive performance"

logistic regression ---- is a ---- method

logistic regression ---- link ---- "https://en.wikipedia.org/wiki/Logistic_regression"

logistic regression ---- subclass of ---- supervised learning

long short-term memory ---- can do ---- anomaly detection

long short-term memory ---- can do ---- machine translation

long short-term memory ---- can do ---- natural language processing

long short-term memory ---- can do ---- speech recognition

long short-term memory ---- can do ---- time series forecasting

long short-term memory ---- description ---- "LSTM - A type of recurrent neural network (RNN) that addresses the vanishing gradient problem. A common LSTM unit is composed of a cell, an input gate, an output gate, and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell. -Wikipedia"

long short-term memory ---- has advantage ---- "does not suffer from the vanishing/exploding gradient problem"

long short-term memory ---- has advantage ---- "relatively insensitive to gap length -Wikipedia"

long short-term memory ---- has disadvantage ---- "has many more parameters than other methods"

long short-term memory ---- is a ---- method

long short-term memory ---- link ---- "https://en.wikipedia.org/wiki/Long_short-term_memory"

machine learning ---- definition ---- "The study of computer algorithms that improve automatically through experience and by the use of data"

machine learning ---- subclass of ---- artificial intelligence

machine translation ---- is a ---- application

naive bayes ---- can do ---- classification

naive bayes ---- can do ---- recommendation

naive bayes ---- description ---- "A family of probabilistic classifiers based on applying Bayes theorem with strong (naive) independence assumptions between the features. Often uses maximum likelihood for parameter estimation. -Wikipedia"

naive bayes ---- has advantage ---- "actually work quite well in many complex real-world situations -Wikipedia"

naive bayes ---- has advantage ---- "only requires a small number of training data to estimate the parameters necessary for classification -Wikipedia (no citation)"

naive bayes ---- has advantage ---- very simple

naive bayes ---- has disadvantage ---- "outperformed by many other classifiers such as boosted trees or random forests -Wikipedia"

naive bayes ---- is a ---- method

naive bayes ---- link ---- "https://en.wikipedia.org/wiki/Naive_Bayes_classifier"

naive bayes ---- subclass of ---- supervised learning

natural language processing ---- is a ---- application

noise reduction ---- is a ---- application

price prediction ---- is a ---- application

principal component analysis ---- can do ---- data analysis

principal component analysis ---- can do ---- dimensionality reduction

principal component analysis ---- can do ---- feature extraction

principal component analysis ---- can do ---- feature learning

principal component analysis ---- description ---- "PCA is the process of computing the principal components (the axes with the highest variance - in other words, the most descriptive directions) and using them to perform a change of basis on the data, sometimes using only the first few principal components and ignoring the rest."

principal component analysis ---- has advantage ---- "Generates *new* features that may be highly descriptive of the data, as opposed to mere feature selection -ASU CSE 575"

principal component analysis ---- has disadvantage ---- "Assumes linearity of data; but non-linear extensions exist"

principal component analysis ---- has disadvantage ---- "when multiple classes overlap in one of the non-principal component dimensions, misclassification will get bad. In such cases, Linear Discriminant Analysis (LDA) may be a better approach."

principal component analysis ---- is a ---- method

principal component analysis ---- link ---- "https://en.wikipedia.org/wiki/Principal_component_analysis"

principal component analysis ---- subclass of ---- unsupervised learning

random forest ---- can do ---- classification

random forest ---- can do ---- regression

random forest ---- description ---- "An ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the individual trees. -Wikipedia"

random forest ---- has advantage ---- "Corrects for decision trees habit of overfitting their training sets"

random forest ---- has advantage ---- "Robust to outlying observations"

random forest ---- has advantage ---- "generally outperforms decision trees -Wikipedia"

random forest ---- has disadvantage ---- not interpretable

random forest ---- is a ---- method

random forest ---- link ---- "https://en.wikipedia.org/wiki/Random_forest"

random forest ---- subclass of ---- supervised learning

recommendation ---- is a ---- application

recurrent neural network ---- can do ---- machine translation

recurrent neural network ---- can do ---- natural language processing

recurrent neural network ---- can do ---- speech recognition

recurrent neural network ---- can do ---- time series forecasting

recurrent neural network ---- description ---- "A class of artificial neural networks where feedback connections between nodes form a directed graph along a temporal sequence. This allows it to exhibit temporal dynamic behavior. RNNs can use their internal state (memory) to process variable length sequences of inputs. -Wikipedia"

recurrent neural network ---- has advantage ---- "has fewer parameters than LSTM"

recurrent neural network ---- has disadvantage ---- "suffers from vanishing/exploding gradient problem"

recurrent neural network ---- is a ---- method

recurrent neural network ---- link ---- "https://en.wikipedia.org/wiki/Recurrent_neural_network"

regression ---- is a ---- application

reinforcement learning ---- definition ---- "An area of machine learning where agents acting in an environment learn how to maximize a reward"

reinforcement learning ---- subclass of ---- machine learning

spam filtering ---- is a ---- application

spectral clustering ---- can do ---- graph clustering

spectral clustering ---- description ---- "Makes use of the spectrum (eigenvalues) of the similarity matrix of the data to perform dimensionality reduction before clustering in fewer dimensions. The similarity matrix is provided as an input and consists of a quantitative assessment of the relative similarity of each pair of points in the dataset. -Wikipedia"

spectral clustering ---- has advantage ---- "does not make strong assumptions on the forms of the clusters -ASU CSE 575 course"

spectral clustering ---- has advantage ---- "easy and efficient even for large data sets as long as the similarity graph is sparse"

spectral clustering ---- has advantage ---- "good clustering results"

spectral clustering ---- has disadvantage ---- sensitive to initial parameters

spectral clustering ---- is a ---- method

spectral clustering ---- link ---- "https://en.wikipedia.org/wiki/Spectral_clustering"

speech recognition ---- is a ---- application

supervised learning ---- definition ---- "A subset of machine learning where the model trains on labelled data"

supervised learning ---- subclass of ---- machine learning

support vector machine ---- can do ---- anomaly detection

support vector machine ---- can do ---- classification

support vector machine ---- can do ---- image classification

support vector machine ---- description ---- "An SVM maps training examples to points in space so as to maximize the width of the gap between the two categories. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. -Wikipedia"

support vector machine ---- has advantage ---- "can handle non-linear classification using kernel trick, mapping inputs into high-dimensional feature spaces"

support vector machine ---- has advantage ---- "one of the most robust prediction methods -Wikipedia"

support vector machine ---- has disadvantage ---- "requires full labeling of input data"

support vector machine ---- is a ---- method

support vector machine ---- link ---- "https://en.wikipedia.org/wiki/Support-vector_machine"

support vector machine ---- subclass of ---- supervised learning

support vector regression ---- can do ---- regression

support vector regression ---- can do ---- time series forecasting

support vector regression ---- has advantage ---- "high degree of generalization when introduced to previously unseen data - Jain et.al."

support vector regression ---- has advantage ---- "much faster than neural networks on large high-resolution data sets -Forecasting energy consumption of multi-family... by Jain et.al."

support vector regression ---- is a ---- method

support vector regression ---- link ---- "https://www.analyticsvidhya.com/blog/2020/03/support-vector-regression-tutorial-for-machine-learning/"

time series forecasting ---- is a ---- application

unsupervised learning ---- subclass of ---- machine learning

video recognition ---- is a ---- application