ITUPublications

International Telecommunication Union
Standardization Sector

ITU-T Focus Group Technical Report

(10/2022)

Focus Group on Artificial Intelligence for Natural Disaster Management

Glossary – Artificial Intelligence for Natural Disaster Management

ITU-T FG-AI4NDM-GLOS Technical Report

WS-Glossary – Artificial Intelligence for Natural Disaster Management

Summary

This Technical Report encompasses the relevant list of definitions within the trans-disciplinary domain of Artificial Intelligence and Natural Disaster Management (AI4NDM).

Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Acknowledgement

This Technical Report was prepared under the leadership of Ms. Monique Kuglitsch, Chair of FG-AI4NDM (Fraunhofer HHI, Germany) and Ms Elena Xoplaki, Justus Liebig University Giessen, Germany.

It is based on the contributions of various authors who participated in the Focus Group activities.

Ms Elena Xoplaki (Justus Liebig University Giessen, Germany) served as the main Editor of this Technical Report. Ms Mythili Menon (FG-AI4NDM Advisor) and Ms Hiba Tahawi (FG-AI4NDM Assistant) served as the FG-AI4NDM Secretariat.

Change Log

This document contains Version 1 of the Focus Group Report – Glossary – Artificial Intelligence for Natural Disaster Management approved at the seventh meeting of FG-AI4NDM on 24-26 October 2022.

Editor: Elena Xoplaki
Justus Liebig University Giessen
Germany
Email: elena.xoplaki@geogr.uni-giessen.de

Table of Contents

ITU-T FG-AI4NDM Deliverable

WS-Glossary – Artificial Intelligence for Natural Disaster Management

1 Scope

This glossary contains selected key terms, along with their definitions, falling within the scope of the ITU/WMO/UNEP Focus Group on Artificial Intelligence for Natural Disaster Management (FG-AI4NDM). It aims to ensure direct and unambiguous communication among working groups, topic groups, decision-makers, stakeholders, practitioners and scientists from different fields, and to provide the appropriate information for understanding the basic concepts of AI and natural disasters.

2 References

[ITU-T E.102] Recommendation ITU-T E.102 (2019), *Terms and definitions for disaster relief systems, network resilience and recovery*.

[ITU-T E.475] Recommendation ITU-T E.475 (2020), *Guidelines for intelligent network analytics and diagnostics*.

[ITU-T E.806] Recommendation ITU-T E.806 (2019), *Measurement campaigns, monitoring systems and sampling methodologies to monitor the quality of service in mobile networks*.

[ITU-T F.791] Recommendation ITU-T F.791 (2018), *Accessibility terms and definitions*.

[ITU-T H.222.0] Recommendation ITU-T H.222.0 (2021), *Information technology – Generic coding of moving pictures and associated audio information: Systems*.

[ITU-T H.752] Recommendation ITU-T H.752 (2015), *Multimedia content provisioning interface for IPTV services*.

[ITU-T I.312] Recommendation ITU-T I.312/Q.1201 (1992), *Principles of intelligent network architecture*.

[ITU-T J.193] Recommendation ITU-T J.193 (2004), *Requirements for the next generation of set-top-boxes*.

[ITU-T J.340] Recommendation ITU-T J.340 (2010), *Reference algorithm for computing peak signal to noise ratio of a processed video sequence with compensation for constant spatial shifts, constant temporal shift, and constant luminance gain and offset*.

[ITU-T J.1014] Recommendation ITU-T J.1014 (2020), *Embedded common interface for exchangeable CA/DRM solutions; Advanced security – ECI-specific functionalities*.

[ITU-T L.1022] Recommendation ITU-T L.1022 (2019), *Circular economy: Definitions and concepts for material efficiency for information and communication technology*.

[ITU-T L.1202] Recommendation ITU-T L.1202 (2015), *Methodologies for evaluating the performance of an up to 400 VDC power feeding system and its environmental impact*.

[ITU-T L.1301] Recommendation ITU-T L.1301 (2015), *Minimum data set and communication interface requirements for data centre energy management*.

[ITU-T L.1430] Recommendation ITU-T L.1430 (2013), *EMC, resistibility and safety requirements and guidance for determining responsibility under co-located information and communication technology installations*.

[ITU-T L.1501] Recommendation ITU-T L.1501 (2014), *Best practices on how countries can utilize ICTs to adapt to the effects of climate change*.

[ITU-T L.1503] Recommendation ITU-T L.1503 (201), *Use of information and communication technology for climate change adaptation in cities*.

[ITU-T M.3363] Recommendation ITU-T M.3363 (2020), *Requirements for data management in the telecommunication management network*.

[ITU-T X.800] Recommendation ITU-T X.800 (1991), *Security architecture for Open Systems Interconnection for CCITT applications*.

[ITU-T X.1217] Recommendation ITU-T X.1217 (2021), *Guidelines for applying threat intelligence in telecommunication network operation*.

[ITU-T X.1235] Recommendation ITU-T X.1235 (2022), *Technologies in countering website spoofing for telecommunication organizations*.

[ITU-T X.1303] Recommendation ITU-T X.1303 (2007), *Common alerting protocol (CAP 1.1)*.

[ITU-T X.1451] Recommendation ITU-T X.1451 (2020), *Risk identification to optimize authentication*.

[ITU-T Y.101] Recommendation ITU-T Y.101 (2000), *Global Information Infrastructure terminology: Terms and definitions*.

[ITU-T Y.2071] Recommendation ITU-T Y.2071 (2015), *Framework of a micro energy grid*.

[ITU-T Y.2233] Recommendation ITU-T Y.2233 (2010), *Requirements and framework allowing accounting and charging capabilities in NGN*.

[ITU-T Y.3172] Recommendation ITU-T Y.3172 (2019), *Architectural framework for machine learning in future networks including IMT-2020*.

[ITU-T Y.3174] Recommendation ITU-T Y.3174 (2020), *Framework for data handling to enable machine learning in future networks including IMT-2020*.

[ITU-T Y.3179] Recommendation ITU-T Y.3179 (2021), *Architectural framework for machine learning model serving in future networks including IMT-2020*.

[ITU-T Y.3500] Recommendation ITU-T Y.3500 (2014), *Information technology - Cloud computing - Overview and vocabulary*.

[ITU-T Y.3514] Recommendation ITU-T Y.3514 (2017), *Cloud computing – Trusted inter-cloud computing framework and requirements*.

[ITU-T Y.3519] Recommendation ITU-T Y.3519 (2018), *Cloud computing – Functional architecture of big data as a service*.

[ITU-T Y.3602] Recommendation ITU-T Y.3602 (2022), *Big data – Functional requirements for data provenance*.

[ITU-T Y.4000] Recommendation ITU-T Y.4000/Y.2060 (2012), *Overview of the Internet of things*.

[ITU-T Y.4051] Recommendation ITU-T Y.4051 (2019), *Vocabulary for smart cities and communities*.

[ITU-T Y.4113] Recommendation ITU-T Y.4113 (2016), *Requirements of the network for the Internet of things*.

[ITU-T Y.4205] Recommendation ITU-T Y.4205 (2019), *Requirements and reference model of IoT-related crowdsourced systems*.

[ITU-T Y.4472] Recommendation ITU-T Y.4472 (2020), *Open data application programming interfaces (APIs) for IoT data in smart cities and communities*.

[ITU-R RS.577-7] Recommendation ITU-R RS.577-7 (2009), *Frequency bands and required bandwidths used for spaceborne active sensors operating in the Earth exploration-satellite (active) and space research (active) services*.

[ITU-R RS.1632] Recommendation ITU-R RS.1632 (2003), *Sharing in the band 5 250-5 350 MHz between the Earth exploration-satellite service (active) and wireless access systems (including radio local area networks) in the mobile service*.

[ITU-R TF.686-2] Recommendation ITU-R TF.686-2 (2002), *Glossary and definitions of time and frequency terms*.

[ITU-R V.573-4] Recommendation ITU-R V.573-4 (2000), *Radiocommunication vocabulary*.

3 Introduction

The efforts for this glossary can be summarized as a process of identifying relevant terms and appropriate definitions that best support and represent the ITU/WMO/UNEP Focus Group on Artificial Intelligence for Natural Disaster Management (FG-AI4NDM). The first step of this process is to screen all FG-AI4NDM input documents for terms of interest for the various disciplines that are involved in FG-AI4NDM.

The terms of interest are stored in alphabetical order and research of the definitions was stored in the online database "Prototype Integrated Database ITU Terms and Definitions," followed and aligned with the standards by the subcommittee on "Artificial Intelligence" ISO/IEC JTC 1/SC 42[1]. Further sources were sought if neither of these two sources provided definitions. In the case of existing definitions that are too general or less related to the context of AI and natural disasters, appropriate definitions were sought and added in the frame of the AI and statistics (machine learning is included as a subdomain of AI) and geoscience disciplines.

These are shown in the document as [G] for general definition, [AI/ST] for AI and statistics derived definitions, and [GEO] for geoscience related definitions.

A definition for a term of interest was found to be adequate, if it fulfils one of the following criteria:

1. The definition falls within the AI/statistics or geoscience disciplines and stems from the ITU database;
2. The definition is based on and/or aligned with the standards of the AI subcommittee ISO/IEC JTC 1/SC 42 and belongs to the AI/statistics or geoscience fields;
3. The general definition derives from the ITU database – the AI/statistics or geoscience definition are further researched and added;
4. The definition falls within the scope of the AI/statistics or geoscience disciplines and comes from a verified source;
5. The general definition derives from a verified source – the AI/statistics or geoscience definition are further researched and added.

[1] https://www.iso.org/committee/6794475.html

In some cases, multiple definitions are provided/found. These definitions are merged to one single definition.

The updated version of this glossary takes into consideration the guidelines of the Standardization Committee for Vocabulary (SCV, SCV-TD156).

This Glossary serves as the basis for the further examination and recollection of relevant definitions to cater to the evolving sphere relating to the application of AI for natural disaster management.

4 Definitions

Accelerometer

[G]: An instrument that measures acceleration (change in velocity per unit time). There are two general types of accelerometers. One measures translational accelerations (changes in linear motions in one or more dimensions), and the other measures angular accelerations (changes in rotation rate per unit time). [b-ESA]

Accessibility

[G]: The degree to which a product, device, service or environment (virtual or real) is available to as many people as possible. [ITU-T F.791]

Accuracy:

[G]: Closeness of the agreement between the result of a measurement and a true value of the measurand. Accuracy is generally characterized by the overall uncertainty of a measured value. See also "uncertainty." [ITU-R TF.686-2]

[AI/ST]: The fraction of predictions that a classification model got right. [b-MachineLearning]

Active learning

[AI/ST]: A training approach in which the algorithm chooses some of the data from which it learns. Active learning is particularly valuable when labeled examples are scarce or expensive to obtain. Instead of blindly seeking a diverse range of labeled examples, an active learning algorithm selectively seeks the particular range of examples it needs for learning. [b-Misra]

Adaboost

[AI/ST]: AdaBoost is an ensemble method that trains and deploys trees in series. [b-Salam]

Adaptive neuro-fuzzy inference system (ANFIS)

[AI/ST]: An artificial neural network based on Takagi–Sugeno fuzzy inference system. Its inference system corresponds to a set of fuzzy IF–THEN rules that have learning capability to approximate nonlinear functions. Hence, ANFIS is considered to be a universal estimator. [b-Cyber]

Aggregation

[G]: Data aggregation is the process where raw data are gathered and expressed in a summary form for statistical analysis. [b-IBM]

AI ethics

[AI/ST]: Branch of applied ethics that focuses on the normative issues raised by the design, development, implementation and use of AI. [b-EC]

Alert system

[G]: A system, within which the NG-STB (see "NG-STB") participates, that allows a service provider to distribute public emergency alarms and information about the public emergency to all of the customers attached to the network. [ITU-T J.193]

Analytical hierarchy process (AHP)

[G]: The analytical hierarchy process, also analytic hierarchy process, is a structured technique for organizing and analyzing complex decisions, based on mathematics and psychology. [b-Forman]

Ancillary data

[AI/ST]: The data other than instrument data required to perform an instrument's data processing. [b-NASA]

Application programming interface (API)

[G]: An API provides a set of interfaces from an application environment to an execution environment. The execution environment provides services to the application environment. [ITU-T I.312]

[AI/ST]: A generic set of application programming interfaces for a machine learning data model (DM) to be used in a machine learning overlay and stored in a machine learning metadata store. [ITU-T Y.3174]

Area under curve (AUC)

[AI/ST]: The area under a curve between two points is calculated by performing the definite integral. In the context of a receiver operating characteristic for a binary classifier, the AUC represents the classifier's accuracy. [b-Springer-1]

Artificial intelligence (AI)

[G]: An interdisciplinary field, usually regarded as a branch of computer science, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning. A computerized system that uses cognition to understand information and solve problems. [b-ISO/IEC 2382-28]

Association

[G]: Association learning is a rule-based machine learning and data mining technique that finds important relations between variables or features in a dataset. Unlike conventional association algorithms measuring degrees of similarity, association rule learning identifies hidden correlations in databases by applying some measure of interestingness to generate an association rule for new searches. [b-DeepAI]

[AI/ST]: Association rule learning is a type of unsupervised learning technique that checks for the dependency of one data item on another data item and maps accordingly so that it can be more profitable. It tries to find some interesting relations or associations among the variables of dataset. It is based on different rules to discover the interesting relations between variables in the database. [b-Java]

Atmospheric reanalysis

[GEO]: Reanalyses are estimates of historical atmospheric temperature and wind or oceanographic temperature and current, and other quantities, created by processing past meteorological or oceanographic data using fixed state-of-the-art weather forecasting or ocean circulation models with

data assimilation techniques. Using fixed data assimilation avoids effects from the changing analysis system that occur in operational analyses. Although continuity is improved, global reanalyses still suffer from changing coverage and biases in the observing systems. [b-ECMWF]

Authenticity

[G]: Property that an entity is what it claims to be. Authenticity is judged on the basis of evidence. [b-ISO/IEC TS 22424-1]

Autocorrelation

[AI/ST]: The correlation between paired values of a function of a mathematical or statistical variable taken at usually constant intervals that indicates the degree of periodicity of the function. [b-Andrews]

Axitra

[GEO]: This is an algorithm that calculates dynamic stress variations resulting from a shear point source in a plane-layered elastic medium. [b-Geo]

Bagging (bootstrap aggregations)

[AI/ST]: A method to train an ensemble where each constituent model trains on a random subset of training examples sampled with replacement. For example, a random forest is a collection of decision trees trained with bagging. The term bagging is short for bootstrap aggregating. [b-MachineLearning]

Bandpass filtering

[G]: A bandpass filter is an electronic circuit or device that allows only signals between specific frequencies to pass through and attenuates/rejects frequencies outside the range. [b-Techno]

Bayesian optimization

[AI/ST]: Bayesian optimization is a sequential design strategy for global optimization of black-box functions that does not assume any functional forms. It is usually employed to optimize expensive-to-evaluate functions. [b-Mocus]

Bayesian Evidence-based Fault Orientation and Real-time Earthquake Slip (BEFORES)

[GEO]: An algorithm that uses real-time high-rate GPS data to simultaneously solve for a distributed slip model and fault geometry in real time as a rupture unfolds. [b-JGR]

Believability

[G]: Data believability is a characteristic of the quality of data. There is no clear measure to assess data believability. This characteristic is associated with other characteristics of data such as trustworthiness, reasonableness, and temporality. [b-MIT]

Benchmark dataset

[G]: Benchmark data are data that have been studied and annotated appropriately. They are used as a point of reference to compare and evaluate tools, methods, and algorithms. [b-Agile]

Bidirectional encoder representation transformer (BERT)

[AI/ST]: A model architecture for text representation. A trained BERT model can act as part of a larger model for text classification or other ML tasks. BERT uses the Transformer architecture, and therefore relies on self-attention. It uses the encoder part of the Transformer whose job is to produce

good text representations, rather than to perform a specific task like classification. It is bidirectional and uses masking for unsupervised training. [b-MachineLearning]

Bias

[AI/ST]: Systematic error introduced into sampling or testing by selecting or encouraging one outcome or answer over others. Statistical bias refers to any type of error or distortion that is found with the use of statistical analyses. [b-Piedmont]

[GEO]: An intercept or offset from an origin. Bias (also known as the bias term) is referred to as b or w0 in machine learning models. [b-MachineLearning]

Biased data

[AI/ST]: Data that are incomplete, too heavily weighted toward specific attributes, and/or unrepresentative of all use cases. Biased data can lead to biased models, as a model is only as good as the data used for training. [b-Forbes]

Big data

[G]: Extensive datasets that require a scalable technology for efficient storage, manipulation, management, and analysis. [b-ISO/IEC 20546]

[AI/ST]: Big data is a term that refers to datasets that have been created on a scale that is difficult to obtain, manage, and manipulate in a timely manner using traditional computing tools [1], but are instead handled with a "scalable architecture for efficient storage, manipulation, and analysis" [b-Nisbet], such as that provided by ML and DL. [b-NIST]

Big data analysis

[G]: Big data analytics is the often complex process of examining big data to uncover information – such as hidden patterns, correlations, market trends and customer preferences – that can help organizations make informed business decisions. [b-Techtarget]

Binary classification model

[AI/ST]: A type of classification task that outputs one of two mutually exclusive classes. For example, a machine learning model that evaluates email messages and outputs either "spam" or "not spam" is a binary classifier. [b-MachineLearning]

Binning

[AI/ST]: Converting a (usually continuous) feature into multiple binary features called buckets or bins, typically based on value range. For example, instead of representing temperature as a single continuous floating-point feature, you could chop ranges of temperatures into discrete bins. Given temperature data sensitive to a tenth of a degree, all temperatures between 0.0 and 15.0 degrees could be put into one bin, 15.1 to 30.0 degrees could be a second bin, and 30.1 to 50.0 degrees could be a third bin. [b-MachineLearning]

Black-box approach

[AI/ST]: In deep learning, models can be very complex. If the models reach a certain level of complexity then it can be very difficult to understand how the model works. Black-box approaches refer to approaches that do not try to extract understanding from the model because of its complexity, but rather assess and evaluate the model solely on its input and output. [b-HDSR]

Block kriging

[GEO]: All the kriging techniques aim at predicting the value of a variable at specific unsampled locations. These locations can be considered as spatial points (or more precisely as pixels in the grid of interpolation). As a consequence, such kriging approaches are also referred to as point kriging methods. When the uncertainty is relatively large, one might want to smooth the interpolated results by performing kriging on a larger area than single pixels. This type of kriging interpolation is known as block kriging. This has the advantage of lowering prediction errors over the map. [b-Matheron] [b-aspexit]

Blocked approach

[AI/ST]: In the statistical theory of the design of experiments, blocking is the arranging of experimental units in groups (blocks) that are similar to one another. Blocking can be used to tackle the problem of pseudoreplication. [b-Addelman]

[GEO]: See "Block kriging."

Boosting

[AI/ST]: A machine learning technique that iteratively combines a set of simple and not very accurate classifiers (referred to as "weak" classifiers) into a classifier with high accuracy (a "strong" classifier) by upweighting the examples that the model is currently misclassifying. [b-MachineLearning]

Classification and regression trees (CART)

[AI/ST]: Classification and regression trees are a set of techniques for classification and prediction. The technique is aimed at producing rules that predict the value of an outcome (target) variable from known values of predictor (explanatory) variables. The predictor variables may be a mixture of categorical and continuous variables. [b-Statistics]

Causality

[G]: Causality (also referred to as causation, or cause and effect) occurs when one event, process, state, or object (a cause) contributes to the production of another event, process, state, or object (an effect), where the cause is partly responsible for the effect, and the effect is partly dependent on the cause. In general, a process has many causes, which are also said to be causal factors for it, and all lie in its past. [b-Bunge]

Chatbots

[AI/ST]: Chatbots consist of computer programs intended to simulate dialogues with human users. [b-Adamopoulou]

Chi-squared test

[AI/ST]: A test derived from the chi-square distribution to compare the goodness of fit of theoretical and observed frequency distributions or to compare nominal data derived from unmatched groups of subjects. [b-Collins]

Class compactness

[AI/ST]: Class compactness refers to how distinct one class (or cluster) is from another in the data. Compactness evaluates how similar the instances of a class are. One example of a metric that does that is the deviation from the mean. [b-Cornell]

Class separability

[AI/ST]: Class separability refers to how distinct one class (or cluster) is from another in the data. Separability evaluates how dissimilar two different classes are. An example can be the distance between the means of the two classes. [b-Springer-2]

Classification

[G]: The act or process of dividing things into groups according to their type. [b-Cambridge]

[AI/ST]: Classification is the process of predicting the class of given data points. Classes are sometimes called as targets/ labels or categories. Classification predictive modeling is the task of approximating a mapping function (*f*) from input variables (*x*) to discrete output variables (*y*). [b-Data]

Classification tree

[AI/ST]: Decision tree learning or induction of decision trees is one of the predictive modelling approaches used in statistics, data mining, and machine learning. It uses a decision tree (as a predictive model) to go from observations about an item (represented in the branches) to conclusions about the item's target value (represented in the leaves). Tree models where the target variable can take a discrete set of values are called classification trees; in these tree structures, leaves represent class labels and branches represent conjunctions of features that lead to those class labels. [b-Shalev] [b-Rokach]

Classifier

[G]: 1. Mechanism that describes behavioral and structural features. 2. Device that separates particles, according to their size, shape and density, by physical means other than screening. [b-ISO 16100-2] [b-ISO 1213-1]

[AI/ST]: ML model used for classification. [b-ISO/IEC TR 29119-11]

Climate change

[G]: Any change in climate over time, whether due to natural variability or as a result of human activity. The Intergovernmental Panel on Climate Change (IPCC) uses a relatively broad definition, referring to a change in the state of the climate that can be identified (e.g., using statistical tests) by changes in the mean and/or the variability of its properties and that persists for an extended period, typically decades or longer. Climate change may be due to natural internal processes or external forcing, or to persistent anthropogenic changes in the composition of the atmosphere or in land use. [ITU-T L.1501]

Climatology

[GEO]: Set of reference simulations, including historical simulations forced with observations, reanalysis or forecasts re-runs. [b-ECMWFTerm]

Cloud computing

[G]: Paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand. Examples of resources include servers, operating systems, networks, software, applications, and storage equipment. [ITU-T Y.3500]

Clustering

[G]: The process of partitioning a set of patterns into disjoint and homogeneous meaningful groups. [ITU-T E.475]

[AI/ST]: Grouping by an algorithm of events, people or things based on their proximity to each other, as reflected in the data gathered on them. [b-Remi]

Coarse to fine

[AI/ST]: In parameter optimization coarse to fine refers to the step wise refinement of the parameters. Initially, in the first step, few values are tested for the parameter, which are far apart. When the best performing of these values is identified, in the next step, a more refined test can be made for many new values but over a much smaller range around the previously best result. This can continue for several steps until the desired precision is achieved. [b-DataScience]

Coefficient of determination

[AI/ST]: Coefficient of determination, denoted R^2 or r^2 and pronounced "R squared," is the proportion of the variation in the dependent variable that is predictable from the independent variable(s). [b-Glantz] [b-Draper] [b-Wright]

Common Alerting Protocol (CAP)

[G]: The Common Alerting Protocol is an XML-based data format for exchanging public warnings and emergencies between alerting technologies. CAP allows a warning message to be consistently disseminated simultaneously over many warning systems to many applications, such as Google Public Alerts and Cell Broadcast. CAP increases warning effectiveness and simplifies the task of activating a warning for responsible officials. [OASIS] [ITU-T X.1303]

Complexity

[G]: The degree of complication of a system or system component, determined by such factors as the number and intricacy of interfaces, the number and intricacy of conditional branches, the degree of nesting, the types of data structures, and other system characteristics. [b-EUTerm]

Computer vision

[G]: Computer vision is an interdisciplinary field that uses computer science techniques to analyze and understand digital images and videos. Computer vision tasks include object recognition, event detection, motion detection, and object tracking, among others. [b-Springer-1]

Conditional inference tree

[AI/ST]: Conditional inference trees is a non-parametric class of decision trees and is also known as unbiased recursive partitioning. It is a recursive partitioning approach for continuous and multivariate response variables in a conditional inference framework. [b-Bath]

Confusion matrix

[AI/ST]: A confusion matrix summarizes the classification performance of a classifier with respect to some test data. It is a two-dimensional matrix, indexed in one dimension by the true class of an object and in the other by the class that the classifier assigns. [b-Ting]

Continuous ranked probability score (CRPS)

[AI/ST]: The continuous ranked probability score is a much-used measure of performance for probabilistic forecasts of a scalar observation. It is a quadratic measure of the difference between the forecast cumulative distribution function (CDF) and the empirical CDF of the observation. [b-Zamo]

Convection

[GEO]: In meteorology, the term is used specifically to describe vertical transport of heat and moisture in the atmosphere, especially by updrafts and downdrafts in an unstable atmosphere. The terms "convection" and "thunderstorms" often are used interchangeably, although thunderstorms are only one form of convection. Cbs, towering cumulus clouds, and ACCAS clouds all are visible forms of convection. However, convection is not always made visible by clouds. Convection which occurs without cloud formation is called dry convection, while the visible convection processes referred to above are forms of moist convection. [b-NOAA-1]

Convective events

[GEO]: Atmospheric convection is the result of a parcel-environment instability, or temperature difference layer in the atmosphere. Different lapse rates within dry and moist air masses lead to instability. Mixing of air during the day which expands the height of the planetary boundary layer leads to increased winds, cumulus cloud development, and decreased surface dew points. Moist convection leads to thunderstorm development, which is often responsible for severe weather throughout the world. Special threats from thunderstorms include hail, downbursts, and tornadoes. [b-NWS] [b-Hahn]

Convolutional encoder-decoder (CED)

[AI/ST]: The CED architecture is composed of an encoder branch containing convolutional layers that feeds to a latent space (also known as hidden state), followed by a decoder branch to construct the output variable. [b-Encoder]

Convolutional layer

[AI/ST]: A layer of a deep neural network in which a convolutional filter passes along an input matrix. [b-MachineLearning]

Convolutional neural networks (CNN)

[AI/ST]: A neural network in which at least one layer is a convolutional layer. [b-MachineLearning]

Convolutional transpose layer

[AI/ST]: The convolutional transpose layer is a convolutional layer used in artificial neural networks. Contrary to the typical convolutional layer, it is used to generate an output feature map that has a spatial dimension greater than that of the input feature map, and is thus used in upsampling. [b-Convolutional]

Copernicus Climate Data Store

[GEO]: Copernicus Climate Data Store is a database of climate and weather data. [b-ClimateData]

Copernicus Open Access Hub

[GEO]: The Copernicus Open Access Hub (previously known as Sentinels Scientific Data Hub) provides complete, free and open access to Sentinel-1, Sentinel-2, Sentinel-3, and Sentinel-5P user products, starting from the In-Orbit Commissioning Review (IOCR). Sentinel data are also available

via the Copernicus Data and Information Access Services (DIAS) through several platforms. [b-Copernicus]

Corpus creation

[AI/ST]: A corpus can be defined as a collection of machine-readable authentic texts (including transcripts of spoken data) that is sampled to be representative of a particular natural language or language variety. Corpora play an essential role in natural language processing (NLP) research as well as a wide range of linguistic investigations. They provide a material basis and a test bed for building NLP systems. [b-Xiao]

Correlation

[G]: Capability to generate an aggregated charging information record (CIR) by combining and analyzing charging events collected from the same transport/service session. [ITU-T Y.2233]

Cross training

[AI/ST]: Cross training refers to the training of different parts of a model with different data. [b-Medium]

Crowdsourced data

[G]: A method to gather active or passive quality of service measurements from a large number of end user devices. [ITU-T E.806]

Crowdsourcing

[G]: The practice of obtaining needed services, ideas, content or other system resources by soliciting contributions from a large, open and potentially undefined group of people, rather than from employees, suppliers or identified experts through an online open call by providing incentives (financial, social, or entertainment) to all or a subset of those crowd members who participate in the crowdsourcing activity. [ITU-T Y.4205]

Cyclones

[GEO]: In meteorology, a cyclone is a large air mass that rotates around a strong center of low atmospheric pressure, counterclockwise in the Northern Hemisphere and clockwise in the Southern Hemisphere as viewed from above (opposite to an anticyclone). [b-AMS]

D-SUMM

[AI/ST]: An information summarizer, which uses the "BERT" natural language processing model to derive situational awareness for a specified area. [b-Khurana]

Darknet framework

[AI/ST]: Darknet is an open source neural network framework written in the C and CUDA languages. [b-Github]

Dashboard

[G]: A dashboard is a type of graphical user interface which often provides at-a-glance views and summaries of different types of information. [b-Cambridge]

Data accessibility

[G]: Possibility to request and obtain the data at any time in a machine-readable format. [b-EUParliament]

Data annotation

[AI/ST]: Data annotation is the categorization and labeling of data to create training data for AI models. [b-Appen]

Data bias

[AI/ST]: Data bias in AI systems is an error that occurs when some elements/factors are more represented and/or considered than others. [b- Hellström]

Data cascades

[AI/ST]: Compounding events causing negative, downstream effects from data issues, resulting in technical debt over time. Data cascades often result from applying conventional AI practices that undervalued data quality. [b-Sambasivan]

Data cleaning

[G]: A process to delete irrelevant data and duplicate data in the original dataset, to smooth the noise data, and process missing values and outliers. [ITU-T X.1217]

Data completeness

[G]: Data completeness refers to the comprehensiveness or wholeness of the data. [b-BDEX]

Data completion

[G]: Data completion refers to techniques that can be used to deal with missing data. For example, if a value is missing for an instance, that value can be replaces with the weighted average of the respective values of the closest neighbors. [b-Laboratory]

Data compression

[G]: In information theory, data compression, source coding, or bit-rate reduction is the process of encoding information using fewer bits than the original representation. Any particular compression is either lossy or lossless. Lossless compression reduces bits by identifying and eliminating statistical redundancy. No information is lost in lossless compression. Lossy compression reduces bits by removing unnecessary or less important information. [b-Wade]

Data consistency

[G]: Data consistency refers to the characteristic of data to be the same even when collected by different agents, or stored in different locations. [b-BigData]

Data cubes

[GEO]: A data cube, also known as a coverage is the digital representation of some spatiotemporal phenomenon. A data cube acts as a function to return values from its range for any direct position within its spatial, temporal or spatiotemporal domain. Data cubes play an important role in geographic information systems (GIS), geospatial content and services, GIS data processing, and data sharing. [b-GIGAS]

Data curation

[G]: The active and ongoing management of data through its lifecycle of interest and usefulness. [b-AcademicPress]

Data custodianship

[G]: Data custodianship mandates responsibilities for the acquisition, management, maintenance, and quality of information. [b-UNGGIM]

Data deletion

[G]: The deletion of data entries or instances that are incomplete and unusable. [b-PersonalData]

Data distillation

[AI/ST]: For a machine learning algorithm to be effective, useful features must be extracted from (often) large amounts of training data. However, this process can be made challenging due to the costs associated with training on such large datasets, both in terms of compute requirements and wall clock time. The idea of distillation plays an important role in these situations by reducing the resources required for the model to be effective. The most widely known form of distillation is model distillation (a.k.a. knowledge distillation), where the predictions of large, complex teacher models are distilled into smaller models. [b-aigoogle]

Data enrichment

[AI/ST]: The process of determining which features might be useful in training a model, and then converting raw data from log files and other sources into said features. [b-Klein]

Data fusion

[G]: Data fusion is the process of integrating multiple data sources to produce more consistent, accurate, and useful information than that provided by any individual data source. [b-Springer-1]

Data integration

[G]: Data integration involves the combination of data residing in different resources and then the supply in a unified view to the users. Data integration is in high demand for both commercial and scientific domains in which they need to merge the data and research results from different repositories. [b-TIBCO]

Data integrity

[G]: The property that the data has not been altered or destroyed in an unauthorized manner. [ITU-T X.800]

Data interoperability

[G]: The ability of two or more systems or components to exchange data and to use the data that have been exchanged. [b-Data4SDGs]

Data leakage

[AI/ST]: Leakage (also known as data leakage or target leakage) is the use of information in the model training process that would not be expected to be available at prediction time, causing the predictive scores (metrics) to overestimate the model's utility when run in a production environment. [b-MIT] [b-Shachar]

Data lifecycle

[G]: A whole range of data processing phases including data planning, data acquisition, data storage, data sharing, data usage, data transmission, and data disposal. [ITU-T M.3363]

Data mining

[G]: Computational process that extracts patterns by analyzing quantitative data from different perspectives and dimensions, categorizing them, and summarizing potential relationships and impacts. [b-ISO/IEC 22989]

Data normalization

[AI/ST]: In statistics and applications of statistics, normalization can have a range of meanings. In the simplest cases, normalization of ratings means adjusting values measured on different scales to a notionally common scale, often prior to averaging. [b-Freedman]

Data ownership

[G]: Legal rights and complete control over a single piece or set of data elements governing the acquisition, use and distribution of data assets. [b-Dodge]

Data protection laws

[G]: Rights-based approaches that provide standards for regulating data processing that both protect the rights of individuals and establish obligations for data controllers and processors. [b-London] [b-WHO]

Data provenance

[G]: Information on the source of data, such as the person responsible for the provision of data, functions applied to data, and information about the computing environment for data processing (e.g., operating system, description of the hardware, locale settings and time zone). Big data provenance is the information that records the historical path of data according to the data lifecycle operations in a big data ecosystem. [ITU-T Y.3602]

Data quality

[G]: Data quality refers to the state of qualitative or quantitative pieces of information. There are many definitions of data quality, but data are generally considered high quality if they are "fit for […] intended uses in operations, decision making and planning." Moreover, data are deemed of high quality if they correctly represent the real-world construct to which they refer. [b-Redman]

Data sparsity

[G]: The number of elements set to zero (or null) in a vector or matrix divided by the total number of entries in that vector or matrix. Feature sparsity refers to the sparsity of a feature vector; model sparsity refers to the sparsity of the model weights. [b-MLMastery-1]

Data split

[AI/ST]: Data split refers to the splitting of the instances of the data to groups. It is often done in machine learning to separate the training set, the validation set, and the test set. [b-MLMastery-2]

Data supply chains

[G]: The data supply chain represents the technological steps and human-involved processes supporting the flow of data through the organization, from its raw state, through transformation and integration, all the way through to the point of consumption or analysis. [b-Immuta]

[AI/ST]: The lifecycle process of data; selection, procurement, transfer, quality assurance, warehousing/storage, data management, transformation, monitoring, and distribution – feeding data pipelines for use in data products. [b-DataSupply]

Data uncertainties

[G]: Uncertainty is a situation in which something is not known, or something that is not known or certain. [b-CambridgeDictionary-1]

[AI/ST]: In computer science, uncertain data are data that contain noise that makes them deviate from the correct, intended or original values. In the age of big data, uncertainty or data veracity is one of the defining characteristics of data. [b-Hariri] [b-IBM]

Data validation

[G]: Data validation means checking the accuracy and quality of source data before using, importing or otherwise processing data. [b-Informatica]

[AI/ST]: In computer science, data validation is the process of ensuring data have undergone data cleansing to ensure they have data quality, that is, that they are both correct and useful. It uses routines, often called "validation rule," "validation constraint," or "check routines," which check for correctness, meaningfulness, and security of data that are input to the system. [b-Essnet]

Data visibility

[G]: Data visibility is the degree of ease through which an entity can monitor, search, display, and analyze data from disparate sources. [b-Sentinal]

Decision support system (DSS)

[G]: A decision support system is an information system that supports business or organizational decision-making activities. [b-Sprague]

Decision tree

[AI/ST]: A decision tree uses a tree-like graph or model as a structure to perform decision analysis. It uses each node to represent a test on an attribute, each branch to represent the outcome of the test, and each leaf node to represent a class label. [b-Springer-1]

Deep belief network (DBN)

[AI/ST]: In machine learning, a deep belief network is a generative graphical model, or alternatively a class of deep neural network, composed of multiple layers of latent variables ("hidden units"), with connections between the layers but not between units within each layer. [b-Hinton]

Deep convection

[GEO]: Deep convection refers to the thermally driven turbulent mixing that moves air parcels from the lower to the upper atmosphere. In the tropics, this generally involves the vertical ascent of warm moist air and, ultimately, precipitation. [b-Columbia]

Deep learning

[AI/ST]: Approach to creating rich hierarchical representations through the training of neural networks with one or more hidden layers. [b-ISO/IEC TR 29119-11]

Deep learning convolutional encoder-decoder

[AI/ST]: See "Convolutional encoder-decoder (CED)."

Denoising

[G]: See "Noise reduction."

[AI/ST]: A common approach to self-supervised learning in which noise is artificially added to the dataset. The model tries to remove the noise. Denoising enables learning from unlabeled examples. The original dataset serves as the target or label and the noisy data as the input. Some masked language models use denoising as follows: noise is artificially added to an unlabeled sentence by masking some of the tokens. The model tries to predict the original tokens. [b-CornellUniversity]

Desert locust

[G]: The desert locust is a species of locust, a periodically swarming, short-horned grasshopper in the family Acrididae. They are found primarily in the deserts and dry areas of northern and eastern Africa, Arabia, and southwest Asia. [b-Draper]

Die-hard network

[G]: Die-hard network is a network system including server function or its concept which is hard to break and toughly survivable with various communication methods. [b-Sprague]

Differential privacy (DP)

[AI/ST]: Differential privacy is a system for publicly sharing information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals in the dataset. The idea behind differential privacy is that if the effect of making an arbitrary single substitution in the database is small enough, the query result cannot be used to infer much about any single individual, and therefore provides privacy. [b- Dalenius]

Digital elevation model (DEM)

[AI/ST]: A digital elevation model is a 3D computer graphics representation of elevation data to represent terrain, commonly of a planet, moon, or asteroid. A "global DEM" refers to a discrete global grid. DEMs are used often in geographic information systems and are the most common basis for digitally produced relief maps. [b-CRC]

[GEO]: DEM refers to a surface in general (e.g., atmospheric layers, groundwater table, etc.) consisting of a dataset that can be used to complement remotely sensed images. It is a set of points defined in a three-dimensional cartesian space (*X*, *Y*, *Z*) that approximates a real surface like a digital counterpart to nature. *X*- and *Y*-axis may be expressed as geographic coordinates whereas the *Z*-axis usually represents the altitude above sea level. [b-UP42]

Digital terrain model (DTM)

[GEO]: A digital terrain model refers to the land surface without considering the persistent objects on the ground (vegetation, buildings, and other artifacts). It can be considered as a synonym of bare-earth DEM and shows the development of the geodesic surface. [b-UP42]

Digital twin

[G]: A digital twin is a virtual model designed to accurately reflect a physical object. [b-IBM-2]

Dimensionality reduction

[AI/ST]: Decreasing the number of dimensions used to represent a particular feature in a feature vector. [b-Guide]

Disaster

[G]: A serious disruption of the functioning of a community or a society involving widespread human, material, economic or environmental losses and impacts, which exceeds the ability of the affected community or society to cope using its own resources. [b-UNDRR]

Disaster risk reduction (DRR)

[G]: Disaster risk reduction is the concept and practice of reducing disaster risks through systematic efforts to analyze and reduce the causal factors of disasters. [b-UNESCO-1]

Discretization

[G]: The transformation of a variable from continuous to discrete. [b-Arizona]

Data drift

[AI/ST]: Data drift is defined as a variation in the production data from the data that were used to test and validate the model before deploying it in production. [b-DataDrift]

[GEO]: Since the climate changes over time, when collecting climate data from a specific area, the data change over time. This is called data drift. When there is data drift and the data are used to create a model, there can be model drift since the model will over time be less effective in newer data. [b-Datatron]

Model drift

[AI/ST]: Model drift (also known as model decay) refers to the degradation of a model's prediction power due to changes in the environment, and thus the relationships between variables. [b-Datatron]

[GEO]: Since model climate differs to some extent from observed climate, climate forecasts will typically 'drift' from the initial observation-based state towards the model's climate. This drift occurs at different time scales for different variables, can obscure the initial-condition forecast information and is usually removed a posteriori by an empirical, usually linear, adjustment. [b-NOAAStudent]

Drifters

[GEO]: A drifter (also known as a drifting buoy) is a floating ocean buoy equipped with meteorological and/or oceanographic sensing instruments linked to transmitting equipment for sending the observed data to collecting centers. [b-ScienceDirect-1]

Early stopping

[AI/ST]: A method for regularization that involves ending model training before training loss finishes decreasing. In early stopping, you end model training when the loss on a validation dataset starts to increase, that is, when generalization performance worsens. [b-MLMastery-3]

Earthflow

[GEO]: A sheet or stream of soil and rock material saturated with water and flowing downslope under the pull of gravity; it represents the intermediate stage between creep and mudflow. Earthflows usually begin in a large basin on the upper part of a slope where debris and weathered material accumulate; the movement, usually set off by heavy rainfall, may be relatively slow or very fast, depending on the amount of water present, the angle of the slope, and other aspects of the terrain. [b-Britannica]

Earthquake

[GEO]: A sudden violent movement of the Earth's surface, sometimes causing great damage. [b-Cambridge]

Earthquake cycle

[GEO]: The earthquake cycle refers to the phenomenon that earthquakes repeatedly occur on the same fault as the result of continual stress accumulation and periodic stress release. Earthquake cycles can occur on a variety of faults including subduction zones and continental faults. Depending on the size of the earthquake, an earthquake cycle can last decades, centuries, or longer. [b-Scholz]

Ecological niche model (ENM)

[GEO]: Species distribution modelling (SDM), also known as environmental (or ecological) niche modelling, habitat modelling, predictive habitat distribution modelling, and range mapping uses computer algorithms to predict the distribution of a species across geographic space and time using environmental data. The environmental data are most often climate data (e.g., temperature, precipitation), but can include other variables such as soil type, water depth, and land cover. SDMs are used in several research areas in conservation biology, ecology, and evolution. These models can be used to understand how environmental conditions influence the occurrence or abundance of a species, and for predictive purposes (ecological forecasting). Predictions from an SDM may be of a species' future distribution under climate change, a species' past distribution in order to assess evolutionary relationships, or the potential future distribution of an invasive species. Predictions of current and/or future habitat suitability can be useful for management applications (e.g., reintroduction or translocation of vulnerable species, reserve placement in anticipation of climate change). [b-Elith]

El Nino-Southern Oscillation (ENSO)

[GEO]: The term El Niño was initially used to describe a warm-water current that periodically flows along the coast of Ecuador and Peru, disrupting the local fishery. It has since become identified with a basin-wide warming of the tropical Pacific Ocean east of the dateline. This oceanic event is associated with a fluctuation of a global-scale tropical and subtropical surface pressure pattern called the Southern Oscillation. This coupled atmosphere-ocean phenomenon, with preferred time scales of two to about seven years, is known as the El Niño-Southern Oscillation. It is often measured by the surface pressure anomaly difference between Tahiti and Darwin or the sea surface temperatures in the central and eastern equatorial Pacific. During an ENSO event, the prevailing trade winds weaken, reducing upwelling and altering ocean currents such that the sea surface temperatures warm, further weakening the trade winds. This event has a great impact on the wind, sea surface temperature, and precipitation patterns in the tropical Pacific. It has climatic effects throughout the Pacific region and in many other parts of the world, through global teleconnections. The cold phase of ENSO is called La Niña. [b-IBE]

Encoder

[G]: An embodiment of an encoding process. [ITU-T H.222.0 (V8)]

[AI/ST]: In general, any ML system that converts from a raw, sparse, or external representation into a more processed, denser, or more internal representation. [b-MachineLearning]

Energy-constrained scenario

[GEO]: A constrained device can be a micro controller, smart device, sensor and actuator, or other type of small computers. In many applications, such a device can gather some information or act on the input of some information. It can be placed in factories, in special appliances, outdoors for collecting temperature data, etc. They can operate under extreme resource constraints-limited energy supply, small battery power, little memory space (flash and random-access memory), and small computing capabilities. Energy-constrained devices are specifically the devices with limited supply of energy. [b-netidee]

Ensemble

[AI/ST]: A collection of models trained independently whose predictions are averaged or aggregated. In many cases, an ensemble produces better predictions than a single model. For example, a random forest is an ensemble built from multiple decision trees. Note that not all decision forests are ensembles. [b-scikit]

Ensemble binary classification model

[AI/ST]: Binary classification is a type of classification task that outputs one of two mutually exclusive classes. An ensemble is a collection of models trained independently whose predictions are averaged or aggregated. In many cases, an ensemble produces better predictions than a single model. An ensemble binary classification model is an ensemble of many binary classification models. [b-ScienceDirect-2]

Epoch

[G]: A long period of time, especially one in which there are new developments and great change. [b-Cambridge]

[AI/ST]: In machine learning, one entire transit of the training data through the algorithm is known as an epoch. [b-Deepchecks]

[GEO]: Formal geochronologic unit, longer than an age and shorter than a period, during which the rocks of the corresponding series were formed. [b-AMS] [b-GeoScience]

European Space Agency (ESA) Data and Information Access Services (DIAS)

[G]: To facilitate and standardize access to data, the European Commission has funded the deployment of five cloud-based platforms. They provide centralized access to Copernicus data and information, as well as to processing tools. These platforms are known as the Data and Information Access Services. [b-Copernicus]

Estimator

[AI/ST]: A statistic intended to estimate a population parameter. [b-Copernicus]

Ethical AI

[AI/ST]: 1. Development, deployment, and use of AI that ensures compliance with ethical norms, including fundamental rights as special moral entitlements, ethical principles, and related core values. 2. "Ethical AI" is an umbrella term that houses a vast set of definitions such as "transparency," "non-maleficence," "responsibility," and "trust". 3. The ethics of artificial intelligence is the ethics of technology specific to robots and other artificial intelligence beings, which is divided into robot ethics and machine ethics. The former is about the concern with the moral behavior of humans as they design, construct, use, and treat artificially intelligent beings, and the latter is about the moral behavior of artificial moral agents (see also inadvertent effects). [b-Leslie] [b-Jobin]

Ethics

[G]: Branch of knowledge concerned with questions about right versus wrong conduct and what constitutes a good or bad life, as well as the justificatory basis for such questions.

[AI/ST]: See "Ethical AI." [b-Springer-1] [WHO GHE]

Evenness

[GEO]: Evenness is a measure of the relative abundance of pixels in the area selected and has a value between > 0 and 1. For a given number of classes (richness), an evenness = 1 is reached when all classes have the same area. Thus, higher values indicate that the number of pixels (area) is more evenly distributed between the land cover classes. Conversely, a value near 0 would indicate that nearly every pixel in the landscape is the same. [b-NASA]

Explainable AI (xAI)

[AI/ST]: Explainable artificial intelligence is a key term in AI design and in the tech community as a whole. It refers to efforts to make sure that artificial intelligence programs are transparent in their purposes and how they work. Explainable AI is a common goal and objective for engineers and others trying to move forward with artificial intelligence progress. [b-Springer-1]

Exploitation

[AI/ST]: See "Exploration-exploitation trade-off."

Exploration

[AI/ST]: See "Exploration-exploitation trade-off."

Exploration-exploitation trade-off

[AI/ST]: Reinforcement learning (RL) is an area of machine learning concerned with how intelligent agents ought to take actions in an environment in order to maximize the notion of cumulative reward. In each step, the intelligent agent can either try to "exploit" its current knowledge and maximize the reward of the next action, given the current knowledge, or, alternatively, perform an action in order to explore the action space and its reward "landscape," in order to receive better rewards in future steps. This is the trade-off between exploitation and exploration. [b-Reinforcement]

Exploratory data analysis (EDA)

[AI/ST]: Exploratory data analysis is an approach/philosophy for data analysis that employs a variety of techniques (graphical and quantitative) to better understand data. It can be seen as an initial examination of data to determine its salient characteristics and assess its quality. [b-DataAnalysis] [b-ISO/IEC 22989]

Exportability

[G]: To send or transport (a commodity, for example) abroad, especially for trade or sale. [b-FreeDictionary]

***F*-measure**

[AI/ST]: *F*-measure provides a single score that balances both the concerns of precision and recall in one number. The *F*-measure is defined as a harmonic mean of precision (*P*) and recall (*R*): $F = 2P*R/(P + R)$. [b-MLMastery-4]

***F*-score**

[AI/ST]: See "F-measure."

F1-score

[AI/ST]: See "F-measure."

Factorial analysis

[AI/ST]: Factorial analysis is a statistical method used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called factors. [b-Jöreskog]

False alarm rate (FAR)

[G]: The number of false alarms, in which an alarm or warning, is given in spite of a non-event, per total number of 'non-events' (times the event did not happen). A false alarm rate is also known as the probability of false detection. [b-Statistics]

False positive

[AI/ST]: An example in which the model mistakenly predicted the positive class. For example, the model inferred that a particular email message was spam (the positive class), but that email message was actually not spam. [b-Github]

False positive rate

[AI/ST]: Defined as: false positive rate = (false positives)/(false positives + true negatives). [b-MachineLearning]

Fault

[G]: Responsibility for wrongdoing or failure. [b-Fault]

[GEO]: A fracture in the Earth's crust along which there has been a displacement of strata. Faults are formed by compressional or tensional tectonic forces and can range from a fraction of a centimeter to hundreds of kilometers in size. [b-Zurich]

Fault slip

[GEO]: A fault's sense of slip is defined as the relative motion of the rock on each side of the fault with respect to the other side. [b-Fault]

Feature

[AI/ST]: An input variable used in making predictions. [b-Fault]

Feature extraction

[AI/ST]: The act of retrieving intermediate feature representations calculated by an unsupervised or pretrained model (for example, hidden layer values in a neural network) for use in another model as input. [b-MachineLearning]

Feature selection

[AI/ST]: Feature selection is the process of reducing the number of input variables when developing a predictive model. [b-MLMastery-5]

Federate testbeds

[G]: A federated testbed is a system that allows each individual testbed to maintain its own autonomy while working with others to share technological resources. [b-IEEE]

Federated learning

[AI/ST]: Federated learning (also known as collaborative learning) is a machine learning technique that trains an algorithm across multiple decentralized edge devices or servers holding local data samples, without exchanging them. [b-Konečný]

Feedback

[AI/ST]: In machine learning, a situation in which a model's predictions influence the training data for the same model or another model. For example, a model that recommends movies will influence the movies that people see, which will then influence subsequent movie recommendation models. [b-NOAAStudent]

Finite element method

[AI/ST]: Mathematical method for finding approximate solutions for partial differential equations or integral equations. [b-Reinhardt]

Finite-discrete element method (FDEM)

[AI/ST]: The combined finite-discrete element method is a recently developed numerical method aimed at modelling failing, fracturing, and/or fragmenting solids. The method combines aspects of both finite elements and discrete elements. [b-Munjiza]

Flash flood

[GEO]: Definition Flood of short duration with a relatively high peak discharge. [b-UNDRR]

Flatten layer

[AI/ST]: Flatten layer is used to make the multidimensional input one dimensional, commonly used in the transition from the convolution layer to the full connected layer. [b-Tensor-1]

Flood hazards

[GEO]: Combination of the probability of flooding and corresponding exposure characteristics such as flood depth, velocity, duration, rise rate, period of occurrence, and water quality. [b-ARCADIS]

Forecast failure method (FFM)

[G]: The forecast failure method for volcanic eruptions is a classical tool applied in the interpretation of monitoring data as potential precursors, providing quantitative predictions of the eruption onset. [b-Frontiers]

Frequency ratio

[AI/ST]: The frequency ratio is defined as the ratio of the frequency of the mistuned blisk sector to the frequency of the tuned blade sector. [b-Proceedings]

Fusion

[G]: Data fusion is the process of integrating information from multiple sources to produce specific, comprehensive, and unified data about an entity. [b-Cheng]

Fuzzy logic

[AI/ST]: A simple form for the many-valued logic, in which the truth values of variables may have any degree of "truthfulness" that can be represented by any real number in the range between 0 (as in completely false) and 1 (as in completely true) inclusive. Consequently, it is employed to handle the concept of partial truth, where the truth value may range between completely true and completely false. In contrast to Boolean logic, where the truth values of variables may have the integer values 0 or 1 only. [b- Chesalov]

Fuzzy rule

[AI/ST]: The fuzzy rule base is a set of if-then statements that store the practical knowledge of human operators about the process. [b-Molecular]

Fuzzy system

[AI/ST]: Fuzzy systems are structures based on fuzzy techniques oriented towards information processing, where the usage of classical sets theory and binary logic is impossible or difficult. [b-Czabanski]

Gaussian Naive Bayes

[AI/ST]: Gaussian Naive Bayes is a variant of Naive Bayes that follows a Gaussian normal distribution and supports continuous data. [b-Gaussian]

General Data Protection Regulation (GDPR)

[G]: Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. [b-EU]

General Regularly-distributed Information in Binary form (GRIB)

[GEO]: GRIB is a World Meteorological Organization (WMO) standard format for archiving and exchanging gridded data. GRIB is a binary format, and the data are packed to increase storage efficiency. [b-WMO-1]

Generalization

[AI/ST]: Refers to a model's ability to make correct predictions on new, previously unseen data as opposed to the data used to train the model. [b-Washington]

Genetic algorithm

[AI/ST]: Algorithm that simulates natural selection by creating and evolving a population of individuals (solutions) for optimization problems. [b-ScienceDirect-3]

Geodesy

[G]: The study of the shape, size, and gravity of the Earth. [b-NOAA]

Geographical information system (GIS)

[G]: A geographic information system is a type of database containing geographic data (that is, descriptions of phenomena for which location is relevant), combined with software tools for managing, analyzing, and visualizing those data. [b-Chang]

Geospatial data

[G]: Information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the earth. This information may be derived from, among other things, remote sensing, mapping, and surveying technologies. [b-WhiteHouse]

Geospatial mapping

[G]: Geospatial mapping is a type of spatial analysis techniques that typically employs software capable of rendering maps processing spatial data, and applying analytical methods to terrestrial or geographic datasets, including the use of geographic information system. [b-IGIGlobal]

Geostationary satellites

[G]: A geostationary satellite is in an orbit that can only be achieved at an altitude very close to 35,786 km (22,236 miles) and that keeps the satellite fixed over one longitude at the equator. [b-ScienceDirect-4]

Global Disaster Alert and Coordination System (GDACS)

[G]: GDACS is a cooperation framework between the United Nations, the European Commission, and disaster managers worldwide to improve alerts, information exchange and coordination in the first phase after major sudden-onset disasters. [b-GDACS]

Global Geodetic Reference Frame

[G]: A globally coordinated approach to geodesy, which is the discipline focused on accurately measuring the shape, rotation, and gravitational field of the Earth. [b-UNGGRF]

Global Health Security Index (GHS)

[G]: The 2021 GHS Index measures the capacities of 195 countries to prepare for epidemics and pandemics. [b-GHS]

Global navigation satellite system (GNSS)

[G]: A system of satellites that provide autonomous geo-spatial positioning with global coverage. The global navigation satellite system receiver calculates its own position on Earth. This positional information can be used in many applications such as mapping, surveying, navigation and mobile geographical information system (GIS). [ITU-T L.94]

Global Positioning System (GPS)

[G]: The Global Positioning System is a United States-owned utility that provides users with positioning, navigation, and timing (PNT) services. This system consists of three segments: the space segment, the control segment, and the user segment. [b-GPS]

Goodness-of-data

[G]: Accuracy, reliability, and validity of data. [b-Nkansah]

Goodness-of-fit

[AI/ST]: How well a model, a theoretical distribution, or an equation matches actual data. [b-Maydeu-Olivares]

Google Earth Engine

[G]: Google Earth Engine combines a multi-petabyte catalog of satellite imagery and geospatial datasets with planetary-scale analysis capabilities. [b-Earth]

Governance

[G]: Governance sets the parameters under which management operates, including how power is distributed and shared, how policies are formulated, priorities set and stakeholders made accountable. Note: Governance is about the definition of the strategic vision and direction, the formulation of the high-level goals and policies of an organization and the overseeing of its management. [b-UNESCO-2]

Gradient boost machine

[AI/ST]: A type of machine learning technique that uses an ensemble of weak prediction models to perform regression and classification tasks. [b-Springer-1]

Gramian angular field

[G]: A Gramian angular field is an image obtained from a time series, representing some temporal correlation between each time point. [b-pyts]

Grid search

[AI/ST]: Grid search is a tuning technique that attempts to compute the optimum values of hyperparameters. It is a search that is performed on the specific parameter values of a model. [b-Medium]

Ground motion model (GMPEs)

[G]: Ground motion prediction equations, also called ground-motion models (GMMs) and attenuation relations, estimate the shaking (strong ground motion) that may occur at a site if an earthquake of a certain magnitude occurs at a nearby location. [b-Douglas]

Ground truth

[AI/ST]: Value of the target variable for a particular item of labelled input data. [b-ISO/IEC 22989]

Ground truth annotation

[AI/ST]: Humanly provided classifications of the data on which the algorithms are trained or against which they are evaluated. [b-UnderstandAI]

Ground truth data

[G]: See "Ground truth"

Hail

[G]: Small, roughly spherical lumps of approximately concentric shells of clear ice and compact snow, collectively, that are precipitated during thunderstorms. [b-ICID]

Hailstorms

[G]: A storm accompanied by hail. [b-IFRC]

Hazard map

[G]: The process of establishing geographically where and to what extent particular phenomena are likely to pose a threat to people, property, infrastructure and economic activities. [b-PNSN]

[GEO]: Hazard maps are developed to illuminate areas that are affected or vulnerable to a particular hazard. They are typically made for natural hazards such as earthquake ground motion, flooding, landslides, liquefaction, and tsunami. Hazard maps are tools that when properly utilized by planners, developers, and engineers, can save lives and economic losses by avoiding exposure to some hazards while designing other development to mitigate or neutralize the potential negative effects of these hazards. [b-GFDRR]

Hidden layer

[AI/ST]: A synthetic layer in a neural network between the input layer (that is, the features) and the output layer (the prediction). Hidden layers typically contain an activation function (such as ReLU) for training. A deep neural network contains more than one hidden layer. [b-Teif]

Hidden Markov model (HMM)

[AI/ST]: A hidden Markov model is a statistical Markov model in which the system being modeled is assumed to be a Markov process – call it X – with unobservable ("hidden") states. As part of the definition, HMM requires that there be an observable process Y whose outcomes are "influenced" by the outcomes of X in a known way. [b-DoD]

Human-in-the-Loop (HITL)

[G]: Human-in-the-loop is defined as a model that requires human interaction. Human-in-the-loop systems allow humans to change the output of the learning systems. Human-in-the-loop simulators always have human input as part of the simulation, and humans influence the outcomes of the simulation exercise such that the outcomes may not be exactly reproducible. [b-ScienceDirect-5]

Hydrography

[GEO]: Science that deals with the measurements and description of the physical features of the oceans, seas, lakes, rivers, and their adjoining coastal areas, with reference to their use for navigational purposes. [b-Oxford]

Hyperparameter

[AI/ST]: Hyperparameters are higher-level properties of a model such as how fast it can learn (learning rate) or complexity of a model. The depth of trees in a decision tree or number of hidden layers in a neural network are examples of hyper parameters. [b-MLGlossary]

Image segmentation

[G]: The division of a complex picture into parts corresponding to regions or objects, so that the picture can then be described in terms of the parts, their properties, and their spatial relationship; also, grouping together those parts on an image description that come from an object in the scene prior to their recognition. [b-McGraw-Hill]

Inference

[G]: In statistics, inference refers to the process of fitting the parameters of a distribution conditioned on some observed data. [b-ScienceDirect-6]

[AI/ST]: Process by which a deployed machine learning model generates a result. Examples of generated result from machine learning model are prediction or classification. [ITU-T Y.3179]

Information and communication technologies (ICTs)

[G]: Technologies and equipment that handle (e.g., access, create, collect, store, transmit, receive, disseminate) information and communication. [ITU-T L.1430]

Information gain ratio (IGR)

[G]: Information gain is the reduction in entropy produced from partitioning a set with attributes and finding the optimal candidate that produces the highest value. [b-Moral]

[AI/ST]: In decision forests, the difference between a node's entropy and the weighted (by number of examples) sum of the entropy of its children nodes. A node's entropy is the entropy of the examples in that node. [b-Quinlan]

Infrasound

[G]: Sound-like waves having a frequency below the audible range, that is, below about 16 Hz. [b-Leventhall]

Inlier

[G]: An inlier is a data value that lies in the interior of a statistical distribution and is in error. Because inliers are difficult to distinguish from good data values they are sometimes difficult to find and correct. [b-Greenacre]

Input layer

[AI/ST]: The first layer (the one that receives the input data) in a neural network. [b-ScienceDirect-7]

Integrated gradients

[AI/ST]: Integrated Gradients aim to explain the relationship between a model's predictions in terms of its features. [b-Tensor-2]

Internal gravity wave (IGW)

[G]: A wave that propagates in density-stratified fluid under the influence of buoyancy forces. [b-AMS]

Internet of Things (IoT)

[G]: A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies. [ITU-T Y.4000]

Interoperability

[G]: The ability of two or more systems or applications to exchange information and to mutually use the information that has been exchanged. [ITU-T Y.101]

Interpolation

[G]: Process of obtaining the value of a function corresponding to a value of the argument intermediate between those given determination of points intermediate between known points on a desired path or contour in accordance with a given mathematical function, for example linear, circular, or higher-order functions. [b-Springer-1]

[AI/ST]: In the mathematical field of numerical analysis, interpolation is a type of estimation, a method of constructing (finding) new data points based on the range of a discrete set of known data points. [b-ISO/IEC 27729] [b-Steffensen]

Interpretability

[G]: Interpretability is the degree to which a human can understand the cause of a decision or to which a human can consistently predict the model's result. [b-Miller] [b-Interpretability]

[AI/ST]: Interpretable machine learning is a useful umbrella term that captures the extraction of relevant knowledge from a machine learning model concerning relationships either contained in data or learned by the model or the level of understanding how the underlying (AI) technology works. [b-ISO/IEC TR 29119-11] [b-Murdoch]

Intersection over Union (IoU)

[G]: The intersection of two sets divided by their union. [b-Standford-1]

[AI/ST]: In machine learning image detection tasks, IoU is used to measure the accuracy of the model's predicted bounding box with respect to the ground-truth bounding box. In this case, the IoU for the two boxes is the ratio between the overlapping area and the total area, and its value ranges from 0 (no overlap of predicted bounding box and ground-truth bounding box) to 1 (predicted bounding box and ground-truth bounding box have the exact same coordinates). [b-Standford-2]

Ionospheric total electron content (TEC)

[G]: Total number of electrons present along a path between a radio transmitter and receiver. [b-NOAA-2]

Ionospheric total electron disturbances

[GEO]: Ionospheric disturbances are a short-term deviation from regular climatology (such as diurnal variations). These disturbances can be local, regional, and sometimes global. [b-Yukitoshi]

Jackknife

[AI/ST]: 1. Example of a resampling method that estimates the variability of a statistic from the variability of that statistic across subsamples, rather than from parametric assumptions. 2. The jackknife is a method used to estimate the variance and bias of a large population. [...] It involves a leave-one-out strategy of the estimation of a parameter (e.g., the mean) in a dataset of N observations (or records). Ideally, N - 1 models are built on the dataset with different factors left out of each model. The estimates of all models are then aggregated into a single estimate of the parameter. The jackknife gets computationally intractable as $N \rightarrow \infty$. [b-Sinharay]

JavaScript Object Notation (JSON)

[G]: Open and text-based exchange format. [b-ISO/TS 23029:2020]

***k*-fold cross validation**

[AI/ST]: Cross validation is a resampling procedure used to evaluate machine learning models on a limited data sample. The procedure has a single parameter called *k* that refers to the number of groups into which a given data sample is split. As such, the procedure is often called k-fold cross validation. [b-MLMastery-6]

***K*-nearest neighbor (KNN or *k*-NN)**

[AI/ST]: The *k*-nearest neighbors' algorithm is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point. [b-IBM-3]

Key-value store system

[G]: A key-value database is a type of nonrelational database that uses a simple key-value method to store data. A key-value database stores data as a collection of key-value pairs in which a key serves as a unique identifier. Both keys and values can be anything, ranging from simple objects to complex compound objects. Key-value databases are highly partitionable and allow horizontal scaling at scales that other types of databases cannot achieve. [b-AWS]

Keyword detection

[AI/ST]: Keyword extraction in natural language processing is the task of finding the words that best describe the subject of a text. [b- Willyan]

Knowledge maps

[G]: A knowledge map is a visual display of captured information and relationships, which enables the efficient communication and learning of knowledge by observers with differing backgrounds at multiple levels of detail. The individual items of knowledge included in such a map can be text, stories, graphics, models, or numbers. [...] Knowledge mapping is defined as the process of associating items of information or knowledge (preferably visually) in such a way that the mapping itself also creates additional knowledge. [b-Vail]

Kolmogorov–Smirnov test (K–S test or KS test)

[AI/ST]: In statistics, the Kolmogorov–Smirnov test is a nonparametric test of the equality of continuous, one-dimensional probability distributions that can be used to compare a sample with a reference probability distribution (one-sample K–S test), or to compare two samples (two-sample K-S test) [b-Naaman]

Labeled data

[AI/ST]: In supervised learning, the "answer" or "result" portion of an example. Each example in a labeled dataset consists of one or more features and a label. For instance, in a housing dataset, the features might include the number of bedrooms, the number of bathrooms, and the age of the house, while the label might be the house's price. In a spam detection dataset, the features might include the subject line, the sender, and the email message itself, while the label would probably be either "spam" or "not spam." [b-MachineLearning]

Land cover

[G]: Land cover maps represent spatial information on different types (classes) of physical coverage of the Earth's surface (e.g., forests, grasslands, croplands, lakes, or wetlands). [b-Copernicus]

Land use

[G]: Land use corresponds to the socio-economic description (functional dimension) of areas: areas used for residential, industrial, or commercial purposes; for farming or forestry; for recreational or conservation purposes, etc. Links with land cover are possible; it may be possible to infer land use from land cover and conversely. However, situations are often complicated and the link is not always evident. Contrary to land cover, land use is difficult to 'observe.' For example, it is often difficult to decide if grasslands are used or not for agricultural purposes. Distinctions between land use and land cover and their definition have impacts on the development of classification systems, data collection, and information systems in general. [b-EEA]

Landslide

[GEO]: Wide variety of processes that result in the downward and outward movement of slope-forming materials including rock, soil, artificial fill, or a combination. [b-ISO 22327]

Landslide inventory map

[GEO]: A landslide inventory map records the location and, where known, the date of occurrence and the types of mass movements that have left discernable traces in an area. [b-Fausto]

Latent space

[AI/ST]: Formally, an abstract multi-dimensional space that encodes a meaningful internal representation of externally observed events. Samples that are similar in the external world are positioned close to each other in the latent space. [b-Baeldung]

Layerwise relevance propagation (LRP)

[AI/ST]: A methodology that allows to visualize the contributions of single pixels to predictions for kernel-based classifiers over bag of words features and for multilayered neural networks. [b-Binder] [b-Bach]

Lead time

[G]: A lead time is the latency between the initiation and completion of a process. [b-Business]

Least developed countries (LDC)

[G]: Low-income countries confronting severe structural impediments to sustainable development. They are highly vulnerable to economic and environmental shocks and have low levels of human assets. [b-UNDESA]

Life cycle

[G]: Consecutive and interlinked stages of a product system, from raw material acquisition or generation from natural resources to final disposal. [b-ISO]

[AI/ST]: The machine learning life cycle is the cyclical process that data science projects follow. It defines each step that an organization should follow to take advantage of machine learning and artificial intelligence to derive practical business value. [b-DataRobot]

[GEO]: In biology, a biological life cycle (or just life cycle or lifecycle when the biological context is clear) is a series of changes in form that an organism undergoes, returning to the starting state. "The concept is closely related to those of the life history, development, and ontogeny, but differs from them in stressing renewal." [b-Rodrigues]

Light detection and ranging (LiDAR)

[G]: System consisting of 1) a photon source (frequently, but not necessarily, a laser), 2) a photon detection system, 3) a timing circuit, and 4) optics for both the source and the receiver that uses emitted laser light to measure ranges to and/or properties of solid objects, gases, or particulates in the atmosphere. [b-ISO/IEC 22989]

Linear discriminant analysis (LDA)

[AI/ST]: Linear discriminant analysis, normal discriminant analysis (NDA), or discriminant function analysis is a generalization of Fisher's linear discriminant, a method used in statistics and other fields, to find a linear combination of features that characterizes or separates two or more classes of objects or events. The resulting combination may be used as a linear classifier, or, more commonly, for dimensionality reduction before later classification. [b-McLachlan]

Linear programming (LP)

[G]: Linear programming, also called linear optimization, is a method to achieve the best outcome (such as maximum profit or lowest cost) in a mathematical model whose requirements are represented by linear relationships. [b-Fukuda]

Log-loss

[AI/ST]: The loss function used in binary logistic regression. [b-ISO/IEC 22989]

Logistic regression

[AI/ST]: A classification model that uses a sigmoid function to convert a linear model's raw prediction into a value between 0 and 1. Although logistic regression is often used in binary classification problems, logistic regression can also be used in multi-class classification problems (where it becomes called multi-class logistic regression or multinomial regression). [b-TowardsScience]

Long short-term memory (LSTM)

[AI/ST]: Type of recurrent neural network that processes sequential data with a satisfactory performance for both long and short span dependencies. It can be considered as a type of cell in a recurrent neural network used to process sequences of data in applications such as handwriting recognition, machine translation, and image captioning. LSTMs address the vanishing gradient problem that occurs when training Recurrent Neural Networks (RNN) due to long data sequences by maintaining history in an internal memory state based on new input and context from previous cells in the RNN. [b-Dictionary]

Long Range (LoRa)

[G]: LoRa is a spread spectrum modulation technique derived from chirp spread spectrum (CSS) technology. Semtech's LoRa is a long range, low power wireless platform that has become the de facto wireless platform of Internet of Things. [b-semtech]

Loss function

[G]: In the making of decisions on the basis of observations on a variate *x*, disadvantage may be suffered through ignorance of the true distribution of *x*. The extent of the disadvantage is often a function of the true distribution and of the decision that is actually made. This is called the loss function. [b-IEEE]

[AI/ST]: A loss function is the difference between the training label values and the prediction made by the model. The parameters of the model are estimated by minimizing the loss function. Different trainers can be configured with different loss functions. [b-MLMastery-7]

Machine learning (ML)

[G]: Processes that enable computational systems to understand data and gain knowledge from it without necessarily being explicitly programmed. [ITU-T Y.3172]

[AI/ST]: A program or system that builds (trains) a predictive model from input data. The system uses the learned model to make useful predictions from new (never-before-seen) data drawn from the same distribution as the one used to train the model. Machine learning also refers to the field of study concerned with these programs or systems. [b-Breuel]

Machine learning operations (ML Ops)

[AI/ST]: ML Ops is a set of practices that combines machine learning, DevOps(*), and data engineering, which aims to deploy and maintain ML systems in production reliably and efficiently. [ITU-T Y.3172]

(*): DevOps (a portmanteau of "development" and "operations") is the combination of practices and tools designed to increase an organization's ability to deliver applications and services faster than traditional software development processes. This speed enables organizations to better serve their customers and compete more effectively in the market (https://www.synopsys.com/glossary/what-is-devops.html).

Machine learning pipeline

[G]: A set of logical nodes, each with specific functionalities, which can be combined to form a machine learning application in a telecommunication network. [ITU-T Y.3172]

Majority voting

[AI/ST]: Simple majority voting is used to ensemble the classifiers in determining detection accuracy. This is an iterative phase in which a threshold (acceptable detection accuracy set) is set and checked with the evaluation results until an optimum result is achieved. [b-ScienceDirect-8]

Maximum entropy (MAXENT)

[AI/ST]: The maximum entropy approach is rooted in information theory and has been successfully applied to many fields including physics and natural language processing. It creates a model that best accounts for the available data but with a constraint that without any additional information the model should maximize entropy. In other words, the model prefers a uniform distribution by maximizing the conditional entropy. [b-Handbook]

Maximum likelihood estimation (MLE)

[AI/ST]: In statistics, MLE is a method of estimating the parameters of an assumed probability distribution, given some observed data. This is achieved by maximizing a likelihood function so that, under the assumed statistical model, the observed data are most probable. The point in the parameter space that maximizes the likelihood function is called the maximum likelihood estimate. [b-Cramer-1]

Mean absolute error (MAE)

[AI/ST]: The mean of the algebraic result of subtracting a true, specified, or theoretically correct value from the value computed, observed, measured, or achieved. [b-ISO/IEC 2382]

Mean absolute percentage error (MAPE)

[AI/ST]: The mean absolute percentage error, also known as mean absolute percentage deviation (MAPD), is a measure of prediction accuracy of a forecasting method in statistics. It usually expresses the accuracy as a ratio defined by the formula:

$$\mathrm{MAPE} = (100\%/n) * \Sigma \mid (y(i)-x(i))/\mid y(i) \mid$$

where $x(i)$ is the actual value and $y(i)$ is the forecast value and i=[1,..,n]. Their difference is divided by the actual value $x(i)$. The absolute value of this ratio is summed for every forecasted point in time and divided by the number of fitted points n. [b-MAPE]

Message Queue Telemetry Transport (MQTT)

[G]: Client server publish/subscribe messaging transport protocol. It is lightweight, open, simple, and designed to be easy to implement. These characteristics make it ideal for use in many situations, including constrained environments such as communication in machine to machine (M2M) and Internet of Things contexts where a small code footprint is required and/or network bandwidth is at a premium. The protocol runs over transmission control protocol/internet protocol, or over other network protocols that provide ordered, lossless, and bi-directional connections. [b-ISO/IEC 20922]

Meta learning

[AI/ST]: A subset of machine learning that discovers or improves a learning algorithm. A meta-learning system can also aim to train a model to quickly learn a new task from a small amount of data or from experience gained in previous tasks. Meta-learning algorithms generally try to achieve the following: improve/learn hand-engineered features (such as an initializer or an optimizer), be more data-efficient and compute-efficient, improve generalization.

Meta-learning is related to few-shot learning. [b-MachineLearning]

Metadata

[G]: Structured, encoded data that describe characteristics of information-bearing entities such as their data descriptions, data about data ownership, access paths, access rights and data volatility to aid in the identification, discovery, assessment and management of the described entities. [ITU-T Y.3519] [ITU-T H.752][2]

[AI/ST]: Additional data associated with the image data beyond the image data. [b-Springer-1]

2 This definition is based on both the indicated references.

Micro-Electro-Mechanical Systems (MEMS)

[G]: Micro-electro-mechanical systems (or microelectronic and microelectromechanical systems) and the related micromechatronics and microsystems constitute the technology of microscopic devices, particularly those with moving parts. [b-MEMS]

Mid-latitudes

[G]: The middle latitudes (also called the mid-latitudes, sometimes midlatitudes, or moderate latitudes) are a spatial region on Earth located between the latitudes 23°26'22" and 66°33'39" North, and 23°26'22" and 66°33'39" South. [b-Moon] [b-Climate]

Missing value treatment

[G]: The handling of missing data during the preprocessing of a dataset as many machine learning algorithms do not support missing values. Treatment depends on the type of missing values and the replacement methodology. [b-Waldner] [b-Analytics] [b-Samuylova]

Monitoring

[AI/ST]: Machine learning monitoring is a practice of tracking and analyzing production model performance to ensure acceptable quality as defined by the use case. It provides early warnings on performance issues and helps diagnose their root cause to debug and resolve. [b-Clapham]

Monotonicity

[G]: In mathematics, monotonicity is a function between ordered sets that preserves or reverses the given order. This concept first arose in calculus, and was later generalized to the more abstract setting of order theory. [b-XGBoost]

[AI/ST]: In model evaluation, monotonicity refers to the application of monotonic constrains in order to enforce a monotonic relationship between the prediction feature space and the target variable. [b-XGBoost]

Multi-hazard early warning/alert systems

[GEO]: The set of capacities needed to generate and disseminate timely and meaningful warning information to enable individuals, communities and organizations threatened by a hazard to prepare and to act appropriately and in sufficient time to reduce the possibility of harm or loss. [ITU-T E.102]

Multilayer-feed-forward

[AI/ST]: Neural network where information is fed from the input layer to the output layer in one direction only and connections between the nodes do not form a cycle. [b-ISO/IEC 22989]

Multivariate analysis

[AI/ST]: Multivariate analysis provides models and procedures for dealing separately with each of a number of variables in estimation, while at the same time providing tests of hypotheses which lead to a single probability statement referring to all variables jointly. [b-Cramer-2]

Naive Bayes

[AI/ST]: A simple learning algorithm that utilizes Bayes rule together with a strong assumption that the attributes are conditionally independent, given the class. [b-Encyclopedia]

Natural disaster management (NDM)

[GEO]: Natural disaster management is a process of effectively preparing for and responding to natural disasters. It involves strategically organizing resources to lessen the harm that they cause. It also involves a systematic approach to managing the responsibilities of disaster prevention, preparedness, response, and recovery. [b-USDept] [b-Tulane]

Natural disaster management risk alert

[GEO]: See "Multi-hazard early warning/alert systems."

Natural disasters

[GEO]: Natural disasters are generally defined as a "potentially damaging physical event" of a predominantly natural origin (e.g., atmospheric, hydrologic, geophysical, oceanographic, or biologic). Adverse effects of these events include injury, mortality, displacements, damage to property (including cultural heritage) and infrastructure, and disturbance to nature and natural resources. Previous version: A serious disruption of the functioning of a community or a society involving widespread human, material, economic or environmental losses and impacts, which exceeds the ability of the affected community or society to cope using its own resources. [b-NOAAStudent]

Natural language processing (NLP)

[AI/ST]: The process of converting a piece of English text into a programmer-friendly data structure that describes the meaning of the natural language text. The natural language processing tasks include part-of-speech tagging, chunking, named entity recognition, and semantic role labelling. [ITU-T E.475]

Near-real time

[G]: Real-time computing (RTC) is the computer science term for hardware and software systems subject to a "real-time constraint," for example, from event to system response. Real-time programs must guarantee response within specified time constraints, often referred to as "deadlines." [b-Computer]

NerveNet

[AI/ST]: NerveNet addresses the problem of learning structured policies for continuous control by explicitly modeling the structure of an agent, which naturally takes the form of a graph. Specifically, serving as the agent's policy network, NerveNet first propagates information over the structure of the agent and then predicts actions for different parts of the agent. [b-NerveNet]

Neural network

[AI/ST]: Network of one or more layers of neurons connected by weighted links with adjustable weights, which takes input data and produces an output. Neural networks are a prominent example of the connectionist approach. Although the design of neural networks was initially inspired by the functioning of biological neurons, most works on neural networks do not follow that inspiration anymore. [b-ISO/IEC 22989]

Neurons

[AI/ST]: Primitive processing element that takes one or more input values and produces an output value by combining the input values and applying an activation function on the result. Examples of nonlinear activation functions are a threshold function, a sigmoid function, and a polynomial function. [b-ISO/IEC 22989]

Next Generation Set-Top-Box (NG-STB)

[G]: A compilation of hardware and software functional entities contained within one or more physical devices, which at a baseline level provides the receiving functions for cable broadcast services. In addition, the NG-STB may support the interactive functions of Internet Protocol (IP)-based services, additional time-critical services between the Hybrid fiber coaxial (HFC) and the home network as well as extension and supplemental services. [ITU-T J.193]

Noise reduction

[G]: The process of removing noise from a signal. [b-Chen]

Non-graphical univariate analysis

[AI/ST]: A class of single-feature data analysis that does not use visual tools. This preliminary data analysis step focuses on four points; that is, for mechanisms that should be examined. These include measures of central tendency (mean, median, and mode), measures of spread (variability, variants, and standard deviation), the shape of the distribution, and the existence of outliers. [b-PennState]

Normalization

[AI/ST]: The process of converting an actual range of values into a standard range of values, typically from -1 to +1 or from 0 to 1. For example, suppose the natural range of a certain feature is from 800 to 6,000. Through subtraction and division, those values can be normalized into the range from −1 to +1. [b-WMO-2]

Nowcasting

[GEO]: Nowcasting is weather forecasting on a very short-term mesoscale period of up to two hours according to the World Meteorological Organization and up to six hours according to other sources. [b-WMO-2]

Numerical simulation

[GEO]: A calculation that is run on a computer following a program that implements a mathematical model for a physical system. Numerical simulations are required to study the behavior of systems whose mathematical models are too complex to provide analytical solutions, as in most nonlinear systems. [b-Nature-1]

Numerical weather prediction (NWP) model

[GEO]: NWP models process current weather observations to forecast future weather. Output is based on current weather observations, which are assimilated into the model's framework and used to produce predictions for temperature, precipitation, and hundreds of other meteorological elements from the oceans to the top of the atmosphere. [b-NECI]

Numerosity reduction

[AI/ST]: In numerosity reduction, the data volume is decreased by selecting an alternative, smaller form of data representation. These techniques can be parametric or non-parametric. For parametric methods, a model can estimate the data so that only the data parameters need to be saved, instead of the actual data, for example, log-linear models. Non-parametric methods are used to store a reduced representation of the data, including histograms, clustering, and sampling. [b-Java]

Onsite approach

[GEO]: Onsite approach (e.g., common in earthquake early warning algorithms) uses a sensor located in one location to warn the same location. [b-Böse]

Open access

[G]: A good or service for which no property rights are recognized. [b-National]

Open data

[G]: Data that are publicly accessible through open standards and protocols or through other means. The use and redistribution of open data can be subject to rules. [ITU-T Y.4472]

Optimizer

[AI/ST]: A specific implementation of the gradient descent algorithm, for instance, ADAptive GRADient descent (AdaGrad) and ADAptive with momentum (Adam). [b-DeepAI]

Orthoimagery

[G]: Orthoimagery (also known as orthophotos, orthophotographs, or orthoimages) is an aerial photograph or satellite imagery geometrically corrected ("orthorectified") such that the scale is uniform: the photo or image follows a given map projection. Unlike an uncorrected aerial photograph, an orthophoto can be used to measure true distances, because it is an accurate representation of the Earth's surface, having been adjusted for topographic relief, lens distortion, and camera tilt. [b-USGS]

Orthomosaics

[GEO]: The output from a process where a number of overlapping photos (e.g., from a drone or aerial camera) are stitched together with distortions removed to create a complete and continuous image representation or map of a portion of the Earth. [b-GEONADIR]

Outliers

[G]: Values that are sometimes present in distributions that are characterized by extreme values that differ greatly from the other observations. [ITU-T E.475]

Output layer

[AI/ST]: The "final" layer of a neural network. The layer containing the answer(s). [b-DeepAI]

Overfitting

[AI/ST]: Creating a model that matches the training data so closely that the model fails to make correct predictions on new data. [b-Washington]

Parasite

[G]: Organism, plant, or animal, that lives in or on another living organism of a different kind and derives subsistence from it without returning any benefit. [b-Ford] [b-RSPB]

Patch

[GEO]: A patch is the best spatial unit with adaptive scale characteristics to characterize the spatial environment and to measure the effects of a hydrometeorological phenomenon. [b-NaturalHazards]

Peak signal-to-noise ratio (PSNR)

[G]: The ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. PSNR is defined as 10*log10 of the ratio of the peak signal energy to the MSE observed between the processed signal and the original signal. PSNR is commonly used to quantify reconstruction quality for images and video subject to lossy compression. [ITU-T J.340]

Pearson correlation coefficient (PCC)

[AI/ST]: In statistics, the Pearson correlation coefficient, also known as Pearson's *r*, the Pearson product-moment correlation coefficient (PPMCC), the bivariate correlation, or colloquially simply as the correlation coefficient is a measure of linear correlation between two sets of data. It is the ratio between the covariance of two variables and the product of their standard deviations; thus, it is essentially a normalized measurement of the covariance, such that the result always has a value between −1 and 1. [b-SPSS]

Peer confrontation

[AI/ST]: Peer confrontation in AI is a type of behavioral evaluation. The assessment is performed through a series of 1-vs-1 or multi-agent 'matches.' The result is relative to the other participants. Given this relative value, in order to allow for a numerical comparison, sophisticated performance metrics can be derived. [b-Hernández-Orallo]

Performance diagrams

[AI/ST]: Evaluation diagrams and graphs that allow the visualization and comparison of model performance and behavior in comparison to observations with descriptive statistics (e.g., Taylor diagram, vector field evaluation diagram). [b-Taylor]

Personal identifiable information (PII)

[G]: Personal data or PII (mainly used in the United States) is any information related to an identifiable person. [b-VAHandbook]

Photoperiod

[GEO]: A recurring cycle of light and dark periods of constant length. [b-Vergara]

Physical sensor/acquisition devices

[GEO]: A device that transforms a physical value into an electrical or logical unit and can therefore sense a physical condition or chemical compound by delivering an electronic signal proportional to the observed characteristic. The sensor can be directly connected with a data stream to the management system or via a conversion device and is an integral component of the electrical grid. [ITU-T L.1301] [ITU-T Y.2071] [ITU-T Y.4113]

Physics-informed machine learning (PIML)

[AI/ST]: Physics-informed machine learning seamlessly integrates data and mathematical physics models, even in partially understood, uncertain, and high-dimensional contexts. [b-Nature-2]

Point source algorithm

[GEO]: Earthquake early warning algorithms that provide fast estimates of earthquake characteristics (such as location or magnitude) using a few seconds of P-wave records. [b-ETH]

Policy mapping

[AI/ST]: Policy mapping in reinforcement learning is an agent's probabilistic mapping from states to actions. [b-Sallan]

Pooling

[AI/ST]: Reducing a matrix (or matrices) created by an earlier convolutional layer to a smaller matrix. Pooling usually involves taking either the maximum or average value across the pooled area. A pooling operation, just like a convolutional operation, divides that matrix into slices and then slides that convolutional operation by strides. Pooling helps enforce translational invariance in the input matrix. Pooling for vision applications is known more formally as spatial pooling. Time-series applications usually refer to pooling as temporal pooling. Less formally, pooling is often called subsampling or downsampling. [b-ScienceDirect-10]

Prediction rate curve

[AI/ST]: The area under the receiver operating characteristics curve analysis used to evaluate model predictability. [b-Lobo]

Prediction/forecasting

[G]: Forecasting is the process of making predictions based on past and present data. [b-Scott]

Preprocessing

[G]: Processing data before they are used to train a model. Preprocessing could be as simple as removing words from an English text corpus that do not occur in the English dictionary, or could be as complex as re-expressing data points in a way that eliminates as many attributes that are correlated with sensitive attributes as possible. Preprocessing can help satisfy fairness constraints. It is the process of transforming raw data into a more understandable format. [b-MachineLearning]

Primary earthquake phase

[GEO]: The primary earthquake phase is the one characterized by the P (primary) seismic waves generated by an earthquake source that propagates within the body of the Earth. The P seismic waves travel as elastic motions at the highest speeds and because of their greater speed than the other two seismic wave classes, P waves are the first to reach any point on the Earth's surface. [b-Britannica]

Principal component analysis (PCA)

[AI/ST]: Constructing new features, which are the principal components of a dataset. The principal components are random variables of maximal variance constructed from linear combinations of the input features. Equivalently, they are the projections onto the principal component axes, which are lines that minimize the average squared distance to each point in the dataset. To ensure uniqueness, all of the principal component axes must be orthogonal. PCA is a maximum-likelihood technique for linear regression in the presence of Gaussian noise on both inputs and outputs. In some cases, PCA corresponds to a Fourier transform, such as the discrete cosine transform used in JPEG image compression. [b-Springer-1]

Probabilistic graphical models (PGMs)

[AI/ST]: Statistical models that encode complex joint multivariate probability distributions using graphs. [b-DataS]

Probabilistic model

[AI/ST]: A regression or classification model that uses not only the weights for each feature, but also the uncertainty of those weights. A probabilistic model generates a prediction and the uncertainty of that prediction or the probability distribution over a set of classes. [b-Probability]

Prompt elasto-gravity signal (PEGS)

[GEO]: The sudden displacement of rock mass induced by an earthquake generates density variations that, in turn, modify the Earth's gravity field. The signal associated with these transient gravity perturbations propagates at the speed of light, much faster than the fastest elastic waves (P-waves). The induced elastic response of direct gravity perturbations is named prompt elastogravity signals. [b-Licciardi]

Proxy

[G]: Application of programming interface implementation that conveys action requests and parameters to a layer implementation somewhere else. [b-ISO/IEC 24727]

[GEO]: In the study of past climates ("paleoclimatology"), climate proxies are preserved physical characteristics of the past that stand in for direct meteorological measurements and enable scientists to reconstruct the climatic conditions over a longer fraction of the Earth's history. [b-Carbonplace]

Quality evaluation

[G]: The process used to determine the satisfiability of a solution with respect to its goal(s). [b-MachineLearning]

Quickpropagation

[AI/ST]: Quickpropagation or QuickProp is a second-order optimization algorithm that uses a simple approximation of the Hessian's diagonal to accelerate optimization and therefore belongs to the class of quasi-Newton algorithms. [b-Computer] [b-Brust]

Radar

[G]: A radiodetermination system based on the comparison of reference signals with radio signals reflected, or retransmitted, from the position to be determined. [b-Radio] [ITU-R RS.1632] [ITU-R RS.577-7]

Random forest

[AI/ST]: Random forests or random decision forests are ensemble learning methods for data classification and regression. They construct a multitude of decision trees during the training and output the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. [b-Springer-1]

Random search

[AI/ST]: Random search is a hyperparameter optimization algorithm that utilizes a defined search space as a bounded domain of hyperparameter values and randomly samples points in that domain. [b-MachineLearning]

Random shuffling

[AI/ST]: An algorithm that modifies a sequence in place by shuffling its contents. [b-Numpy]

Real-time data streams

[G]: Data or services (e.g., broadcasting) that are transmitted with virtually no delay. [ITU-T F.791]

Real-time volcano monitoring (RT-VM)

[GEO]: Monitoring a volcano requires scientists to use of a variety of techniques that can hear and see activity inside a volcano and detect signs of change that forewarn of volcanic reawakening. To fully understand a volcano's behavior, monitoring should include several types of observations (earthquakes, ground movement, volcanic gas, rock chemistry, water chemistry, and remote satellite analysis) on a continuous or near-real-time basis. [b-Volcano]

Really simply syndication (RSS)

[G]: A family of Web feed formats used to publish often updated content such as blog entries, news headlines, or podcasts. An RSS document, which is called a "feed," "web feed," or "channel," contains either a summary of content from an associated web site or the full text. RSS makes it possible for people to keep up with their favorite web sites without having to check them manually. This flow of content between websites and users is called "web syndication." [b-RSS]

Recall

[AI/ST]: In pattern recognition, information retrieval, object detection, and classification (machine learning), precision and recall are performance metrics that apply to data retrieved from a collection, corpus or sample space. Precision (also called positive predictive value) is the fraction of relevant instances among the retrieved instances, while recall (also known as sensitivity) is the fraction of relevant instances that were retrieved. Both precision and recall are therefore based on relevance (see definition of "relevance"). [b-Powers]

Receiver operating characteristic (ROC)

[AI/ST]: A plot of the true positive rate against the false positive rate at all classification thresholds. This is used to evaluate the performance of a classification model at different classification thresholds. The area under the ROC curve can be interpreted as the probability that the model correctly distinguishes between a randomly chosen positive observation (e.g., "spam") and a randomly chosen negative observation (e.g., "not spam"). [b-CheatSheet]

Rectified linear unit activation function (ReLu)

[AI/ST]: An activation function with the following rules. If input is negative or 0, output is 0. If input is positive, output is equal to input. [b-MachineLearning]

Recurrent neural network (RNN)

[AI/ST]: Neural network in which outputs from both the previous layer and the previous processing step are fed into the current layer. [b-ISO/IEC 22989]

Regression model

[AI/ST]: A type of model that outputs continuous (typically floating point) values. [b-MachineLearning]

Regularization

[AI/ST]: The penalty on a model's complexity. Regularization helps prevent overfitting. [b-Springer-1]

Reinforcement learning (RL)

[AI/ST]: An area of machine learning concerned with how intelligent agents ought to take actions in an environment in order to maximize the notion of cumulative reward. Reinforcement learning is one of three basic machine learning paradigms, alongside supervised learning and unsupervised learning. [b-van Otterlo]

Relevance

[G]: In information science and information retrieval, relevance denotes how well a retrieved document or set of documents meets the information need of the user. Relevance may include concerns such as timeliness, authority, or novelty of the result. [b-Survey]

Reliability

[G]: Reliability is the probability and/or ability of a system, product, or component to perform and maintain under stated conditions as required for a specified period of time. [ITU-T L.1022] [ITU-T Y.3514] [ITU-T L.1202]

Remote sensing satellite

[G]: A satellite whose purpose is remote observation by reception of electromagnetic waves using active or passive sensors. [ITU-R V.573-4]

Remotely piloted aircraft System (RPAS)

[G]: A remotely piloted aircraft, its associated remote pilot station(s), the required command and control links, and any other components as specified in the type design. [b-ICAO]

Reproducibility

[AI/ST]: Reproducibility is obtaining consistent results using the same input data; computational steps, methods, and code; and conditions of analysis. This definition is synonymous with "computational reproducibility." [b-NAP]

Residual neural network (ResNet)

[AI/ST]: An artificial neural network (ANN), which is a gateless or open-gated variant of the HighwayNet, the first working very deep feedforward neural network with hundreds of layers. Skip connections or shortcuts are used to jump over some layers. Typical ResNet models are implemented with double- or triple-layer skips that contain nonlinearities (ReLU) and batch normalization in between. [b-MachineLearning]

Resiliency metrics

[G]: Variables that can be measured (if quantifiable) or otherwise tracked (if qualitative) that represent an indicator. An indicator in the resiliency domain may be qualities, traits, or states of a system that suggest ("indicate") or hint at the effectiveness, progress, or success of resiliency, which is the ability to deal continually with the challenges of a changing climate (including changes in average conditions, extreme events, and related disruptions) now and in the future. This may entail coping, adapting, and/or transforming (at a personal, community, or systems level) so as to maintain that ability. [b-MachineLearning]

Reuse

[G]: Process by which a product or its parts, having reached the end of their first use, are used for the same purpose for which they were conceived. [b-NAP]

Risk

[G]: Effect of uncertainty on objectives, possibility of loss or injury, someone or something that creates or suggests a hazard, or the chance that an investment (such as stock or commodity) will lose value. [b-Springer-1] [b-Kaiming]

Risk management

[G]: The reduction of disaster risk as an expected outcome, a goal focused on preventing new risk, reducing existing risk and strengthening resilience, as well as a set of guiding principles, including primary responsibility of states to prevent and reduce disaster risk, all-of-society and all-of-state institutions engagement. [b-Terminology] [ITU-T X.1451] [b-Sendai]

RiverCore

[GEO]: Fixed nodes whose primary function is to measure a river's depth in real time. They can be used in flash flood early warning systems. [b-Sensors]

Robot

[G]: Automation system with actuators that performs intended tasks in the physical world, by means of sensing its environment and a software control system. A robot includes the control system and interface of a control system. The classification of a robot as industrial robot or service robot is done according to its intended application. In order to properly perform its tasks, a robot makes use of different kinds of sensors to confirm its current state and perceive the elements composing the environment in which it operates. [b-MachineLearning] [b-ISO/IEC 22989]

Robust inference

[G]: The development of procedures that are still reliable and reasonably efficient under small deviations from the model, that is, when the underlying distribution lies in a neighborhood of the assumed model. [b-Robust]

Robustness

[G]: Property of the implementation of a specified Electronic Commerce Indicator (ECI) secure function representing the effort and/or cost involved to compromise the security of the implemented secure function. [ITU-T J.1014]

[AI/ST]: The characteristic of a model when its generated output and overall performance is satisfactory and admissible, even though data may derive from various probability distributions, contain outliers, and diverge from parametric distributions, expected data ranges, and other model assumptions. [b-Earth]

Root mean square error (RMSE)

[AI/ST]: The root mean square error of an estimator of a population parameter is the square root of the mean square error (MSE). The mean square error is defined as the expected value of the square of the difference between the estimator and the parameter. [b-RMSE]

Sample selection

[AI/ST]: Process to select a subset of data samples intended to present patterns and trends similar to that of the larger dataset being analyzed. Ideally, the subset of data samples will be representative of the larger dataset. [b-ISO/IEC 22989]

Sampling

[G]: See "Sample Selection."

Synthetic aperture radar (SAR)

[G]: A technique for producing fine-resolution images from a resolution-limited radar system. It requires that the radar be moving in a straight line, either on an airplane or orbiting in space. The basic principle of any imaging radar is to emit an electromagnetic signal (which travels at the speed of light) toward a surface and record the amount of signal that bounces/echoes back, or "backscatters," and its time delay. The resulting radar imagery is built up from the strength and time delay of the returned signal, which depends primarily on the roughness and electrical conducting properties of the observed surface and its distance from the orbiting radar. [b-NASAJPL]

Satellite sensor

[G]: A measuring instrument in the Earth exploration satellite service or in the space research service by means of which information is obtained by transmission and reception of electromagnetic waves. [ITU-R V.573-4]

Scale factoring

[G]: In mathematics, a number used as a multiplier in a scaling operation. [b-Fan] [GEO]: In geodesy, one of seven parameters used in the datum shift method called Helmert transformation. [b-MachineLearning]

Secure file transfer

[G]: A file transferring process, which is based in a secure file transfer protocol which provides encryption and security in transactions. [b-Kaczmarczyk]

Seed

[G]: The fertilized ripened ovule of a flowering plant containing an embryo and capable normally of germination to produce a new plant, broadly: a propagative plant structure (such as a spore or small dry fruit). [b-Deepchecks]

[AI/ST]: A random seed (or seed state, or just seed) is a number (or vector) used to initialize a pseudorandom number generator. [b-Geology]

Seismic cycle

[GEO]: The seismic cycle refers to the observation that earthquakes repeatedly rupture a given part of a fault. The term "cycle" does not however imply that earthquakes are a periodic or regularly repeating event. The seismic cycle can be divided into three periods, consisting of inter-seismic slip, co-seismic slip, and post-seismic slip. [b-Britannica]

Seismic network

[GEO]: Set of stations coordinated by a single centralizing body responsible for the acquisition, processing, and dissemination of seismological data. The existence of networks is fundamental because the unique recording of an earthquake is practically unusable if only for the localization of the source. [b-Evision]

Seismic wave

[GEO]: A vibration generated by an earthquake, explosion, or similar energetic source and propagated within the Earth or along its surface. [b-Evision]

Seismogenesis

[GEO]: Any process that generates earthquakes. [b-Geology]

Self-supervised learning

[AI/ST]: An unsupervised learning method where the supervised learning task is created out of the unlabeled input data. [b-ISO/IEC 22989]

Self training

[AI/ST]: A variant of self-supervised learning that is particularly useful when all of the following conditions are true. The ratio of unlabeled examples to labeled examples in the dataset is high. This is a classification problem. Self training works by iterating over the following two steps until the model stops improving: step 1 - use supervised machine learning to train a model on the labeled examples; step 2 - use the model created in step 1 to generate predictions (labels) on the unlabeled examples, moving those in which there is high confidence into the labeled examples with the predicted label. [b-Saltelli]

Semi-supervised learning

[AI/ST]: Machine learning that makes use of both labelled and unlabeled data during training. [b-DZone]

Sensitivity

[G]: Sensitivity analysis is the study of how the uncertainty in the output of a mathematical model or system (numerical or otherwise) can be divided and allocated to different sources of uncertainty in its inputs. [b-Elmenreich]

[AI/ST]: Sensitivity is a measure of the proportion of actual positive cases that got predicted as positive (or true positive). [b-Shannon]

Sensor fusion

[G]: Sensor fusion is the process of combining sensor data or data derived from disparate sources such that the resulting information has less uncertainty than would be possible when these sources were used individually. [b-MachineLearning]

Shannon information

[AI/ST]: In information theory, the information content, self-information, surprisal, or Shannon information is a basic quantity derived from the probability of a particular event occurring from a random variable. It can be thought of as an alternative way of expressing probability. The Shannon information is closely related to entropy, which is the expected value of the self-information of a random variable, quantifying how surprising the random variable is "on average." This is the average amount of self-information an observer would expect to gain about a random variable when measuring it. [b-Western]

Singular value decomposition (SVD)

[AI/ST]: In linear algebra, SVD is a factorization of a real or complex matrix. It generalizes the eigendecomposition of a square normal matrix with an orthonormal eigenbasis to any m x n matrix. [b-TowardsData]

Slip-zone

[GEO]: The zone that slips during the rupture of an earthquake. The relative displacement of formerly adjacent points on opposite sides of a fault, measured on the fault surface. [b-USGS]

Slow slip

[GEO]: Slow slip is the phenomenon of slow-slip events, which are long-lived shear-slip events at subduction interfaces and the physical processes responsible for the generation of slow earthquakes. They are slow thrust-sense displacement episodes that can have durations up to several weeks, and are thus termed "slow." [b-Michael]

Smart seismic networks

[G]: A smart seismic network system uses seismic sensors to perform tomography of the subsurface based on ambient noise cross-correlation measurements. This system can reproduce an image of the subsurface that represents the velocity variations in the subsurface when smart seismic sensors are deployed in the area of interest and they use wireless network communication and in-situ computing. [b-Valero]

Smoothing

[AI/ST]: In statistics and image processing, to smooth a dataset is to create an approximating function that attempts to capture important patterns in the data, while leaving out noise or other fine-scale structures/rapid phenomena. In smoothing, the data points of a signal are modified so individual points higher than the adjacent points (presumably because of noise) are reduced, and points that are lower than the adjacent points are increased leading to a smoother signal. [b-Simonoff]

Sniffer

[GEO]: System mountable to any vehicle – it can be a drone – used in the recovery process of drifters after a hydrometeorological event and to obtain the collected data. [b-Roberts]

Snowpack

[GEO]: A mass of lying snow that is compressed and hardened by its own weight. Snowpack forms from layers of snow that accumulate in geographic regions and high altitudes where the climate includes cold weather for extended periods during the year. [ITU-T L.1503]

Social vulnerability index (SVI)

[G]: A web-based tool that allows users to examine the communities that may be most vulnerable to external stressors such as floods, forest fires, power outages, and winter storms. The SVI compiles 16 factors at the census tract level in four categories: socioeconomic status, household/disability, minority/language, and housing/transportation. [b-DHHS]

Spatial dependence

[AI/ST]: Spatial dependence is defined as the property of random variables taking values at pairs of locations a certain distance apart, which are more similar (positive autocorrelation) or less similar (negative autocorrelation) than expected for randomly associated pairs of observations. [b-Legendre]

Spatial interpolation

[AI/ST]: The process of using points with known values to estimate values at other unknown points. [b-MachineLearning]

Spatial resolution

[G]: The degree to which fine detail of an object can be reproduced in a radiographic, fluoroscopic, television, or other image. The smallest object or highest spatial frequency of a given contrast that is just perceptible. [b-ECGlossary]

[GEO]: In physics and geosciences, the term spatial resolution refers to the linear spacing of a measurement or the physical dimension that represents a pixel of the image. [b-Thomas]

Specificity

[AI/ST]: Specificity is defined as the proportion of actual negatives, which were predicted as the negative (or true negative). [b-DZone]

Speech-to-speech machine translation (SSMT)

[AI/ST]: Translating speech from one language to speech in another language. This can be done with a cascade of automatic speech recognition (ASR), text-to-text machine translation (MT), and text-to-speech (TTS) synthesis sub-systems, which is text-centric. Recently, works on speech-to-speech translation without relying on intermediate text representation is emerging. [ML-Learning]

Stacking

[AI/ST]: Using a machine learning model to learn how to best combine the predictions from contributing ensemble members. [b-Springer]

Static model

[G]: Models that represent a phenomenon at a given point in time or that compare a phenomenon at different points in time. [b-Springer-11]

Stick-slip

[G]: The stick-slip phenomenon, slip-stick phenomenon, or simply stick-slip, is the spontaneous jerking motion that can occur while two objects are sliding over each other. [b-Heslot]

[GEO]: The behavior of seismically-active faults can be explained using a stick-slip model, with earthquakes being generated during the periods of rapid slip. [b-Scholz]

Stochastic gradient descent (SGD)

[AI/ST]: A gradient descent algorithm in which the batch size is one. In other words, SGD relies on a single example chosen uniformly at random from a dataset to calculate an estimate of the gradient at each step. [b-DeepAI]

Structural similarity index measure (SSIM)

[G]: A method for predicting the perceived quality of digital television and cinematic pictures, as well as other kinds of digital images and videos. [b-Wang]

Subduction

[GEO]: In plate tectonics, subduction is the process of one lithospheric plate descending beneath another at active convergent continental margins. [b-Sendai]

Success rate curve

[AI/ST]: A curve based on the comparison between a prediction image and a landslide used in a modeling, which can help determine how well the resulting landslide susceptibility maps have classified the areas of existing landslides. [b-Deng]

Supervised learning

[AI/ST]: Learning strategy in which the correctness of acquired knowledge is tested through feedback from an external knowledge source. [b-ISO/IEC 2382]

Support vector machines (SVM)

[AI/ST]: SVMs are supervised learning models with associated learning algorithms that analyze data for classification and regression analysis. [ITU-T X.1235]

Sustainable development

[G]: Development that meets the environmental, social, and economic needs of the present without compromising the ability of future generations to meet their own needs. [ITU-T Y.4051]

Symmetric extreme dependency index (SEDI)

[G]: A skill score appropriate for extreme events. It provides meaningful results in the case of rare events where the hit rate and false alarm rate decrease towards zero. It is defined for a binary event and thus requires a threshold to be set. [b-ECMWFTerm]

Synthetic aperture radar (SAR)

[GEO]: A radar which produces an image by integrating data from successive antenna positions, thereby constructing an effective antenna aperture much larger than the "real" physical aperture. SAR payloads are carried on some Earth observation satellites and some planetary probes. [b-WMO-2]

Systematic errors

[G]: A systematic error is not determined by chance but is introduced by an inaccuracy (as of observation or measurement) inherent in the system. [b-ECMWFTerm]

***t*-distributed stochastic neighbor embedding (*t*-SNE)**

[AI/ST]: A statistical method for visualizing high-dimensional data by giving each datapoint a location in a two- or three-dimensional map. [b-Roweis]

Taylor diagrams

[G]: Taylor diagrams are mathematical diagrams designed to graphically indicate which of several approximate representations (or models) of a system, process, or phenomenon is most realistic. [b-Taylor]

Thermal imaging

[G]: 1. A very powerful remote sensing technique for a number of reasons, particularly when used to elucidate field studies relating to animal ecology. Thermal imaging data are collected at the speed of light in real time from a wide variety of platforms, including land, water, and air-based vehicles. [b-ScienceDirect-11]

2. Infrared thermography (IRT), thermal video, and/or thermal imaging, is a process whereby a thermal camera captures and creates an image of an object by using infrared radiation emitted from the object in a process, which are examples of infrared imaging science. [b-Hapke]

Time series forecasting

[AI/ST]: Time series forecasting occurs when you make scientific predictions based on historical time stamped data. It involves building models through historical analysis and using them to make observations and drive future strategic decision-making. [b-Time]

Topic modeling

[AI/ST]: An unsupervised machine learning technique that is capable of scanning a set of documents, detecting word and phrase patterns within them, and automatically clustering word groups and similar expressions that best characterize a set of documents. [b-Topic]

Total electron content (TEC)

[GEO]: The total number of electrons present along a path between a radio transmitter and receiver. Radio waves are affected by the presence of electrons. The more electrons in the path of the radio wave, the more the radio signal will be affected. For ground to satellite communication and satellite navigation, TEC is a good parameter to monitor for possible space weather impacts. [b-NOAA-2]

Transfer learning (TL)

[AI/ST]: A research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. [b-West]

Transferability

[AI/ST]: It is the ability to acquire and reuse knowledge in machine learning. [b-Jiang]

Traveling ionospheric disturbances (TIDs)

[GEO]: Quasi-periodic variations of ionospheric densities in the Earth's upper atmosphere, believed to be the ionospheric signatures of atmospheric gravity waves. [b-Hines]

Tsunami

[G]: A wave, or train of waves, produced by a disturbance such as a submarine earthquake displacing the sea floor, a landslide, a volcanic eruption, or an asteroid impact. [b-UNDRR]

Underfitting

[AI/ST]: Producing a model with poor predictive ability because the model has not captured the complexity of the training data. Many problems can cause underfitting, including: training on the wrong set of features, training for too few epochs, training at too low a learning rate, training with too high a regularization rate, or providing too few hidden layers in a deep neural network. [b-IBM-4]

UNet

[AI/ST]: A neural network with a U-shape, where connections exist between the horizontally corresponding layers of the contracting input branch and the expanding output branch. It was designed to work with fewer training images and to yield more precise segmentations. [b-Springer-1]

Univariate analysis

[AI/ST]: Univariate analysis is a common method for understanding data. It is the technique of comparing and analyzing the dependency of a single predictor and a response variable, for instance, the mean of a population distribution. [b-DeepAI]

Unmanned aerial vehicles (UAV)

[G]: A pilotless aircraft, in the sense of Article 8 of the Convention on International Civil Aviation, which is flown without a pilot-in-command on-board and is either remotely and fully controlled from another place (ground, another aircraft, space) or programmed and fully autonomous. [b-UAS]

Unsupervised learning

[AI/ST]: A learning strategy that consists of observing and analyzing different entities and determining that some of their subsets can be grouped into certain classes, without any correctness test being performed on acquired knowledge through feedback from external knowledge sources. [b-ISO/IEC 2382]

Validation

[G]: A process of checking a specification to ensure that it is syntactically and semantically correct and represents the intended behavior. [519]

[AI/ST]: A process used, as part of training, to evaluate the quality of a machine learning model using the validation set. Because the validation set is disjoint from the training set, validation helps ensure that the model's performance generalizes beyond the training set. [b-Wasserman]

Variance

[G]: In probability theory and statistics, variance is the expectation of the squared deviation of a random variable from its population mean or sample mean. Variance is a measure of dispersion, meaning it is a measure of how far a set of numbers is spread out from their average. [b-Wasserman]

Vector-borne disease

[G]: Vector-borne diseases are illnesses caused by parasites, viruses, and bacteria that are transmitted by vectors. Every year there are more than 700,000 deaths from diseases such as malaria, dengue, schistosomiasis, human African trypanosomiasis, leishmaniasis, Chagas disease, yellow fever, Japanese encephalitis, and onchocerciasis. [b-WHO]

Virus

[G]: A virus is a submicroscopic infectious agent that replicates only inside the living cells of an organism. [b-Wu]

Volcano

[G]: A vent or fissure in the Earth's surface from which lava and volatiles are extruded. [b-Cortés]

Volcano-independent seismic recognition (VI.VSR)

[G]: VI.VSR trains universal recognition models with data of several volcanoes to obtain portable and robust characteristics. [b-Cortés]

Weakly supervised learning

[AI/ST]: Weakly supervised learning is an umbrella term covering a variety of studies that attempt to construct predictive models by learning with weak supervision. [b-Zhou]

Weather station

[GEO]: A location where meteorological observations such as surface, upper air, and climatological observations are taken. [b-AMS]

Wildfires

[G]: A large, destructive fire that spreads quickly over woodland or brush. [b-AMS]

Wildland–urban interface

[G]: A transitional zone between unoccupied land and human development. The term has been particularly used to refer to any area where human-made improvements have been built close to, or within, natural terrain and flammable vegetation, and where there is a great potential for wildfires. [b-AMS]

Windstorm

[G]: A storm marked by high wind with little or no precipitation. [b-UNDRR]

Wireless sensor network (WSN)

[G]: Wireless sensor networks refer to networks of spatially dispersed and dedicated sensors that monitor and record the physical conditions of the environment and forward the collected data to a central location. [b-Ullo]

WSG84 projection

[G]: The WGS defines a reference frame for the Earth, for use in geodesy and navigation. The latest revision is WGS 84 dating from 1984 (last revised in 2004), which will be valid up to about 2010. [b-WGS]

eXtreme Gradient Boosting (XGBoost)

[AI/ST]: An open-source software library which provides a regularizing gradient boosting framework for C++, Java, Python, R, Julia, Perl, and Scala. [b-eXtreme]

Bibliography

[b-AcademicPress] Barry M. Horowitz, *Chapter 5 – Policy Issues Regarding Implementations of Cyber Attack: Resilience Solutions for Cyber Physical Systems*, Editor(s): William Lawless, Ranjeev Mittu, Donald Sofge, Ira S. Moskowitz, Stephen Russell, Artificial Intelligence for the Internet of Everything, Academic Press, 2019, Pages 87-100, ISBN 9780128176368 https://doi.org/10.1016/B978-0-12-817636-8.00005-3.

[b-Adamopoulou] Adamopoulou, E., Moussiades, L. (2020), *Chatbots: History, technology, and applications*. https://doi.org/10.1016/j.mlwa.2020.100006

[b-Addelman] Addelman S. (1969), *The Generalized Randomized Block Design. The American Statistician*, Volume 4.

[b-Agile] What makes a good benchmark dataset? https://agilescientific.com/blog/2019/4/3/what-makes-a-good-benchmark-dataset

[b-aigoogle] Training Machine Learning Models More Efficiently with Dataset Distillation: https://ai.googleblog.com/2021/12/training-machine-learning-models-more.html

[b-AMS] American Meteorological Society. Glossary of Meteorology: https://glossary.ametsoc.org/wiki/Welcome

[b-Analytics] All You Need To Know About Different Types Of Missing Data Values And How To Handle It: https://www.analyticsvidhya.com/blog/2021/10/handling-missing-value/

[b-Andrews] Applied Statistics – Lesson 5: Correlation Coefficient. Andrews University.

[b-Appen] What is Data Annotation? https://appen.com/blog/data-annotation/

[b-ARCADIS] A Floods Working Group (CIS) (2016). Resource document Flood Risk Management, Economics and Decision Making Support, ARCADIS: http://ec.europa.eu/environment/water/flood_risk/pdf/WGF_Resource_doc.pdf

[b-Arizona] Transformations of Random Variables (2009). University of Arizona

[b-aspexit] Spatial data interpolation: TIN, IDW, kriging, block kriging, co-kriging. What are the differences? https://www.aspexit.com/spatial-data-interpolation-tin-idw-kriging-block-kriging-co-kriging-what-are-the-differences/

[b-AWS] What Is a Key-Value Database? AWS: https://aws.amazon.com/nosql/key-value/?nc1=h_ls

[b-Bach] Bach S, Binder A, Montavon G, Klauschen F, Müller K-R, Samek W (2015), *On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation*.

[b-Bath] Bath G. (2012), *The Software Test Engineer's Handbook: A Study Guide for the ISTQB Test Analyst and Technical Analyst Advanced Level Certificates.*

[b-Baeldung] Latent Space in Deep Learning. Baeldung: https://www.baeldung.com/cs/dl-latent-space

[b-BDEX] What Is Data Completeness? Why One Missing Piece Matters: https://www.bdex.com/blog/why-is-data-completeness-important/

[b-BigData] Data Consistency Theory and Case Study for Scientific Big Data (2019). University of Science and Technology Beijing, Beijing.

[b-Binder] Binder Alexander, Grégoire Montavon, Sebastian Bach, Klaus-Robert Müller, Wojciech Samek (2016), *Layer-wise Relevance Propagation for Neural Networks with Local Renormalization Layers*. Cornell University.

[b-Böse] Böse, Erdik M., Wenzel1 F. (2007), *A New Approach to Earthquake Early Warning*. Earthquake Early Warning Systems, 65–83.

[b-Breuel] Breuel, Cristiano (2022), *ML Ops: Machine Learning as an Engineering Discipline*. Towards Data Science. Retrieved 01 April. https://www.synopsys.com/glossary/what-is-devops.html

[b-Bristol] Point estimates and population parameters. University of Bristol

[b-Britannica] Britannica Dictionary: https://www.britannica.com/

[b-Brust] Brust, Clemens-Alexander & Sickert, Sven & Simon, Marcel & Rodner, Erik & Denzler, Joachim. (2016), *Evaluation of QuickProp for Learning Deep Neural Networks – A Critical Review*.

[b-Bunge] Bunge M. (2012), *Causality and Modern Science*.

[b-Business] Business Dictionary: http://www.businessdictionary.com/definition/manufacturing-lead-time.html

[b-Cambridge] Cambridge Dictionary: https://dictionary.cambridge.org/

[b-CambridgeDictionary-1] Uncertainty: https://dictionary.cambridge.org/dictionary/english/uncertainty

[b-Carbonplace] Carbonplace EU: http://www.carbonplace.eu/info-rss

[b-Chang] Chang, Kang-tsung (2016), *Introduction to Geographic Information Systems* (9th ed.). McGraw-Hill. p. 2. ISBN 978-1-259-92964-9

[b-CheatSheet] Glossary.CheatSheet: https://ml-cheatsheet.readthedocs.io/en/latest/glossary.html

[b-Chen] Chen, Yangkang, Fomel, Sergey (November–December 2015), *Random noise attenuation using local signal-and-noise orthogonalization*. Geophysics. 80 (6)

[b-Cheng] Cheng B., Stanley R. (2012), *Data fusion by using machine learning and computational intelligence techniques for medical image analysis and classification*. Missouri University of Science and Technology: https://dl.acm.org/doi/book/10.5555/2520640

[b-Chesalov] Chesalov A., Vlaskin A., Bakanach M. (2022). *Artificial Intelligence Glossarium: 1000 terms*.

[b-Clapham] Clapham, Christopher; Nicholson, James (2014), Oxford Concise Dictionary of Mathematics (Fifth Edition.). Oxford University Press

[b-Climate] Climate Prediction Net: https://www.climateprediction.net/climate-science/glossary/mid-latitudes/

[b-ClimateData] Climate Data Store https://cds.climate.copernicus.eu/#!/home

[b-Collins] Collins Dictionary: https://www.collinsdictionary.com/de/

[b-Columbia] Deep Thoughts on Deep Convection. Columbia Climate School: https://news.climate.columbia.edu/2009/03/01/deep-thoughts-on-deep-convection/

[b-Computer] Neither Quick Nor Proper (2016), *Evaluation of QuickProp for Learning Deep Neural Networks*. Computer Vision Group.

[b-Computer] Real-time constraints in a rapid prototyping language. Computer Languages (1993). Volume 18, Issue 2, 77-103.

[b-Convolutional] What is Transposed Convolutional Layer? https://towardsdatascience.com/what-is-transposed-convolutional-layer-40e5e6e31c11

[b-Copernicus] Copernicus Global Land Service (CGLS): https://land.copernicus.eu/global/products/lc

[b-Copernicus] Copernicus Open Access Hub https://scihub.copernicus.eu/

[b-Copernicus] Data and Information Access Services: https://www.copernicus.eu/en/access-data/dias

[b-Cornell] G-softmax: Improving Intra-class Compactness and Inter-class Separability of Features. Cornell University https://arxiv.org/abs/1904.04317

[b-CornellUniversity] Transfer learning for self-supervised, blind-spot seismic denoising. Cornell University: https://arxiv.org/abs/2209.12210

[b-Cortés] Cortés, G.R., Carniel, P. Lesage, M.Á. Mendoza, and I. Della Lucia, 2021, *Practical Volcano-Independent Recognition of Seismic Events: VULCAN.ears Project*. Front. Earth Sci., 8, 616676.

[b-Dalenius] Dalenius, Tore (1977), *Towards a methodology for statistical disclosure control.*

[b-Cramer-1] Cramer, J.S. (1986), *Econometric Applications of Maximum Likelihood Methods*. New York, NY: Cambridge University Press.

[b-Cramer-2] Cramer EM, and Bock RD (1966), *Multivariate Analysis. Review of Educational Research*. Dec., 36(5):604-617. doi:10.2307/1169484. Accessed 20220913.

[b-CRC] i, Z., Zhu, Q. and Gold, C. (2005), *Digital terrain modeling: principles and methodology*, CRC Press.

[b-Cyber] J.-S. R. Jang (1993), *ANFIS: adaptive-network-based fuzzy inference system*, in IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, no. 3, pp. 665-685, May-June, doi: 10.1109/21.256541.

[b-Czabanski] Czabanski. R, Jezewski. M, Leski.J (2017). *Theory and Applications of Ordered Fuzzy Numbers* pp 23–43.

[b-Data] Machine Learning Classifiers https://towardsdatascience.com/machine-learning-classifiers-a5cc4e1b0623

[b-Data4SDGs] Data Interoperability Collaborative https://www.data4sdgs.org/initiatives/data-interoperability-collaborative

[b-DataAnalysis] Interpreting Exploratory Data Analysis (EDA) https://www.datasciencecentral.com/interpreting-exploratory-data-analysis-eda/

[b-DataDrift] Why data drift detection is important and how do you automate it in 5 simple steps: https://towardsdatascience.com/why-data-drift-detection-is-important-and-how-do-you-automate-it-in-5-simple-steps-96d611095d93

[b-DataRobot] Machine Learning Life Cycle. DataRobot: https://www.datarobot.com/wiki/machine-learning-life-cycle/

[b-DataS] Introduction to Probabilistic Graphical Models: https://towardsdatascience.com/introduction-to-probabilistic-graphical-models-b8e0bf459812

[b-DataScience] What is Coarse-to-Fine in the context of neural networks? https://datascience.stackexchange.com/questions/47921/what-is-coarse-to-fine-in-the-context-of-neural-networks

[b-DataSupply] Learning from Machines: The Data Supply Chain: https://towardsdatascience.com/learning-from-machines-the-data-supply-chain-4380f420bb2c

[b-Datatron] What is Model Drift. Datatron: https://datatron.com/what-is-model-drift/

[b-DeepAI] What is Association Learning? Deep AI https://deepai.org/machine-learning-glossary-and-terms/association-learning

[b-Deepchecks] DeepCheck Glossary: https://deepchecks.com/

[b-Deng] Deng, X., L. Li, and Y. Tan (2017), *Validation of Spatial Prediction Models for Landslide Susceptibility Mapping by Considering Structural Similarity*. IJGI, 6, 103, https://doi.org/10.3390/ijgi6040103.

[b-DHHS] Environmental Public Health Tracking. Department of Health and Human Services: https://www.nh.gov/epht/highlights/documents/social-vulnerability-index.pdf

[b-Dialogic] Telecommunications Service Provider (TSP): https://www.dialogic.com/glossary/telecommunications-service-provider-tsp

[b-Dictionary] A Dictionary of Statistical Terms (1982).

[b-DoD] Undersecretary of Defense for Acquisition Technologies (1998). DoD Modelling and Simulation Glossary.

[b-Dodge] Dodge, Y (2003), *The Oxford Dictionary of Statistical Terms*, OUP. ISBN 0-19-920613-9 (entry for normalization of scores)

[b-Douglas] Douglas J., *GMPE compendium*: http://www.gmpe.org.uk/

[b-Draper] Draper, J. (1980), *The Direction of Desert Locust Migration*. Journal of Animal Ecology. 49 (3): 959–974.

[b-Draper] Draper, N. R.; Smith, H. (1998), *Applied Regression Analysis*. Wiley-Interscience.

[b-DZone] ML Metrics: Sensitivity vs. Specificity.DZone: https://dzone.com/articles/ml-metrics-sensitivity-vs-specificity-difference

[b-Earth] Google Earth Engine: https://earthengine.google.com/

[b-Earth] Robustness Metrics: How Are They Calculated, When Should They Be Used and Why Do They Give Different Results? Earth's Future: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017EF000649

[b-EC] European Commission (2019), Independent High-Level Expert Group on Artificial Intelligence. *Ethics Guidelines for Trustworthy AI*.

[b-ECGlossary] EC Glossary-Quality assurance and radiation protection in diagnostic radiology.

[b-ECMWF] ECMWF – Forecast Charts and Data. https://www.ecmwf.int/en/forecasts

[b-ECMWFTerm] Terminology: https://confluence.ecmwf.int/display/COPSRV/Terminology

[b-EEA] European Environment Agency Glossary: https://www.eea.europa.eu/help/glossary#c4=10&c0=all&b_start=0

[b-Elith] Elith, Jane; Leathwick, John R. (2009), *Species Distribution Models: Ecological Explanation and Prediction Across Space and Time*. Annual Review of Ecology, Evolution, and Systematics. 40 (1): 677–697.

[b-Elmenreich] Elmenreich, W. (2002), *Sensor Fusion in Time-Triggered Systems*. Vienna, Austria: Vienna University of Technology. p. 173.

[b-Encoder] What is an encoder decoder model? https://towardsdatascience.com/what-is-an-encoder-decoder-model-86b3d57c5e1a

[b-Encyclopedia] Encyclopedia of Machine Learning, 713–714.

[b-ESA] New generation of sensors. European Space Agency https://www.esa.int/Applications/Observing_the_Earth/FutureEO/Swarm/New_generation_of_sensors

[b-Essnet] Essnet Validat Foundation (2016), *Methodology for data validation 1.0*.

[b-ETH] Earthquake Early Warning. ETH Zurich: http://seismo.ethz.ch/en/research-and-teaching/fields_of_research/earthquake-early-warning

[b-EU] Regulation (EU) 2016/679 of the European Parliament and of the Council.

[b-EUParliament] Proposal for a regulation on the deployment of alternative fuels infrastructure, and repealing Directive 2014/94/EU of the European Parliament and of the Council.

[b-EUTerm] European Union Terminology https://iate.europa.eu/entry/result/1756736

[b-Evision] F.F. Evison, Seismogenesis, Tectonophysics, Volume 9, Issues 2–3, 1970, Pages 113-128, ISSN 0040-1951.

[b-eXtreme] GitHub – Awesome XGBoost: https://github.com/dmlc/xgboost/tree/master/demo#machine-learning-challenge-winning-solutions

[b-Fan] Fan C, Karati A., Yang S. (2021), *Reliable file transfer protocol with producer anonymity for Named Data Networking*.

[b-Fault] The Properties of Fault Slip: https://web.archive.org/web/20100625133126/http://www.data.scec.org/Module/sec1pg14.html

[b-Fausto] Fausto Guzzetti, Alessandro Cesare Mondini, Mauro Cardinali, Federica Fiorucci, Michele Santangelo, Kang-Tsung Chang (2012), *Landslide inventory maps: New tools for an old problem*, Earth-Science Reviews 112, 42-66.

[b-Forbes] On The Semantics Of Data Bias: Reducing Bias Versus Creating Inclusive AI. Forbes https://www.forbes.com/sites/forbescommunicationscouncil/2021/06/21/on-the-semantics-of-data-bias-reducing-bias-versus-creating-inclusive-ai/

[b-Ford] Terminology of Forest Science, Technology, Practice, and Products'. F.C. Ford-Robertson (ed.). Society of American Foresters. 1971

[b-Forman] Forman, Ernest H., Saul I. Gass (2001), *The analytical hierarchy process – an exposition*. Operations Research. 49 (4): 469–487. doi:10.1287/opre.49.4.469.11231

[b-FreeDictionary] The Free Dictionary: https://www.thefreedictionary.com/exportability

[b-Freedman] Freedman D., Pisani R., Purves R. (2007), *Statistics – Descriptive Statistics*.

[b-Frontiers] Probabilistic Enhancement of the Failure Forecast Method Using a Stochastic Differential Equation and Application to Volcanic Eruption Forecasts (2019), https://www.frontiersin.org/articles/10.3389/feart.2019.00135/full

[b-Fukuda] Fukuda, Komei, Terlaky, Tamás (1997), Thomas M. Liebling; Dominique de Werra (eds.), *Criss-cross methods: A fresh view on pivot algorithms*. Mathematical Programming, Series B. 79 (1–3): 369–395.

[b-Gaussian] Gaussian Naive Bayes: https://iq.opengenus.org/gaussian-naive-bayes/

[b-GDACS] Global Disaster Alert and Coordination System: https://gdacs.org/

[b-Geo] Cotton,F., Coutant, O. (1997), *Dynamic stress variations due to shear faults in a plane-layered medium*. Geophysical Journal International, Volume 128, Issue 3, March, Pages 676–688.

[b-Geology] Earthquakes and the Seismic Cycle: http://www.geology.wisc.edu/homepages/chuck/public_html/Classes/Mtn_and_Plates/eq_cycle.html

[b-GEONADIR] What is an Orthomosaic?: GEONADIR: https://geonadir.com/what-is-an-orthomosaic/

[b-GeoScience] GeoScienceWorld. GSA Bulletin. Volume 113, Number 6, June 2001, Notes on geochronologic and chronostratigraphic units (5.6.2020) and A.G.I. & A.G.I. Supp. = American Geological Institute. Glossary of Geology and Related Sciences. Washington, DC., 1957m 325 pp; supplement, 1960, 72 pp.

[b-GFDRR] Knowledge Note 5-1 cluster 5: Hazard and Risk Information and Decision Making.

[b-GHS] Global Health Security Index: https://www.ghsindex.org/

[b-GIGAS] GIGAS (2008), *Data Harmonization – ICT for environmental management and energy efficiency.*

[b-Github] Darknet: https://github.com/pjreddie/darknet

[b-Github] ML-Exercise-04-Spam-Classifier.GitHub.

[b-Glantz] Glantz, Stanton A., Slinker, B.K. (1990), *Primer of Applied Regression and Analysis of Variance*. McGraw-Hill.

[b-GPS] The Global Positioning System: https://www.gps.gov/systems/gps/

[b-Greenacre] Greenacre M., Ayhan H.O., *Identifying inliers*. Universitat Pompeu Fabra.

[b-Guide] A beginner's guide to dimensionality reduction in Machine Learning.

[b-Hahn] Hahn G., Manabe S. (1975), *The Role of Mountains in the South Asian Monsoon Circulation*. Journal of Atmospheric Sciences.

[b-Handbook] Handbook of Statistics Volume 31 (2013).

[b-Hapke] Hapke B. (2012), *Theory of Reflectance and Emittance Spectroscopy*. Cambridge University Press.

[b-Hariri] Hariri R., Fredericks E.M, Bower K. (2019), *Uncertainty in big data analytics: survey, opportunities, and challenges*.

[b-HDSR] Rudin C., Radin J. (2019), *Why Are We Using Black Box Models in AI When We Don't Need To*? A Lesson From an Explainable AI Competition. Harvard Data Science Review.

[b-Hellström] Hellström, T., Dignum, V., & Bensch, S. (2020), *Bias in Machine Learning – What is it Good for*? arXiv preprint arXiv:2004.00686

[b-Hernández-Orallo] Hernández-Orallo, J. (2017), *Evaluation in artificial intelligence: from task-oriented to ability-oriented measurement*. Artif Intell Rev 48, 397-447.

[b-Heslot] F. Heslot, T. Baumberger, B. Perrin, B. Caroli, and C. Caroli, Phys. Rev. E 49, 4973 (1994), Sliding Friction: Physical Principles and Applications – Bo N.J. Persson Ruina, Andy. *Slip instability and state variable friction laws*. Journal of Geophysical Research 88.B12 (1983): 10359-10.

[b-Hines] Hines, C.O. (1960), *Internal atmospheric gravity waves at ionospheric heights*. Can. J. Phys., 38, 1441–1481.

[b-Hinton] Hinton G (2009), *Deep belief networks*. Scholarpedia. 4 (5): 5947.

[b-IBE] El Niño-Southern Oscillation (ENSO). IBE Climate Services: https://drought.climateservices.it/en/glossary/el-nino-southern-oscillation-enso/

[b-IBM] What is the k-nearest neighbors algorithm? IBM: https://www.ibm.com/topics/knn

[b-IBM-1] Data aggregation. IBM. https://www.ibm.com/docs/da/tnpm/1.4.2?topic=data-aggregation

[b-IBM-2] IBM (2021). Learning the Parameters of Bayesian Networks from Uncertain Data

[b-IBM-3] What is a digital twin? IBM: https://www.ibm.com/topics/what-is-a-digital-twin

[b-IBM-4] What is underfitting? IBM: https://www.ibm.com/topics/underfitting

[b-ICAO] International Civil Aviation Organization (ICAO): https://www.icao.int/safety/UA/Documents/ICAO%20RPAS%20CONOPS.pdf

[b-ICID] ICID – International Commission on Irrigation and Drainage

[b-IEEE] 5G Testbed. IEEE Future Networks: https://futurenetworks.ieee.org/topics/5g-testbed

[b-IEEE] A Review of Uncertainty Quantification in Deep Learning: Techniques, Applications and Challenges.

[b-IFRC] Hailstorms. IFRC https://www.ifrc.org/our-work/disasters-climate-and-crises/what-disaster/hailstorms

[b-IGIGlobal] What is Geospatial Mapping. IGI Global: https://www.igi-global.com/dictionary/geospatial-influence-in-science-mapping/60106

[b-Immuta] DataOps Dilemma: Survey Reveals Gap in the Data Supply Chain. Immuta: https://www.immuta.com/downloads/dataops-dilemma-survey-reveals-gap-in-the-data-supply-chain/

[b-Informatica] What is Data Validation? Informatica https://www.informatica.com/services-and-training/glossary-of-terms/data-validation-definition.html

[b-Interpretability] Interpretability. GitHub https://christophm.github.io/interpretable-ml-book/interpretability.html

[b-ISO 1213-1:2020] ISO 1213-1:2020, *Coal and coke – Vocabulary – Part 1: Terms relating to coal preparation*.

[b-ISO 16100-2] ISO 16100-2, *Industrial automation systems and integration – Manufacturing software capability profiling for interoperability – Part 2: Profiling methodology*.

[b-ISO/IEC 20546] ISO/IEC 20546, *Information technology – Big data – Overview and vocabulary*.

[b-ISO/IEC 20922] ISO/IEC 20922, *Information technology – Message Queuing Telemetry Transport (MQTT) v3.1.1*.

[b-ISO/IEC 22989] ISO/IEC 22989, *Information technology – Artificial intelligence – Artificial intelligence concepts and terminology*: https://www.iso.org/obp/ui/#iso:std:iso-iec:22989:ed-1:v1:en

[b-ISO/IEC 2382] ISO/IEC 2382, *Information technology – Vocabulary*.

[b-ISO/IEC 2382-28] ISO/IEC 2382-28, *Information technology – Vocabulary – Part 28: Artificial intelligence – Basic concepts and expert systems*.

[b-ISO/IEC 24727] ISO/IEC 24727, *Identification cards – Integrated circuit card programming interfaces – Part 4: Application programming interface (API) administration*.

[b-ISO/IEC TR 29119-11] ISO/IEC TR 29119-11, *Software and systems engineering – Software testing – Part 11: Guidelines on the testing of AI-based systems*.

[b-ISO/IEC TS 22424-1] ISO/IEC TS 22424-1, *Digital publishing – EPUB3 preservation – Part 1: Principles*.

[b-ISO/TS 23029:2020] ISO/TS 23029:2020, *Web-service-based application programming interface (WAPI) in financial services*.

[b-ISO] *Life cycle perspective*. ISO: https://committee.iso.org/sites/tc207sc1/home/projects/published/iso-14001---environmental-manage/life-cycle.html

[b-ISO 22327] ISO 22327, *Security and resilience – Emergency management – Guidelines for implementation of a community-based landslide early warning system*.

[b-ISO/IEC 27729] ISO/IEC 27729, *Information and documentation – International standard name identifier (ISNI).*

[b-ITU-T L.94] Recommendation ITU-T L.94, *Use of global navigation satellite systems to create a referenced network map*.

[b-Java] Associate Rule Learning:JavatPoint https://www.javatpoint.com/association-rule-learning

[b-Java] Numerosity Reduction in Data Mining. Java Point: https://www.javatpoint.com/numerosity-reduction-in-data-mining

[b-JGR] S.E. Minson, Murray, R., Langbein, R., Gomberg, J. (2014), *Real-time inversions for finite fault slip models and rupture geometry based on high-rate GPS data*. JGR Solid Earth https://doi.org/10.1002/2013JB010622

[b-Jiang] Jiang, J., Y. Shu, J. Wang, and M. Long (2022), *Transferability in Deep Learning: A Survey*. https://doi.org/10.48550/ARXIV.2201.05867 .

[b-Jobin] Jobin, A., Ienca, M. & Vayena (2019), *E. The global landscape of AI ethics guidelines*. Nat Mach Intell 1, 389–399. https://doi.org/10.1038/s42256-019-0088-2

[b-Jöreskog] Jöreskog, Karl G. (1983), *Factor Analysis as an Errors-in-Variables Model*. Principals of Modern Psychological Measurement. Hillsdale: Erlbaum. pp. 185–196.

[b-Kaczmarczyk] Kaczmarczyk. K, Miałkowska. K (2022), *Backtesting comparison of machine learning algorithms with different random seed*. Procedia Computer Science, Volume 207

[b-Kaiming] Kaiming Zhang, Xiangyu Ren, Shaoqing Sun, Jian (2016), *Deep Residual Learning for Image Recognition*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE. pp. 770–778. arXiv:1512.03385.

[b-Khurana] Khurana, D., Koli, A., Khatter, K., Singh, S. (2022), *Natural language processing: state of the art, current trends and challenges Multimedia Tools and Applications*.

[b-Klein] Klein A. (2004), *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making*.

[b-Konečný] Konečný, Jakub, McMahan, Brendan, Ramage, Daniel (2015), *Federated Optimization: Distributed Optimization Beyond the Datacenter*.

[b-Laboratory] Laboratory Data Completion Status https://www.tititudorancea.com/z/laboratory_data_completion_status.htm

[b-Legendre] Legendre, P., L. Legendre, L. Legendre, and L. Legendre (1998), *Numerical ecology*. 2nd English ed. Elsevier, 853.

[b-Leslie] Leslie (2019), *Understanding artificial intelligence ethics and safety – A guide for the responsible design and implementation of AI systems in the public sector*, 2019, Turing Institute, UK,

[b-Leventhall] Leventhall G. (2007), *What is infrasound?* Progress in Biophysics and Molecular Biology. Vol 93.

[b-Licciardi] Licciardi, A., Bletery, Q., Rouet-Leduc, B. et al. (2022), *Instantaneous tracking of earthquake growth with elastogravity signals*. Nature 606, 319-324.

[b-Lobo] Lobo, Jorge M. et al. (2008), *AUC: a misleading measure of the performance of predictive distribution models*. Global Ecology and Biogeography 17: 145-151.

[b-London] Council-EN, based on: The Principles of Good Data Management, Intra-Governmental Group on Geographic Information, Office of the Deputy Prime Minister: London, July 2005 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/14867/Good_dataMan.pdf

[b-MachineLearning] Machine Learning Glossary: https://developers.google.com/machine-learning/glossary

[b-MAPE] MAPE (Mean Absolute Percentage Error): https://docs.oracle.com/en/cloud/saas/planning-budgeting-cloud/pfusu/insights_metrics_MAPE.html

[b-Matheron] Matheron, G. (1963), *Principles of geostatistics*. Economic Geology; 58 (8): 1246–1266. doi: https://doi.org/10.2113/gsecongeo.58.8.1246

[b-Maydeu-Olivares] Maydeu-Olivares. A, Forero. C (2010), *Goodness of fit testing. International Encyclopedia of Education* (pp. 190-196)

[b-McGraw-Hill] Skolnik M. (1990), *Radar Handbook*. McGraw-Hill.

[b-McLachlan] McLachlan, G.J. (2004), *Discriminant Analysis and Statistical Pattern Recognition*. Wiley Interscience.

[b-Medium] What Is Grid Search? https://medium.com/fintechexplained/what-is-grid-search-c01fe886ef0a

[b-Medium] What on earth is machine learning? https://medium.com/tekton-labs/what-on-earth-is-machine-learning-9f9894343cb1

[b-MEMS] An introduction to micro electro mechanical systems (MEMS)

[b-Michael] Michael R. Forrest. Slow Earthquakes https://www.scec.org/

[b-Miller] Miller T. (2019), *Explanation in artificial intelligence: Insights from the social sciences*. Artif. Intell. 267:1–38. doi: 10.1016/j.artint.2018.07.007

[b-Misra] Misra S., He. J. (2020), *Chapter 4 – Stacked neural network architecture to model the multifrequency conductivity/permittivity responses of subsurface shale formations. Machine Learning for Subsurface Characterization.*

[b-MIT] Massachusetts Institute of Technology. Introduction to Machine-Learning.

[b-MIT] Prat, N., Madnick, S. (2007), *Measuring Data Believability: A Provenance Approach*. Massachusetts Institute of Technology (MIT).

[b-MLGlossary] ML Glossary: https://ml cheatsheet.readthedocs.io/en/latest/glossary.html

[b-MLMastery-1] A Gentle Introduction to Sparse Matrices for Machine Learning Machine Learning Mastery.

[b-MLMastery-2] Train-Test Split for Evaluating Machine Learning Algorithms.

[b-MLMastery-3] Use Early Stopping to Halt the Training of Neural Networks At the Right Time: https://machinelearningmastery.com/how-to-stop-training-deep-neural-networks-at-the-right-time-using-early-stopping/

[MLMastery-4] How to Calculate Precision, Recall, and F-Measure for Imbalanced Classification: https://machinelearningmastery.com/precision-recall-and-f-measure-for-imbalanced-classification/#:~:text=Precision%20quantifies%20the%20number%20of,and%20recall%20in%20one%20number .

[b-MLMastery-5] How to Choose a Feature Selection Method For Machine Learning: https://machinelearningmastery.com/feature-selection-with-real-and-categorical-data/

[b-MLMastery-6] A Gentle Introduction to k-fold Cross-Validation. Machine Learning Mastery: https://machinelearningmastery.com/k-fold-cross-validation/

[b-MLMastery-7] Loss and Loss Functions for Training Deep Learning Neural Networks.

[b-MLMastery-8] What Is Meta-Learning in Machine Learning?: https://machinelearningmastery.com/meta-learning-in-machine-learning/

[b-Mocus] Mocus J. (1989), *Bayesian Approach to Global Optimization, Mathematics and its Applications*. Soviet Series Band 37

[b-Molecular] Module in Chemistry, Molecular Sciences and Chemical Engineering, 2018: https://www.sciencedirect.com/topics/chemical-engineering/fuzzy-rules

[b-Moon] J. Moon, W.K. Lee, C. Song, S.G. Lee, S.B. Heo, A. Shvidenko, F. Kraxner, M. Lamchin, E.J. Lee, Y. Zhu, D. Kim, G. Cui (2016), *An introduction to mid-latitude ecotone: sustainability and environmental challenges*. СИБИРСКИЙ ЛЕСНОЙ ЖУРНАЛ

[b-Moral] Moral-Gracia S., Castellano J., Mantas C., Montella A., Abellán J. (2019), *Decision Tree Ensemble Method for Analyzing Traffic Accidents of Novice Drivers in Urban Areas*. Entropy, 21(4), 360.

[b-Munjiza] A. Munjiza, D.R.J. Owen, N. (1995), *Bicanic A combined finite–discrete element method in transient dynamics of fracturing solids*. Int. J. Engng. Comput., 12, pp. 145-174.

[b-Murdoch] Murdoch, W.J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019), *Definitions, methods, and applications in interpretable machine learning*. Proceedings of the National Academy of Sciences, 116(44), 22071-22080.

[b-Naaman] Naaman, M. (2021), *On the tight constant in the multivariate Dvoretzky–Kiefer–Wolfowitz inequality*.

[b-NAP] National Academies of Sciences, Engineering, and Medicine; Policy and Global Affairs; Committee on Science, Engineering, Medicine, and Public Policy; Board on Research Data and Information; Division on Engineering and Physical Sciences; Committee on Applied and Theoretical Statistics; Board on Mathematical Sciences and Analytics; Division on Earth and Life Studies; Nuclear and Radiation Studies

Board; Division of Behavioral and Social Sciences and Education; Committee on National Statistics; Board on Behavioral, Cognitive, and Sensory Sciences; Committee on Reproducibility and Replicability in Science. Reproducibility and Replicability in Science. Washington (DC): National Academies Press (US); 2019 May 7. PMID: 31596559.

[b-NASA] Thoughts on Producing "Ancillary Data" in support of Space Science Flight Projects.

[b-NASA] Shannon Diversity: Richness and Evenness. NASA: https://modis.ornl.gov/documentation/shannon.html

[b-NASAJPL] NASA Jet Propulsion Laboratory – Overview.

[b-National] The Unpredictable Certainty (1997). National Academy Press.

[b-NaturalHazards] A framework for spatio-temporal scales and concepts from different disciplines: The 'vulnerability cube. Natural Hazards 68(3).

[b-Nature-1] Numerical simulations. Nature Portfolio: https://www.nature.com/subjects/numerical-simulations#:~:text=A%20numerical%20simulation%20is%20a,as%20in%20most%20nonlinear%20systems

[b-Nature-2] Karniadakis, G.E., Kevrekidis, I.G., Lu, L. et al. (2021), *Physics-informed machine learning*. Nat Rev Phys 3, 422–440

[b-NECI] Numerical Weather Prediction. National Centers for Environmental Information: https://www.ncei.noaa.gov/products/weather-climate-models/numerical-weather-prediction

[b-NerveNet] Nervenet: Learning Structured Policy With Graph Neural Networks.

[b-netidee] Constrained Devices: What are the challenges when you work with constrained devices. Netidee: https://www.netidee.at/comatrix/constrained-devices

[b-NIST] NIST Big Data Public Working Group. Draft of Big Data Definition.

[b-Nkansah] Nkansah, P., Amankwaa, A., *Adequacy, Accessibility, and Goodness of Data*. Journal of African Business 2(1):95-107.

[b-NOAA-1] Glossary – NOAA's National Weather Service.

[b-NOAA-2] Total Electron Content. National Oceanic and Atmospheric Administration (NOAA): https://www.swpc.noaa.gov/phenomena/total-electron-content

[b-NOAAStudent] Students Glossary. National Oceanic and Atmospheric Administration: https://www.adp.noaa.gov/Students/Glossary.aspx

[b-Numpy] numpy.random.shuffle: https://numpy.org/doc/stable/reference/random/generated/numpy.random.shuffle.html

[b-NWS] United States. National Weather Service. Shallow/Deep Convection: https://www.weather.gov/

[b-Oxford] The Oxford Essential Dictionary of the U.S. Military (2001). Berkley Books, Oxford Reference Online. Oxford University Press. DGT.

[OASIS] OASIS (2022). Common Alerting Protocol, v. 1.1

[b-PennState] Analysis of Discrete Data: https://online.stat.psu.edu/statprogram/stat504

[b-PersonalData] Deletion of personal data – FAQ's: https://www.datenschutz-notizen.de/deletion-of-personal-data-faqs-3328108/

[b-Piedmont] Piedmont, R. Bias Statistical. Encyclopedia of Quality of Life and Well-Being Research pp 382–383

[b-PNSN] Hazard Maps. Pacific Northwest Seismic Network: https://pnsn.org/outreach/hazard-maps-and-scenarios/hazard-maps

[b-Powers] Powers, David M W (2011), *Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation* (PDF). Journal of Machine Learning Technologies. 2 (1): 37–63.

[b-Probability] Probability calibration (2014): https://jmetzen.github.io/2015-04-14/calibration.html

[b-Proceedings] Recent Advances in Structural Integrity Analysis (2014), Proceedings of the International Congress (APCF/SIF-2014).

[b-pyts] Gramian Angular Field. A Python Package for Time Series Classification: https://pyts.readthedocs.io/en/0.10.0/auto_examples/image/plot_gaf.html

[b-Quinlan] Quinlan, J. Ross (1986), *Induction of decision trees*. Machine learning 1.1: 81-106.

[b-Radio] Radio Regulations (2004) – Art. 1 § 1.100.

[b-Redman] Redman, Thomas C. (30 December 2013). Data Driven: Profiting from Your Most Important Business Asset. Harvard Business Press.

[b-Reinforcement] Exploration versus exploitation in reinforcement learning: a stochastic control approach. Cornell University: https://www.manifold.ai/exploration-vs-exploitation-in-reinforcement-learning

[b-Reinhardt] Reinhardt, J. (1985), *Analysis of Approximation Methods for Differential and Integral Equations*.

[b-Remi] Definition reference: Council-EN, based on: Official blog of Remi A.I – A walkthrough on A.I Clustering.

[b-RMSE] Root mean square error (RMSE): https://ec.europa.eu/eurostat/cros/content/root-mean-square-error-rmse_en

[b-Roberts] Robert Edwards, Victor Rangel, Ismael Perez, Omar Álvarez-Cárdena. Flash Flood Early Warning System in Colima, Mexico. September 2020. DOI: 10.3390/s20185231.

[b-Robust] Robust inference. Encyclopedia of Mathematics. http://encyclopediaofmath.org/index.php?title=Robust_inference&oldid=50961

[b-Rodrigues] Rodrigues, Juliany Cola Fernandes, Godinho, Joseane Lima Prado, De Souza, Wanderley (2014), *Biology of Human Pathogenic Trypanosomatids: Epidemiology, Lifecycle and Ultrastructure*. Proteins and Proteomics of Leishmania and Trypanosoma. Subcellular Biochemistry. Vol. 74. pp. 1–42.

[b-Rokach] Rokach, L. Maimon, O. (2014), *Data Mining with Decision Trees: Theory and Applications*. World Scientific

[b-Roweis] Roweis, Sam; Hinton, Geoffrey (2002), *Stochastic neighbor embedding*. Neural Information Processing Systems.

[b-RSPB] Types of Parasites. RSPB: https://www.rspb.org.uk/birds-and-wildlife/natures-home-magazine/birds-and-wildlife-articles/food-chains/parasites/

[b-RSS] Really Simple Syndication. LearnU: https://learn-u.com/lesson/rss-really-simple-syndication/

[b-Salam] Salam.D., Hussain, R., (2020), *Handbook of Data Science Approaches for Biomedical Engineering: Laskar: Noninvasive fracture characterization based on the classification of sonic wave travel times.*

[b-Sallan] Sallan B., Hinton G. (2004), *Reinforcement Learning with Factored States and Actions*. Journal of Machine Learning Research 5. 1063–1088.

[b-Saltelli] Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. (2008). *Global Sensitivity Analysis*: The Primer. John Wiley & Sons.

[b-Sambasivan] Nithya Sambasivan, Shivani Kapania, Hannah Highfill, Diana Akrong, Praveen Paritosh, and Lora M Aroyo (2021), *Everyone wants to do the model work, not the data work*: Data Cascades in High-Stakes AI. In CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, Yokohama, Japan. ACM, New York, NY, USA 15 Pages. https://doi.org/10.1145/3411764.3445518

[b-Samuylova] Samuylova E. (2020), *Machine Learning Monitoring: What It Is, and What We Are Missing*: https://towardsdatascience.com/machine-learning-monitoring-what-it-is-and-what-we-are-missing-e644268023ba

[b-Scholz] Scholz, C.H. (2002), *The mechanics of earthquakes and faulting* (2 ed.). Cambridge University Press. pp. 81–84. ISBN 978-0-521-65540-8. Retrieved 6 December 2011.

[b-Scholz] Scholz, Christopher H., ed. (2002), *The seismic cycle*, The Mechanics of Earthquakes and Faulting (2 ed.), Cambridge: Cambridge University Press, pp. 244–299.

[b-ScienceDirect-1] Drifter: https://www.sciencedirect.com/topics/earth-and-planetary-sciences/drifter

[b-ScienceDirect-10] Convolutional Layer. Science Direct: https://www.sciencedirect.com/topics/mathematics/convolutional-layer

[b-ScienceDirect-11] Thermal Imaging: https://www.sciencedirect.com/topics/earth-and-planetary-sciences/thermal-imaging

[b-ScienceDirect-2] Ensemble Modeling: https://www.sciencedirect.com/topics/computer-science/ensemble-modeling

[b-ScienceDirect-3] Evolutionary Algorithm. Science Direct: https://www.sciencedirect.com/topics/computer-science/evolutionary-algorithm#:~:text=An%20evolutionary%20algorithm%20(EA)%20is,within%20specific%20constraints%20%5B45%5D.

[b-ScienceDirect-4] Geostationary Satellite. Science Direct: https://www.sciencedirect.com/topics/earth-and-planetary-sciences/geostationary-satellite

[b-ScienceDirect-5] Tomaszewski J. (2021), *Overview of the role of artificial intelligence in pathology: the computer as a pathology digital assistant*: https://www.sciencedirect.com/topics/computer-science/human-in-the-loop

[b-ScienceDirect-6] Inferential Statistics. Statistical Methods for Physical Science.

[b-ScienceDirect-7] Neural Networks. ScienceDirect: https://www.sciencedirect.com/topics/neuroscience/neural-networks

[b-ScienceDirect-8] Majority Voting. Science Direct: https://www.sciencedirect.com/topics/computer-science/majority-voting

[b-ScienceDirect-9] Artificial Neural Network. Science Direct: https://www.sciencedirect.com/topics/mathematics/artificial-neural-network

[b-scikit] Ensemble methods.scikit learn: https://scikit-learn.org/stable/modules/ensemble.html

[b-Scott] J. Scott Armstrong and Fred Collopy (1992), *Error Measures For Generalizing About Forecasting Methods: Empirical Comparisons*. International Journal of Forecasting. 8: 69–80.

[b-semtech] What Is LoRa®? Semtech: https://www.semtech.com/lora/what-is-lora

[b-Sendai] United Nations (2015). Sendai Framework for Disaster Risk Reduction.

[b-Sensors] Flash Flood Early Warning System in Colima, Mexico.Sensors (20): https://www.researchgate.net/publication/344306938_Flash_Flood_Early_Warning_System_in_Colima_Mexico

[b-Sentinal] Data Visibility: A Guide to the What, Why, and How. SentinalOne: https://www.sentinelone.com/blog/data-visibility/

[b-Shachar] Shachar; Kaufman; Saharon Rosset; Claudia Perlich (January 2011, *Leakage in Data Mining: Formulation, Detection, and Avoidance*. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 6: 556–563. doi:10.1145/2020408.2020496

[b-Shalev] Shalev-Schwartz.S, David.S (2014), *Decision-Tree, Understanding Machine-Learning*. Cambridge University Press.

[b-Shannon] C.E. Shannon, *A Mathematical Theory of Communication*, Bell Systems Technical Journal, Vol. 27, pp 379–423.

[b-Simonoff] Simonoff, Jeffrey S. (1998), *Smoothing Methods in Statistics*, 2nd edition. Springer.

[b-Sinharay] S. Sinharay, in International Encyclopedia of Education (Third Edition), 2010 (978-0-08-044894-7)

[b-Nisbet] Nisbet, R. Ph.D., Ken Yale D.D.S., J.D., in Handbook of Statistical Analysis and Data Mining Applications (Second Edition), 2018, https://doi.org/10.1016/C2012-0-06451-4, ISBN 978-0-12-416632-5

[b-Sprague] Sprague, R (1980), *A Framework for the Development of Decision Support Systems*.

[b-Springer-1] AI: A Glossary of Terms (2019) Springer https://link.springer.com/content/pdf/bbm%3A978-3-319-94878-2%2F1.pdf

[b-Springer-11] Static Model. Science Direct: https://www.sciencedirect.com/topics/earth-and-planetary-sciences/static-model

[b-Springer-2] Class Separability https://www.sciencedirect.com/topics/engineering/class-separability

[b-SPSS] SPSS Tutorials: Pearson Correlation: https://libguides.library.kent.edu/SPSS/PearsonCorr

[b-Standford-1] Generalized Intersection over Union. Standford University: https://giou.stanford.edu/

[b-Standford-2] Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression. Stanford University

[b-Statistics] Classification and Regression Trees (CART) https://www.statistics.com/glossary/classification-and-regression-trees-cart/

[b-Statistics] False Alarm Ratio: Definition. Statistics How to : https://www.statisticshowto.com/false-alarm-ratio-definition/

[b-Steffensen] Steffensen, J.F. (2006), Interpolation (Second ed.). Mineola, N.Y.

[b-Survey] Information retrieval models in Neural Networks Framework: A survey.

[b-Taylor] Taylor, K.E. (2001), *Summarizing multiple aspects of model performance in a single diagram*, J. Geophys. Res., 106(D7), 7183-7192.

[b-Techno] Technopedia https://www.techopedia.com/

[b-Techtarget] Big Data Analytics: https://www.techtarget.com/searchbusinessanalytics/definition/big-data-analytics

[b-Teif] Teif, V.B., Rippe, K. (2010), *Statistical–mechanical lattice models for protein–DNA binding in chromatin*. J. Phys.: Condens. Matter. 22 (41): 414105.

[b-Tensor-1] Flatten. Tensor Space: https://tensorspace.org/html/docs/layerFlatten.html

[b-Tensor-2] Integrated Gradients. Tensorflow: https://www.tensorflow.org/tutorials/interpretability/integrated_gradients

[b-Terminology] Terminology. Resilience Metrics: https://resiliencemetrics.org/terminology

[b-Thomas] Thomas, C., Ranchin, T., Wald, L., Chanussot, J. (2008), *Synthesis of multispectral images to high spatial resolution: a critical review of fusion methods based on remote sensing physics*. IEEE Transactions on Geoscience and Remote Sensing. 46 (5).

[b-TIBCO] What is Data Integration? https://www.tibco.com/reference-center/what-is-data-integration

[b-Time] Time Series Forecasting: Definition, Applications, and Examples: https://www.tableau.com/learn/articles/time-series-forecasting

[b-Ting] Ting K.M. (2011), Confusion Matrix. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30164-8_157

[b-Topic] Topic Modeling: An Introduction. MonkeyLearnhttps://monkeylearn.com/blog/introduction-to-topic-modeling/

[b-TowardsData] Towards Data Science. Understanding Singular Value Decomposition and its Application in Data Science.

[b-TowardsScience] Introduction to Logistic Regression: https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148

[b-Tulane] What Is Disaster Management? Understanding Emergencies From Prevention to Mitigation. Tulane University – School of Public Health and Tropical Medicine.

[b-UAS] Unmanned Aircraft Systems (UAS). International Civil Aviation Organization

[b-Ullo] Ullo S., Sinha G. (2020), *Advances in Smart Environment Monitoring Systems Using IoT and Sensors*. MDPI Journal

[b-UnderstandAI] The Solid Facts of Ground Truth Annotations. Understand AI: https://understand.ai/blog/annotation/machine-learning/autonomous-driving/2021/05/31/ground-truth-annotations.html

[b-UNDESA] About the LDC Category. UNDESA: https://www.un.org/development/desa/dpad/least-developed-country-category.html

[b-UNDRR] United Nations Office for Disaster Risk Reduction Terminology.

[b-UNESCO-1] Disaster Risk Reduction: Why is Disaster Risk Reduction Important. UNESCO: https://en.unesco.org/disaster-risk-reduction

[b-UNESCO-2] Concept of Governance. International Bureau of Education: http://www.ibe.unesco.org/en/geqaf/technical-notes/concept-governance

[b-UNGGIM] Strategic Pathway. UNGGIM https://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/IGIF_SP4-Data_FIRST_DRAFT.pdf

[b-UNGGRF] Global Geodetic Reference Frame for Sustainable Development: https://www.unggrf.org/articles/global-geodetic-reference-frame-sustainable-development

[b-UP42] Everything you need to know about Digital Elevation Models (DEMs), Digital Surface Models (DSMs), and Digital Terrain Models (DTMs): https://up42.com/blog/tech/everything-you-need-to-know-about-digital-elevation-models-dem-digital

[b-USDept] Natural Disaster Response and Recovery. US Department of Interior.

[b-USGS] United States Geological Survey and Glossary.

[b-VAHandbook] Management of data breaches involving sensitive personal information (SPI) (2012) – VA Handbook.

[b-Vail] Vail, E.F (1999), *Mapping Organizational Knowledge*, in: Knowledge Management Review, Issue 8, May/June, pp. 10-15.

[b-Valero] Valero, M., F. Li, and W. Song (2019), *Smart seismic network for shallow subsurface imaging and infrastructure security*. IJSNET, 31, 10, https://doi.org/10.1504/IJSNET.2019.101569 .

[b-van Otterlo] van Otterlo, M., Wiering, M. (2012), *Reinforcement learning and markov decision processes. Reinforcement Learning. Adaptation, Learning, and Optimization*. Vol. 12. pp. 3–42.

[b-Vergara] Vergara B.S, Chang TT., *The Flowering Response of the rice plant to photoperiod*.

[b-Volcano] Comprehensive monitoring provides timely warnings of volcano reawakening. Volcano Hazards. https://www.usgs.gov/programs/VHP/comprehensive-monitoring-provides-timely-warnings-volcano-reawakening

[b-Wade] Wade G. (1994), *Signal Coding and Processing, Channel Coding*.

[b-Waldner] Waldner JB (2008), *Nanocomputers and Swarm Intelligence*. London: ISTE John Wiley & Sons. p. 205.

[b-Wang] Wang, Zhou, Bovik, A.C., Sheikh, H.R., Simoncelli, E.P. (2004-04-01). *Image quality assessment: from error visibility to structural similarity*. IEEE Transactions on Image Processing. 13 (4): 600–612.

[b-Washington] Generalization and Overfitting.Western Washington University.

[b-Wasserman] Wasserman.L (2005), *All of Statistics: a concise course in statistical inference*. Springer texts in statistics. p. 51.

[b-West] West, Jeremy, Ventura, Dan, Warnick, Sean (2007), *Spring Research Presentation: A Theoretical Foundation for Inductive Transfer*. Brigham Young University, College of Physical and Mathematical Sciences. Archived from the original on 2007-08-01.

[b-Western] Western Oregon University. The Singular Value Decomposition in Symmetric (L¨owdin) Orthogonalization and Data Compression.

[b-WGS] World Geodetic System – 1984 (WGS-84) Manual.

[b-WhiteHouse] The White House, Office of the Press Secretary, Executive Order 12906: Coordinating Geographic Data Access, 11 April 1994 (01.04.2022).

[b-WHO] Vector-borne diseases. World Health Organization (WHO): https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases

[b-WHO] World Health Organization Global, WHO Guidance – Ethics and governance of artificial intelligence for health. https://www.who.int/publications/i/item/9789240029200

[b-Willyan] Willyan D. Abilhoa, Leandro N. de Castro (2014), *A keyword extraction method from twitter messages represented as graph*, Applied Mathematics and Computation 240308-325.

[b-WMO-1] World Meteorological Organization GRIB (GRIdded Binary): https://docs.safe.com/fme/html/FME_Desktop_Documentation/FME_ReadersWriters/grib/grib.htm

[b-WMO-2] World Meteorological Organization: https://web.archive.org/web/20160605094419/http://www.eumetcal.org/eurometglossary/nowcast.htm

[b-WHO GHE] World Health Organization Global Network of WHO Collaborating Centres for Bioethics (2015). Global Health Ethics. Key issues. ISBN 978 92 4 154911 0. ISBN (PDF) 978 92 4 069403 3.

[b-Wright] Wright, Sewall (1921), *Systems of Mating*. Genetics. 6 (2): 111–178. doi:10.1093/genetics/6.2.111

[b-Wu] Wu KJ (15 April 2020), *There are more viruses than stars in the universe. Why do only some infect us? – More than a quadrillion quadrillion individual viruses exist on Earth, but most are not poised to hop into humans. Can we find the ones that are*? National Geographic Society.

[b-XGBoost] Monotonic Constraints. XGBoost Tutorial: https://xgboost.readthedocs.io/en/stable/tutorials/monotonic.html

[b-Xiao] Xiao, R. (2010), Corpus creation. In N. Indurkhya & F. Damerau (eds) The Handbook of Natural Language Processing (2nd ed.), 147-165. London: CRC Press.

[b-Yukitoshi] Yukitoshi Nishimura, Olga Verkhoglyadova, Yue Deng, Shun-Rong Zhang (2021), *Cross-Scale Coupling and Energy Transfer in the Magnetosphere-Ionosphere-Thermosphere System*, 1st Edition, December.

[b-Zamo] Zamo, M., & Naveau, P. (2018), *Estimation of the continuous ranked probability score with limited information and applications to ensemble weather forecasts*. Mathematical Geosciences, 50(2), 209-234. http://dx.doi.org/10.1007/s11004-017-9709-7

[b-Zhou] Zhou, Z.-H. (2018), *A brief introduction to weakly supervised learning*. National Science Review, 5, 44–53, https://doi.org/10.1093/nsr/nwx106.

[b-Zurich] Folds, Faults and Rock Deformation. University of Zurich.