CryoslideRisk Workshop Glossary for Mass Movement Terms

* indicates definitions from Van Everdingen, R. 2005. *Multi-language glossary of permafrost and related ground-ice terms*. National Snow and Ice Data Center/World Data Center for Glaciology, Boulder, CO. World Wide Web Address: http://nsidc.org/fgdc/glossary

§ indicates definitions from Turner, A. K. and Schuster, R. L., Eds. 1996. *Landslides: Investigation and Mitigation*: Transportation Research Board, Special Report 247, National Academy Press, Washington, D.C.

***Active layer:** The layer of ground that is subject to annual thawing and freezing in areas underlain by permafrost

***Active layer detachment slide:** a hillslope failure that has a failure plane at the base of the active layer

***Buried ice:** A form of ground ice. Ice formed or deposited on the ground surface and later covered by sediments.

***Creep of frozen ground:** The slow deformation (or time-dependent shear strain) that results from long-term application of a stress too small to produce failure in the frozen material

***Cryolithology:** The study of the genesis, structure and lithology of frozen earth materials

***Cryosphere:** That part of the earth's crust, hydrosphere, and atmosphere subject to temperatures below 0°C for at least part of each year

***Debris flow:** A sudden and destructive variety of landslide, in which loose material on a slope, with more than 50 percent of particles larger than sand size, is mobilized by saturation and flows down a channel or canyon.

§**Factor of Safety:** For landslides, the factor of safety (FS) is defined as the sum of the driving forces over the sum of the resisting forces. A FS of 1 indicates that the slope is on the verge of failure or has already failed.

Frozen debris lobe: A slow-moving landslide in permafrost

***Ground ice:** massive ice bodies, interstitial and pore ice within the permafrost

***Permafrost:** Ground that remains at or below 0°C for two years or longer

***Permafrost zonation:** Continuous (permafrost occurs everywhere across the landscape)

***Rock glacier:** A mass of rock fragments and finer material, on a slope, that contains either an ice core or interstitial ice, and shows evidence of present movement

***Retrogressive thaw slump:** A slope failure resulting from thawing of ice-rich permafrost. Retrogressive thaw slumps consist of a steep headwall that retreats in a retrogressive fashion due to thawing, and a debris flow formed by the mixture of thawed sediment and meltwater that slides down the face of the headwall and flows away.

§**Strength parameters:** One way to describe the strength of the soil or rock is through its strength parameters, cohesion (C), and angle of internal friction (φ). These values are determined through laboratory testing.

***Talik:** A layer or body of unfrozen ground in an area of permafrost

§**Types of landslides:** Landslides can be grouped into **fall, topple, slide, spread,** and **flow.**

Rockfall: detachment of soil or rock from a steep slope

Topple: forward rotation of a mass of soil or rock out of the slope

Slide: downslope movement of soil or rock on a rupture surface

Spread: extension of a cohesive soil or rock mass and subsidence of this mass into softer underlying material

Flow: spatially continuous movement, with distribution of velocities in the displaced mass resembling a viscous fluid

CryoslideRisk Workshop Glossary for Machine Learning Terms
(https://ml-cheatsheet.readthedocs.io/en/latest/glossary.html)

Accuracy
Percentage of correct predictions made by the model.

Algorithm
A method, function, or series of instructions used to generate a machine learning model. Examples include linear regression, decision trees, support vector machines, and neural networks.

Attribute
A quality describing an observation (e.g. color, size, weight). In Excel terms, these are column headers.

Bias metric
What is the average difference between your predictions and the correct value for that observation?

- **Low bias** could mean every prediction is correct. It could also mean half of your predictions are above their actual values and half are below, in equal proportion, resulting in low average difference.
- **High bias** (with low variance) suggests your model may be underfitting and you're using the wrong architecture for the job.

Bias term
Allow models to represent patterns that do not pass through the origin. For example, if all my features were 0, would my output also be zero? Is it possible there is some base value upon which my features have an effect? Bias terms typically accompany weights and are attached to neurons or filters.

Classification
Predicting a categorical output.

- **Binary classification** predicts one of two possible outcomes (e.g. is the email spam or not spam?)
- **Multi-class classification** predicts one of multiple possible outcomes (e.g. is this a photo of a cat, dog, horse or human?)

Classification Threshold
The lowest probability value at which we're comfortable asserting a positive classification. For example, if the predicted probability of being diabetic is > 50%, return True, otherwise return False.

Clustering
Unsupervised grouping of data into buckets.

Confusion Matrix

Table that describes the performance of a classification model by grouping predictions into 4 categories.

- **True Positives**: we *correctly* predicted they do have diabetes
- **True Negatives**: we *correctly* predicted they don't have diabetes
- **False Positives**: we *incorrectly* predicted they do have diabetes (Type I error)
- **False Negatives**: we *incorrectly* predicted they don't have diabetes (Type II error)

Deep Learning

Deep Learning is derived from one machine learning algorithm called perceptron or multi layer perceptron that gain more and more attention nowadays because of its success in different fields like, computer vision to signal processing and medical diagnosis to self-driving cars. As all other AI algorithms deep learning is from decades, but now today we have more and more data and cheap computing power that make this algorithm really powerful to achieve state of the art accuracy. In modern world this algorithm knowns as artificial neural network. deep learning is much more than traditional artificial neural network. But it was highly influenced by machine learning's neural network and perceptron network.

Dimension

Dimension for machine learning and data scientist is differ from physics, here Dimension of data means how much feature you have in you data ocean(data-set). e.g in case of object detection application, flatten image size and color channel(e.g 28*28*3) is a feature of the input set. In case of house price prediction (maybe) house size is the data-set so we call it 1 dimentional data.

Epoch

An epoch describes the number of times the algorithm sees the entire data set.

False Positive Rate

Defined as

$$FPR=1-Specificity=\frac{FalsePositives}{FalsePositives+TrueNegatives}$$

The False Positive Rate forms the x-axis of the ROC curve.

Feature

With respect to a dataset, a feature represents an attribute and value combination. Color is an attribute. "Color is blue" is a feature. In Excel terms, features are similar to cells. The term feature has other definitions in different contexts.

Feature Selection

Feature selection is the process of selecting relevant features from a data-set for creating a Machine Learning model.

Feature Vector

A list of features describing an observation with multiple attributes. In Excel we call this a row.

Gradient Accumulation

A mechanism to split the batch of samples—used for training a neural network—into several mini-batches of samples that will be run sequentially. This is used to enable using large batch sizes that require more GPU memory than available.

Label

The "answer" portion of an observation in supervised learning. For example, in a dataset used to classify flowers into different species, the features might include the petal length and petal width, while the label would be the flower's species.

Learning Rate

The size of the update steps to take during optimization loops like Gradient Descent. With a high learning rate we can cover more ground each step, but we risk overshooting the lowest point since the slope of the hill is constantly changing. With a very low learning rate, we can confidently move in the direction of the negative gradient since we are recalculating it so frequently. A low learning rate is more precise, but calculating the gradient is time-consuming, so it will take us a very long time to get to the bottom.

Loss

Loss = true_value(from data-set)- predicted value(from ML-model) The lower the loss, the better a model (unless the model has over-fitted to the training data). The loss is calculated on training and validation and its interpretation is how well the model is doing for these two sets. Unlike accuracy, loss is not a percentage. It is a summation of the errors made for each example in training or validation sets.

Machine Learning

Mitchell (1997) provides a succinct definition: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E." In simple language machine learning is a field in which human made algorithms have an ability learn by itself or predict future for unseen data.

Model

A data structure that stores a representation of a dataset (weights and biases). Models are created/learned when you train an algorithm on a dataset.

Neural Networks

Neural Networks are mathematical algorithms modeled after the brain's architecture, designed to recognize patterns and relationships in data.

Normalization

Restriction of the values of weights in regression to avoid overfitting and improving computation speed.

Noise

Any irrelevant information or randomness in a dataset which obscures the underlying pattern.

Observation

A data point, row, or sample in a dataset. Another term for instance.

Outlier

An observation that deviates significantly from other observations in the dataset.

Overfitting

Overfitting occurs when your model learns the training data too well and incorporates details and noise specific to your dataset. You can tell a model is overfitting when it performs great on your training/validation set, but poorly on your test set (or new real-world data).

Parameters

Parameters are properties of training data learned by training a machine learning model or classifier. They are adjusted using optimization algorithms and unique to each experiment.

Examples of parameters include:

- weights in an artificial neural network
- support vectors in a support vector machine
- coefficients in a linear or logistic regression

Precision

In the context of binary classification (Yes/No), precision measures the model's performance at classifying positive observations (i.e. "Yes"). In other words, when a positive value is predicted, how often is the prediction correct? We could game this metric by only returning positive for the single observation we are most confident in.

P=TruePositivesTruePositives+FalsePositivesP=TruePositivesTruePositives+FalsePositives

Recall

Also called sensitivity. In the context of binary classification (Yes/No), recall measures how "sensitive" the classifier is at detecting positive instances. In other words, for all the true observations in our sample, how many did we "catch." We could game this metric by always classifying observations as positive.

R=TruePositivesTruePositives+FalseNegativesR=TruePositivesTruePositives+FalseNegatives

Recall vs Precision

Say we are analyzing Brain scans and trying to predict whether a person has a tumor (True) or not (False). We feed it into our model and our model starts guessing.

- **Precision** is the % of True guesses that were actually correct! If we guess 1 image is True out of 100 images and that image is actually True, then our precision is 100%! Our results aren't helpful however because we missed 10 brain tumors! We were super precise when we tried, but we didn't try hard enough.
- **Recall**, or Sensitivity, provides another lens which with to view how good our model is. Again let's say there are 100 images, 10 with brain tumors, and we correctly guessed 1 had a brain tumor. Precision is 100%, but recall is 10%. Perfect recall requires that we catch all 10 tumors!

Regularization

Regularization is a technique utilized to combat the overfitting problem. This is achieved by adding a complexity term to the loss function that gives a bigger loss for more complex models

Reinforcement Learning

Training a model to maximize a reward via iterative trial and error.

ROC (Receiver Operating Characteristic) Curve

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").

Segmentation

It is the process of partitioning a data set into multiple distinct sets. This separation is done such that the members of the same set are similar to each otherand different from the members of other sets.

Supervised Learning

Training a model using a labeled dataset.

Test Set

A set of observations used at the end of model training and validation to assess the predictive power of your model. How generalizable is your model to unseen data?

Training Set

A set of observations used to generate machine learning models.

Transfer Learning

A machine learning method where a model developed for a task is reused as the starting point for a model on a second task. In transfer learning, we take the pre-trained weights of an already trained model (one that has been trained on millions of images belonging to 1000's of classes, on several high power GPU's for several days) and use these already learned features to predict new classes.

True Positive Rate

Another term for recall, i.e.

$$TPR=\frac{TruePositives}{TruePositives+FalseNegatives}$$

The True Positive Rate forms the y-axis of the ROC curve.

Underfitting

Underfitting occurs when your model over-generalizes and fails to incorporate relevant variations in your data that would give your model more predictive power. You can tell a model is underfitting when it performs poorly on both training and test sets.

Unsupervised Learning

Training a model to find patterns in an unlabeled dataset (e.g. clustering).

Validation Set

A set of observations used during model training to provide feedback on how well the current parameters generalize beyond the training set. If training error decreases but validation error increases, your model is likely overfitting and you should pause training.

Variance

How tightly packed are your predictions for a particular observation relative to each other?

- **Low variance** suggests your model is internally consistent, with predictions varying little from each other after every iteration.
- **High variance** (with low bias) suggests your model may be overfitting and reading too deeply into the noise found in every training set.