The goal of supervised learning is to build an artificial system that can learn the mapping between the input and the output, and can predict the output of the system given new inputs.

Overfitting is an undesirable machine learning behaviour that occurs when the machine learning model gives accurate predictions for training data but not for new data.

In this case, the machine learning model learns the details and noise in the training data instead of the underlying patterns.

K-fold cross-validation helps to mitigate the risk of overfitting and provides a more reliable assessment of how well the model is expected to perform on unseen data.

K-fold cross-validation is used to assess the performance and generalization capability of a machine learning model. It involves partitioning the dataset into k subsets, using k-1 subsets for training, and the remaining subset for validation. This process is repeated k times, and the performance measures are averaged.

Backpropagation is the essence of neural net training. It is the practice of fine-tuning the weights of a neural net based on the error rate (i.e. loss) obtained in the previous epoch (i.e. iteration.) Proper tuning of the weights ensures lower error rates, making the model reliable by increasing its generalization.

Ensemble Learning helps improve machine learning results by combining several models. This approach allows the production of better predictive performance compared to a single model. Bagging and Boosting are two types of Ensemble Learning.

In Bagging model that learns from each other independently in parallel and combines them for determining the model average.

The simplest way of combining predictions that   
belong to the same type.

It aim to decrease variance, not bias.

Each model receives equal weight & built independently.

In Boosting model learn sequentially and adaptively to improve model predictions of a learning algorithm.

A way of combining predictions that   
belong to the different types.

It aim to decrease bias, not variance.

Models are weighted according to their performance & are influenced   
by the performance of previously built models.

Gradient Descent is an optimization algorithm for finding a local minimum of a differentiable function. Gradient descent in machine learning is simply used to find the values of a function's parameters (coefficients) that minimize a cost function as far as possible.

Support Vector Machine (SVM) is a powerful machine learning algorithm used for linear or nonlinear classification, regression, and even outlier detection tasks. SVMs can be used for a variety of tasks, such as text classification, image classification, spam detection, handwriting identification, gene expression analysis, face detection, and anomaly detection.

Feature engineering is the process of selecting and transforming variables when creating a predictive model using machine learning. It is used to make the process more accurate and increases the prediction power of the algorithms by selecting the most critical variable, highlighting patterns and eliminating the redundant and irrelevant variables.

Decision trees are capable of handling numerical as well as categorical data, but, while implementing, we need to prepare the data for classification. There are two methods to handle the categorical data before training: one-hot encoding and label encoding.

* Label Encoding: If categorical data is label encoded, the decision tree can naturally interpret the encoded values as ordinal, assuming there is an inherent order among the categories.
* One-Hot Encoding: Allows the decision tree to make binary decisions based on the presence or absence of a specific category, avoiding assumptions of ordinal relationships.

Classification is the process of finding or discovering a model or function that helps in separating the data into multiple categorical classes i.e. discrete values.

* In this algorithm, we try to find the best possible decision boundary which can separate the two classes with the maximum possible separation.
* Evaluation metrics like Precision, Recall, and F1-Score are used here to evaluate the performance of the classification algorithms.
* Examples of classification algorithms are: Logistic Regression, Decision Trees, Random Forest, Support Vector Machines (SVM), Naive Bayes.

Regression is the process of finding a model or function for distinguishing the data into continuous real values instead of using classes or discrete values.

* In this algorithm, we try to find the best-fit line which can represent the overall trend in the data.
* Evaluation metrics like Mean Squared Error R2-Scoreand  are used here to evaluate the performance of the regression algorithms.
* Examples of regression algorithms are: Linear Regression, Polynomial Regression, Ridge Regression, Lasso Regression, Support Vector Regression (SVR).

Unsupervised learning refers to a class of problems in machine learning where a model is used to characterize or extract data, that is, a type of machine learning that learns from data without human supervision.

Example of a use case: Customer segmentation, or understanding different customer groups around which to build marketing or other business strategies.

A complex model can also be prone to overfitting, which means it memorizes the data and fails to generalize to new situations. On the other hand, a simple model can have fewer parameters, features, layers, or interactions that make it easier to train and interpret. However, a simple model can also be prone to underfitting, which means it misses the data and fails to capture the underlying patterns.

Random forest algorithm improve on decision trees by generating a collection of decision trees, instead of relying on a single decision tree.

Each trained on a random subset of the input data. The idea behind this approach is that by combining multiple models, the overall prediction accuracy will be improved.

The K-NN algorithm works by finding the K nearest neighbors to a given data point based on a distance metric, that is, the closer two given points are to each other, the more related and similar they are.

For example, if we have a dataset of tomatoes and bananas. KNN will store similar measures like shape and color. When a new object comes for instance, apple, it will check its similarity with the color (red or yellow) and shape.