A Brief Description of the ML Algorithms

Naive Bayes: Based on Bayes' theorem, Naive Bayes assumes all features are independent of each other. It calculates the probability of each class and the conditional probability of each class given a feature value, and then predicts the class with the highest probability.

Logistic Regression: A statistical method that estimates the probability of a binary outcome based on one or more predictor variables. The output is transformed using the logistic function to ensure it lies between 0 and 1.

SVM (Support Vector Machine): A linear classifier that aims to find the hyperplane that best separates the classes. For non-linear data, it uses the kernel trick to transform the data into a higher-dimensional space.

Decision Tree: A flowchart-like structure where each internal node represents a feature test, each branch is an outcome of the test, and leaf nodes represent class labels. It recursively splits the data to maximize the separation of classes.

Bagging: An ensemble method that trains multiple instances of a base algorithm on random subsets of the data. The final prediction is obtained by aggregating (e.g., majority vote for classification) the predictions of each instance.

Random Forest: An ensemble of decision trees trained on random subsets of data and features. The final classification is determined by majority voting among the trees.

Extra Trees: Like Random Forest, but it chooses random thresholds for splits on features rather than the optimal thresholds.

Ada Boosting: An iterative algorithm that adjusts the weights of training instances based on the errors of the previous classifier. New classifiers are trained to correct the mistakes of their predecessors.

Gradient Boosting: Builds an additive model in a forward stage-wise fashion. New trees are fit to the negative gradient (residual errors) of the loss function.

KNN (K-Nearest Neighbors): A non-parametric method that classifies a data point based on how its neighbors are classified. For a given data point, it finds the 'k' training examples that are closest to the point and returns the most common output value among them.

Neural Networks: Inspired by the structure of the human brain, a neural network consists of interconnected layers of nodes (or "neurons"). Each connection has a weight, and each neuron applies an activation function to its input. During training, the network adjusts these weights to minimize the difference between its predictions and actual outputs. Deep learning involves using neural networks with many layers (deep neural networks) to model complex patterns and representations in data.