Appendix 1: Glossary

TERM	MEANING
SVM	Support vector machines (SVMs) are a set of supervised learning methods used for classification, regression and outliers detection.
MLP	A multilayer perceptron (MLP) is a class of feedforward artificial neural network (ANN), which is mostly for classification tasks.
TRAINING DATA	The training data is an initial set of data used to help a program understand how to apply ML technologies to learn and produce sophisticated results
TESTING DATA	Training data's output is available to model whereas testing data is the unseen data for which predictions have to be made in order to evaluate model performance
CROSS-VALIDATION	It 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 that a given data sample is to be split into.
BINARY CLASSIFICATION	Binary classification is the task of classifying the elements of a set into two groups on the basis of a classification rule.
SVM KERNEL	SVM algorithms use a set of mathematical functions that are defined as the kernel. It allows us to operate in the original feature space without computing the coordinates of the data in a higher dimensional space.
SOFTMAX FUNCTION	Softmax function is an activation function for MLP that turns numbers aka logits into probabilities that sum to one. It outputs a vector that represents the probability distributions of a list of potential outcomes.
SMOTE	he Synthetic Minority Over-sampling Technique (SMOTE) is an oversampling approach that creates synthetic minority class samples.
GRADIENT DESCENT	Gradient descent is an optimization algorithm used to minimize some function by iteratively moving in the direction of steepest descent as defined by the negative of the gradient.
CONFUSION MATRIX	Confusion matrix is used for evaluating the quality of the output of a classifier on the iris data set.
PRECISON	It means the percentage of results which are relevant.
RECALL	It refers to the percentage of total relevant results correctly classified by the algorithm.
F1 SCORE	The F-score, also called the F1-score, is a measure of a model's accuracy on a dataset. It is used to evaluate binary classification systems, which classify examples into 'positive' or 'negative'.
LEARNING RATE	The learning rate is a configurable hyperparameter used in the training of neural networks that has a small positive value, often in the range between 0.0 and 1.0.
MOMENTUM	Momentum is a term used in gradient descent algorithm.
EARLY-STOPPING	Early stopping is a form of regularization used to avoid overfitting when training a learner with an iterative method, such as gradient descent. Such methods update the learner so as to make it better fit the training data with each iteration
OVERFITTING	Overfitting is a modeling error that occurs when a function is too closely fit to a limited set of data points.
HYPERPARAMETER TUNING	hyperparameter tuning is the problem of choosing a set of optimal hyperparameters for a learning algorithm. A hyperparameter is a parameter whose value is used to control the learning process.
LEARNING CURVE	A learning curve is a plot of model learning performance over experience or time. Learning curves are a widely used diagnostic tool in machine learning for algorithms that learn from a training dataset incrementally
AUC - ROC CURVE	It is one of the most important evaluation metrics for checking any classification model's performance. It is also written as AUROC (Area Under the Receiver Operating Characteristics)