**Machine Learning Algorithms – Pros and Cons**

**Advantages of Decision trees are:**

* Simple to understand and to interpret. Trees can be visualised.
* Requires little data preparation. Other techniques often require data normalisation, dummy variables need to be created and blank values to be removed.
* Able to handle both numerical and categorical data. Other techniques are usually specialised in analysing datasets that have only one type of variable. See [algorithms](https://scikit-learn.org/stable/modules/tree.html#tree-algorithms) for more information.
* They work well for both regression and classification tasks.
* They require relatively less effort for training the algorithm.
* They can be used to classify non-linearly separable data.
* They're very fast and efficient compared to KNN and other classification algorithms.
* Uses a white box model. If a given situation is observable in a model, the explanation for the condition is easily explained by boolean logic. By contrast, in a black box model (e.g., in an artificial neural network), results may be more difficult to interpret.

**The disadvantages of decision trees include:**

* Decision-tree learners can create over-complex trees that do not generalise the data well. This is called overfitting. Mechanisms such as pruning - setting the minimum number of samples required at a leaf node or setting the maximum depth of the tree are necessary to avoid this problem.
* Decision trees can be unstable because small variations in the data might result in a completely different tree being generated. This problem is mitigated by using decision trees within an ensemble.
* The problem of learning an optimal decision tree is known to be NP-complete under several aspects of optimality and even for simple concepts. Consequently, practical decision-tree learning algorithms are based on heuristic algorithms such as the greedy algorithm where locally optimal decisions are made at each node. Such algorithms cannot guarantee to return the globally optimal decision tree. This can be mitigated by training multiple trees in an ensemble learner, where the features and samples are randomly sampled with replacement.
* Decision tree learners create biased trees if some classes dominate. It is therefore recommended to balance the dataset prior to fitting with the decision tree.

## **Pros and Cons of Support Vector Machines -**

Every classification algorithm has its own advantages and disadvantages that come into play according to the dataset being analyzed. Some of the advantages of SVMs are as follows:

* The very nature of the Convex Optimization method ensures guaranteed optimality. The solution is guaranteed to be a global minimum and not a local minimum.
* SVM is an algorithm which is suitable for both linearly and nonlinearly separable data (using kernel trick). The only thing to do is to come up with the regularization term, C.
* It works really well with clear margin of separation
* It uses a subset of training points in the decision function (called support vectors), so it is also memory efficient.
* SVMs work well on small as well as high dimensional data spaces. It works effectively for high-dimensional datasets because of the fact that the complexity of the training dataset in SVM is generally characterized by the number of support vectors rather than the dimensionality. Even if all other training examples are removed and the training is repeated, we will get the same optimal separating hyperplane.

## **Disadvantages of SVMs are as follows:**

* They are not suitable for larger datasets because the training time with SVMs can be high and much more computationally intensive.
* It also doesn’t perform very well, when the data set has more noise i.e. target classes are overlapping.
* They are less effective on noisier datasets that have overlapping classes.

## **Pros and Cons of Random Forests**

* Both training and prediction are very fast, because of the simplicity of the underlying decision trees. In addition, both tasks can be straightforwardly parallelized, because the individual trees are entirely independent entities.
* The multiple trees allow for a probabilistic classification: a majority vote among estimators gives an estimate of the probability.
* The nonparametric model is extremely flexible, and can thus perform well on tasks that are under-fit by other estimators.

A primary disadvantage of random forests is that the results are not easily interpretable: that is, if you would like to draw conclusions about the meaning of the classification model, random forests may not be the best choice.