**Decision Tree for classification**

Classification is a two-step process, learning step and prediction step.

A decision is tree is a tree where the node represents the features, the branch represents a decision rule and the leaf node represents the outcome. A decision has a root which is the starting point of the tree.

Decision trees can handle high dimensional data with good accuracy

How decision tree works

According to datacamp.com there are 3 ways to perform a decision tree

**Attribute Selection Measures:** Attribute selection measure is a heuristic for selecting the splitting that partition data into the best possible manner

ASM provides a rank to each feature by explaining the given dataset. the best score attribute will be selected as s splitting attribute.

There are many more selecting algorithms of measures:

**Information gain:** it computes the difference between entropy before split and average entropy after split of the dataset given certain attributes

**Gain Ratio:** it biased the attribute with more outcomes. It takes the attribute that has large distinct value.

Gini index takes a binary split for each attribute. it will consider the Gini index that is minimum as a splitting attribute

**Advantages of decision tree**

It is can be easy to visualize and interpret the decision tree. It helps determine nonlinear pattern

It is easy to implement because, there is no need to normalize the columns

Decision tree is useful for handling data with numerical and categorial features with lees than 100 categories

**Disadvantages of decision tree**

Decision tree is very sensitive to noisy data

We must balance the dataset before creating the decision three to avoid any biased

To determine how well a test condition performs, we need to compare the degree of impurity of the parent before spliting with degree of the impurity of the child nodes after splitting. The larger their differnce, the better the test condition. The measurment of node impurity/purity are:

Gini Index

Entropy In Machine Learning, it is frequently used as an impurity measure: a set’s entropy is zero when

it contains instances of only one class

Misclassification Error (Humanoriented)

Stop the split procedure

A possible way to expend the node until either all the records to the same class have identical attribute values

Algorithm for Decision tree Induction

The decision tree induction algorithm works by recursively selecting the best attribute to split the data and expanding the leaf nodes of the tree until the stopping cirterion is met. The choice of best split test condition is determined by comparing the impurity of child nodes and also depends on which impurity measurement is used

Assumption

Decision Trees make very few assumptions about the training data (as opposed to linear models, which

obviously assume that the data is linear, for example). If left unconstrained, the tree structure will adapt

itself to the training data, fitting it very closely, and most likely overfitting it

Hopefully by now you are convinced that Decision Trees have a lot going for them: they are simple to

understand and interpret, easy to use, versatile, and powerful. However they do have a few limitations.

First, as you may have noticed, Decision Trees love orthogonal decision boundaries (all splits are

perpendicular to an axis), which makes them sensitive to training set rotation

More generally, the main issue with Decision Trees is that they are very sensitive to small variations in

the training data.

[**http://mines.humanoriented.com/classes/2010/fall/csci568/portfolio\_exports/lguo/decisionTree.html**](http://mines.humanoriented.com/classes/2010/fall/csci568/portfolio_exports/lguo/decisionTree.html)

<https://www.datacamp.com/community/tutorials/decision-tree-classification-python>