Decision Tree in Machine Learning

Decision Tree in Machine Learning

Decision Tree in Machine Learning is used for supervised learning [classification and regression]. Decision Tree exploits **correlation between features** and **non-linearity in the features**.

What is a Decision Tree ?

Wondering what a Decision Tree would be ? You might have come across the programmatic representation of a decision tree which is a nested if-else.

Let us consider the following pseudo logic, where we are trying to classify the given living-thing into either human, bird or plant :

```
if(displacement is
present){
```

```
if(displacement is present){
    if(wings are present AND feathers are present){
        living-thing is bird
    } else if(hands are present){
        living-thing is human
    }
} else if(displacement is absent){
    living-thing is plant
}
```

In the above pseudo code,

Output variable is category of living-thing whose value could be human or bird or plant.

Input variable is living-thing

Features of input data taken into consideration are displacement[whose values are present/absent], wings[whose values are present/absent], feathers[whose values are present/absent] and hands[whose values are present/absent]. So we have four features whose values are discrete.

What is the need of Decision Tree in Machine Learning

In the traditional programs, the above if-else-if code is hand written. Efforts put by a human being in identifying the rules and writing this piece of code where there are four features and one input are relatively less.

But could you imagine the efforts required if the number of features are in hundreds or thousands. Its becomes

a tedious job with nearly impossible timelines. **Decision Tree could learn these rules from the training data**. Despite other classifiers like Naive Bayes Classifier or other linear classifiers, Decision Tree could capture the non-linearity of a feature or any relation between two or more features.

Mentioning about capturing relation among features, in the above example, the features : wings and feathers are co-related. For the considered example(or data set), their values are related in a way such that their collective value is deciding on the decision flow.

Example Dataset :

In machine learning, input dataset for the Decision Tree algorithm would be the list of feature values with the corresponding categorical value. A sample of the dataset as shown in the below table :

Input	Output	Features			
living-being	category	wings	hands	feathers	displacement
Joe	human	absent	present	absent	present
Parrot	bird	present	absent	present	present
Jean	human	absent	present	absent	present
Hibiscus	plant	absent	absent	absent	absent
Eagle	bird	present	absent	present	present
Rose	plant	absent	absent	absent	absent

Each row in the above table represents an observation/experiment.

In practical scenarios, the number of features could be from single digit number to thousands, and the data set would contain single digit number to millions of entries/observations/experiments.
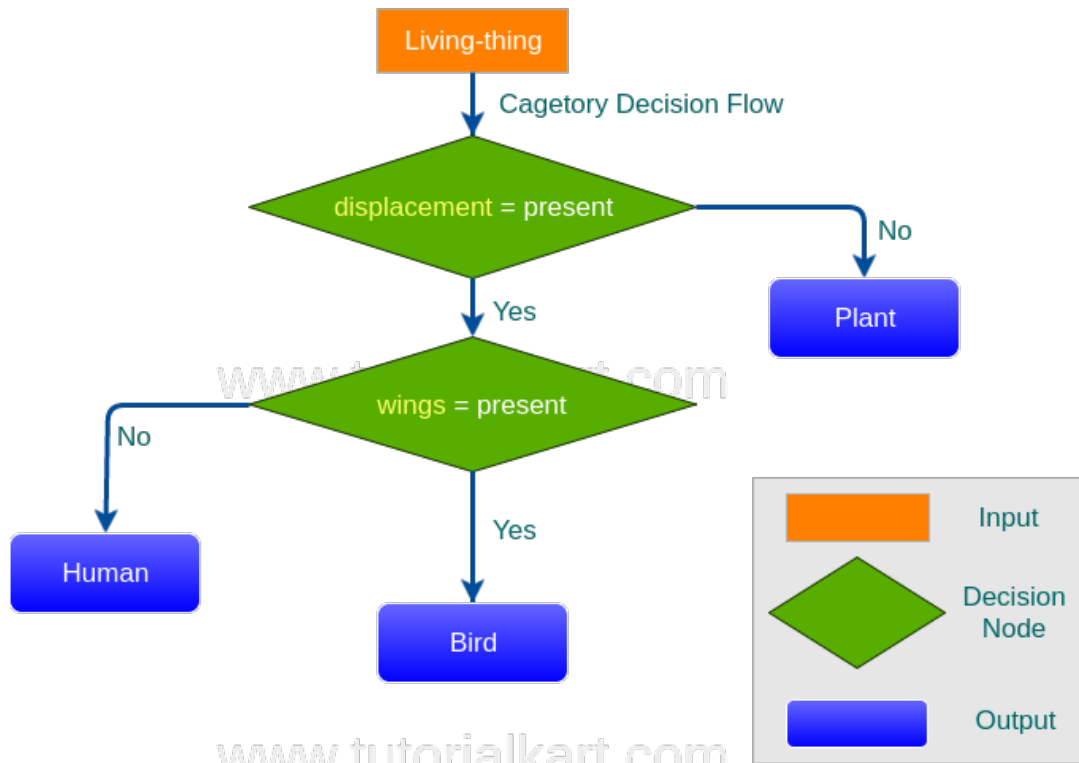
How is a Decision Tree built from the Dataset ?

The common way to build a Decision Tree is to use a greedy approach. Consider you are greedy on the number of Decision Nodes. The number of Decision Nodes should be minimal. By testing a feature value, the Dataset is broken into sub-Datasets, with a condition that the split gives maximum benefit to the classification i.e., the feature value considered(among all the possible feature value combinations) is the best available to categorize the given data set into two subsets. In each sub-Dataset, a new feature value combination is chosen, as in the former split, to divide it into smaller sub-Datasets, with the same condition that the split gives maximum benefit to the classification. The process is repeated until a Decision Node is not required to further split the sub-Dataset, and almost all of the samples in that sub-Dataset belong to a single category.

Please refer to Greedy Algorithm[https://en.wikipedia.org/wiki/Greedy_algorithm] for more insight.

Flowchart representation of Decision Tree :

The graphical representation of Decision Tree for the Dataset mentioned above would be as shown in the following diagram :

Example of Decision Tree in Machine Learning

From the above flowchart, it is evident that the Decision Tree has made use of only two features [displacement, wings] as the other two features are redundant. Thus reducing the number of Decision Nodes.

Conclusion :

In this Machine Learning tutorial, we have seen what is a Decision Tree in Machine Learning, what is the need of it in Machine Learning, how it is built and an example of it. Decision Tree is a building block in Random Forest Algorithm where some of the disadvantages of Decision Tree are overcome.

Machine Learning Tutorials
Machine Learning Algorithms
� Important considerations while working with a Machine Learning Algorithm
⊩ Decision Tree
⊩ Random Forest

Deep Learning
DeepLearning4j
⊩ Setup Environment & dl4j-examples