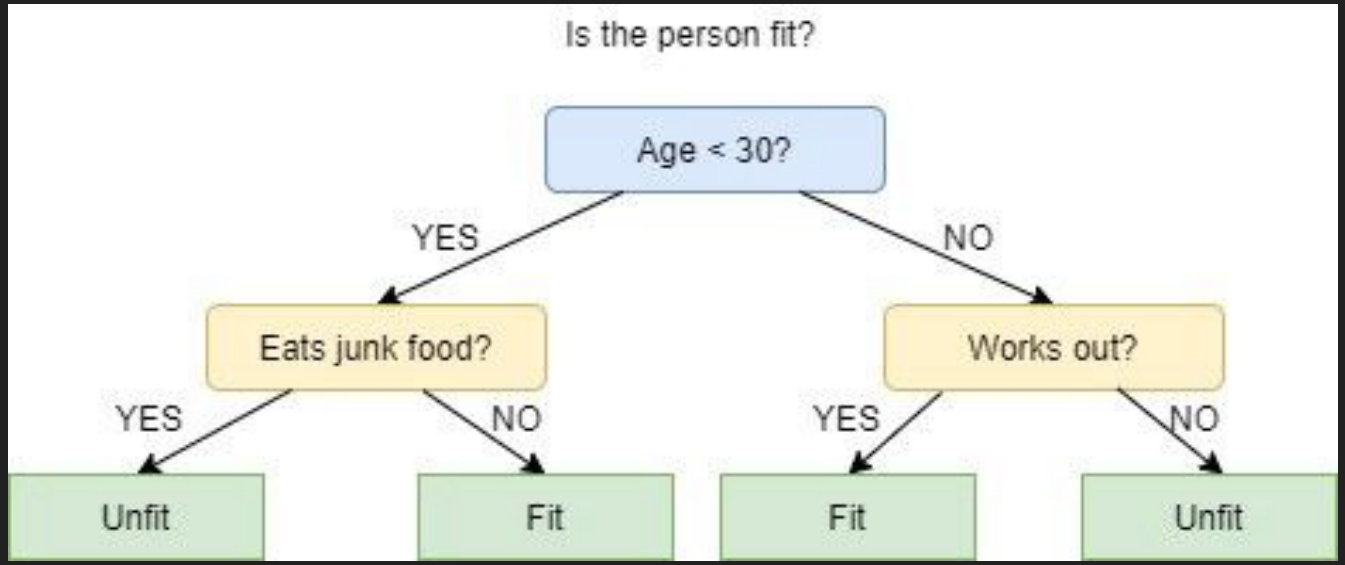
ID3 Algorithm | ML LAB 3 | VTU

By Prajwal Mani

What is Decision Trees?

- In simple words, a decision tree is a structure that contains nodes (rectangular boxes) and edges(arrows) and is built from a dataset (table of columns representing features/attributes and rows corresponds to records).
- Each node is either used to make a decision (known as decision node) or represent an outcome (known as leaf node).

Example of decision trees



What is ID3?

- ID3 stands for Iterative Dichotomiser 3 and is named such because the algorithm iteratively (repeatedly) dichotomizes(divides) features into two or more groups at each step.
- ID3 uses a top-down greedy approach to build a decision tree.
- Most generally ID3 is only used for classification problems with nominal features only.

How does ID3 select the best feature?

- ID3 uses **Information Gain** or just Gain to find the best feature.
- Information Gain calculates the reduction in the entropy and measures how well a given feature separates or classifies the target classes. The feature with the highest Information Gain is selected as the best one.
- **Entropy** is the measure of disorder and the Entropy of a dataset is the measure of disorder in the target feature of the dataset.

$$\text{Entropy}(S) = -\sum p_i * \log_2(p_i) \; ; \; i = 1 \text{ to } n$$

ID3 algorithm steps

1. Calculate the Information Gain of each feature.
2. Considering that all rows don’t belong to the same class, split the dataset **S** into subsets using the feature for which the Information Gain is maximum.
3. Make a decision tree node using the feature with the maximum Information gain.
4. If all rows belong to the same class, make the current node as a leaf node with the class as its label.
5. Repeat for the remaining features until we run out of all features, or the decision tree has all leaf nodes.

Entropy and info gain formula

Entropy

$$Entropy = \sum_{i=1}^{C} -p_i * \log_2(p_i)$$

info gain

$$Gain(T, X) = Entropy(T) - Entropy(T, X)$$