Decision Tree ID3 a supervised algorithm for classification

The ID3 algorithm is a decision tree algorithm that uses **entropy** to determine the best feature to split the data at each node of the tree. The algorithm follows the following steps:

1. Calculate the entropy of the current data set with respect to the target class labels: The entropy of a set of data with respect to the target class labels can be calculated using the formula:

Entropy(S) = -∑(p(i) * log2(p(i)))

where p(i) is the proportion of data belonging to class i in the set S and log2(p(i)) is the base-2 logarithm of p(i).

2. For each feature, calculate the entropy of the data set with respect to the target class labels for each possible split of the data based on the feature: For each feature, we calculate the weighted average of the entropies of the subsets created by splitting the data based on the feature's possible values. The formula to calculate the entropy for each feature is:

InformationGain(f) = Entropy(S) - ∑ (|Sv|/|S|) * Entropy(Sv)

where f is the feature, Sv is the subset of data for a specific value of feature f, |Sv| is the number of data points in Sv, and |S| is the number of data points in the original set S.

3. Choose the feature that results in the maximum information gain: The feature that results in the greatest reduction in entropy (i.e., the maximum information gain) is chosen as the best feature to split the data on.

4. Create a new tree node for the chosen feature and split the data set into subsets based on the feature's possible values: A new tree node is created for the chosen feature, and the data set is split into subsets based on the feature's possible values.

5. Repeat the process for each subset, until all the data in a subset belong to the same class or there are no more features to split the data on: The algorithm is repeated for each subset until either all the data in a subset belong to the same class or there are no more features to split the data on.

The final result of the ID3 algorithm is a decision tree that can be used to make predictions for new data by traversing the tree and returning the class label associated with each leaf node.