9.0 Gini Index, Entropy & Information Gain

1. Decision Trees split the target variable into different sub groups, which are relatively homogenous.(i.e. say subgroups of 1s and subgroups of 0s).
2. **(Definition of Homogenous: same, similar or alike.)**
3. A decision tree takes a statement / condition and makes a decision on whether the condition holds or not.
4. The conditions are represented along the branches & the outcome of the condition, as applied to the target variable, is shown on the node.
5. Arrows leading away from a node indicate a condition which is being applied to the node. Those pointing to a node indicate a condition that is being satisfied.
6. The decision space is split into smaller spaces leading to more and more homogenous subgroups and finally to a prediction.
7. Remember that the goal of machine learning is to decrease uncertainty or disorders from the dataset and hence use of decision trees.
8. **Entropy** is the quantitative measure of the randomness / disorder of the information being processed.
9. **High** Value of Entropy => Randomness is system is **high**, therefore making accurate predictions is **tough**.
10. **Low** Value of Entropy => Randomness is system is **low**, therefore making accurate predictions is **easier**.
11. Information Gain is the measure of how much information a feature provides about a class. **Low entropy** leads to increased Information Gain whereas **High entropy** leads to decreased Information Gain.
12. Information Gain computes the difference between **entropy before split** and average entropy **after split** of the dataset based on a given feature.
13. The split made in a Decision Tree is said to be pure if all the data points are accurately separated into different classes.
14. **Gini Impurity** measures the likelihood that a randomly selected datapoint would be incorrectly classified by a specific node.