Nurtai Maksat CSSE – 1605

**Decision Tree Classification**

The decision tree is one of the most popular machine learning algorithms. Decision trees are used for both classification tasks and regression. On this task, I used the classification and will try to explain what I learned. I had a question at the beginning of work. So what is a decision tree? Searched, found a lot of articles in the medium. Let me try to answer this question. - A decision tree is a tree where each node represents a feature, each branch represents a decision and each leaf represents an outcome.

There are several decision tree construction algorithms:

1. Gini
2. Entropy

I will use the Entropy algorithm. And so, in the process, I used all the possible parameters. As a result, I learned how the parameters affect our model. The first parameter is max\_depth. This indicates how deep the tree can be. The deeper the tree, the more splits it has and it captures more information about the data. The second parameter is min\_samples\_split. min\_samples\_split represents the minimum number of samples required to split an internal node. When we increase this parameter, the tree becomes more constrained, as it has to consider more samples at each node. I tried to increase this parameter but the result worsened. The third parameter, which I experimented, is min\_samples\_leaf. min\_samples\_leaf - the minimum number of samples that should be at the leaf node. This parameter is similar to min\_samples\_splits, however it describes the minimum number of samples of the samples on the leaves.