# Homework 4

For this analysis I'll use the UCI ML wine recognition dataset to predict wine cultivator using data from chemical analysis of wine. My main model will be a gradient boosting classifier with 200 regression trees. It'll be compared later to a multi-layer perceptron with one hidden layer with 100 neurons.

Task 4

For this task I'll compare the impact of flavonoids (organic compounds produced by wine grapes) on the probability of wine being from cultivator encoded as "1".
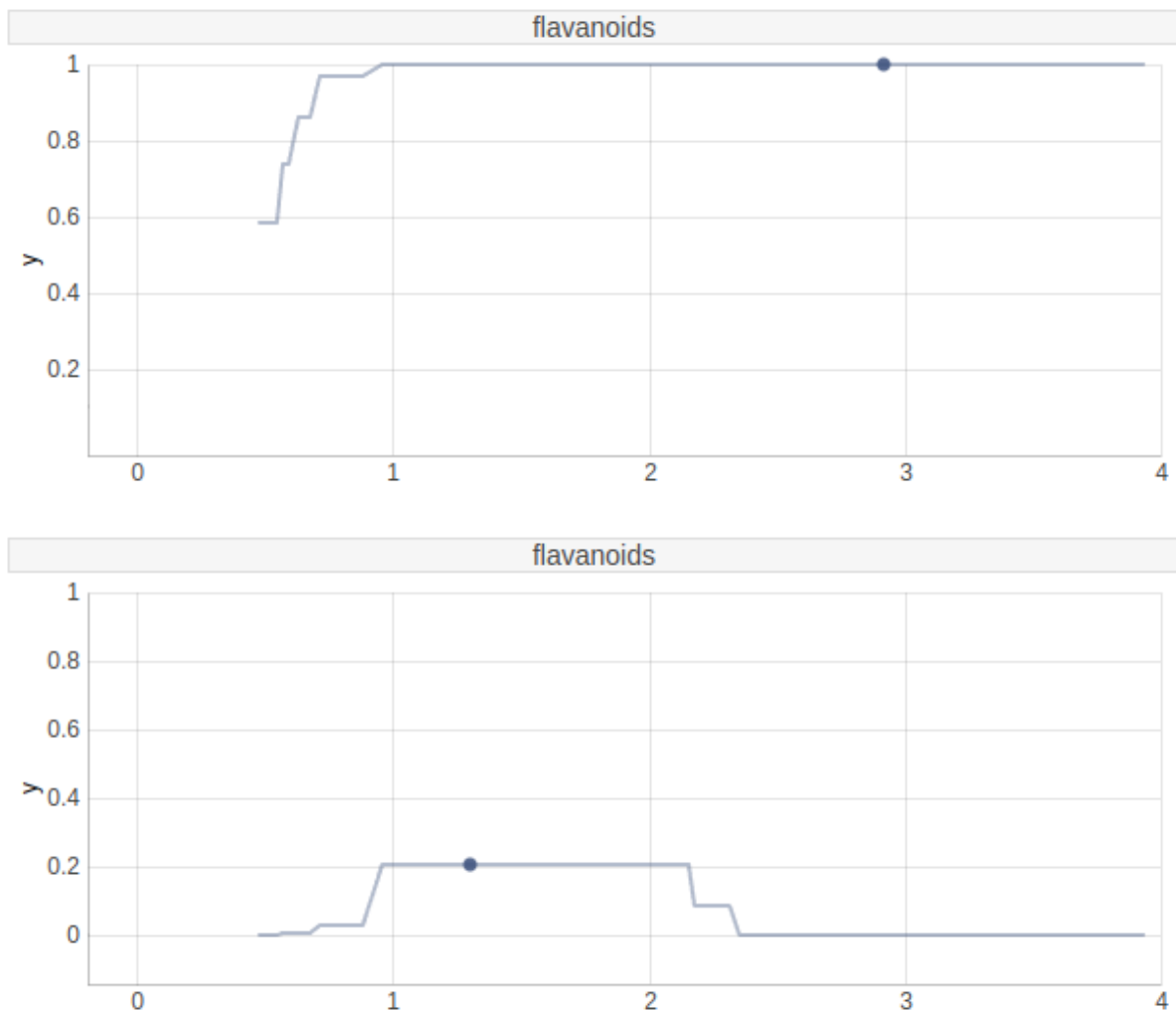
As we can see from the first example, the minimal amount of considered substance results in lower probability, which quickly rises to 1, as flavonoids approach 1. Situation looks completely different for the second example, here highest probability is obtained when the amount of flavonoids is between 1 and 2. But when this substance is in higher concentration probability decreases to 0. The highest probability in the second case is 5 times lower than in the first case. The reason for these differences is that other variables in those two examples are different, and changing only one variable may result in observations out of distribution due to interactions between variables.

Task 5
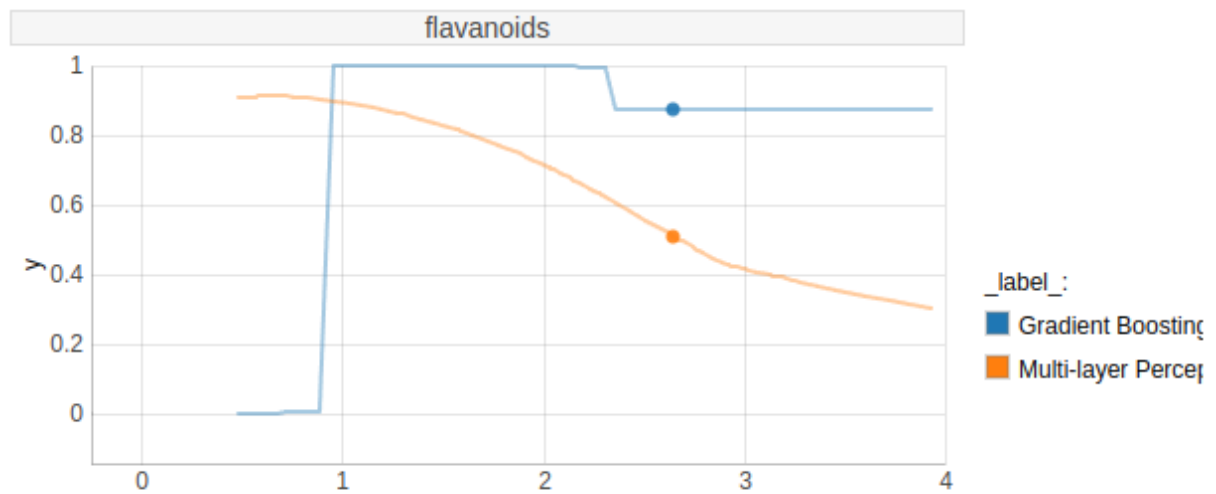
Here we are looking at the same substance. Interesting fact is that for multi-layer perceptrons an increasing amount of flavonoids decreases the probability of wine being from cultivator encoded as "1", while for gradient boosting classifier it increases for lower concentration of substance, and is mostly stable (slightly decreases) for higher concentrations.