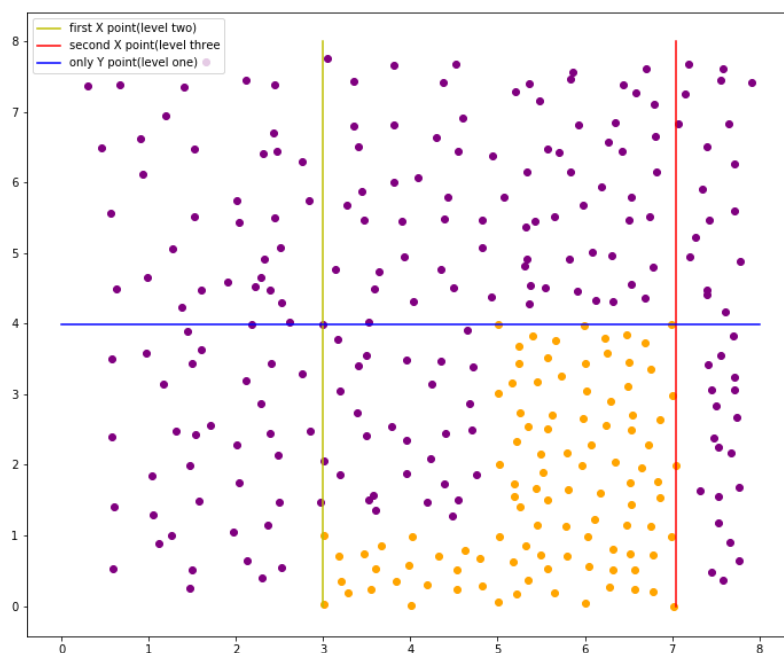
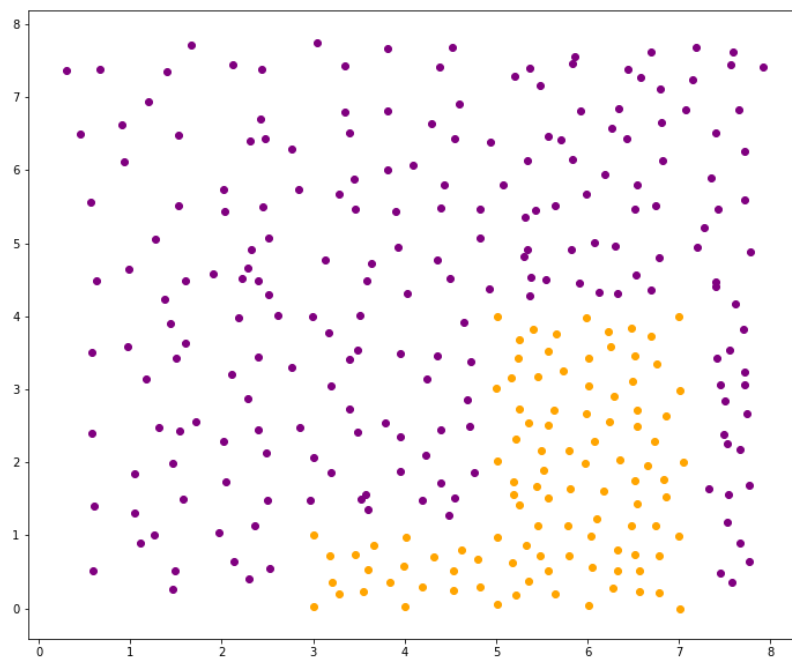
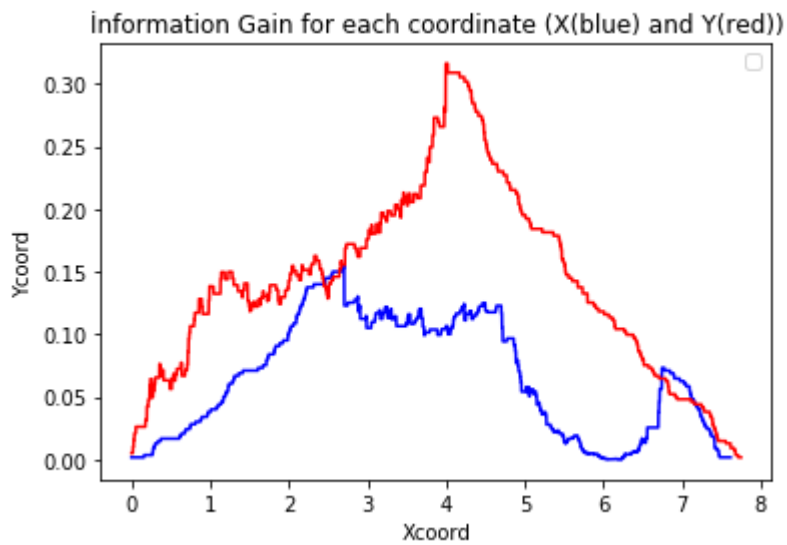


Machine Learning Assignment-3

Sefa Alp 041502017

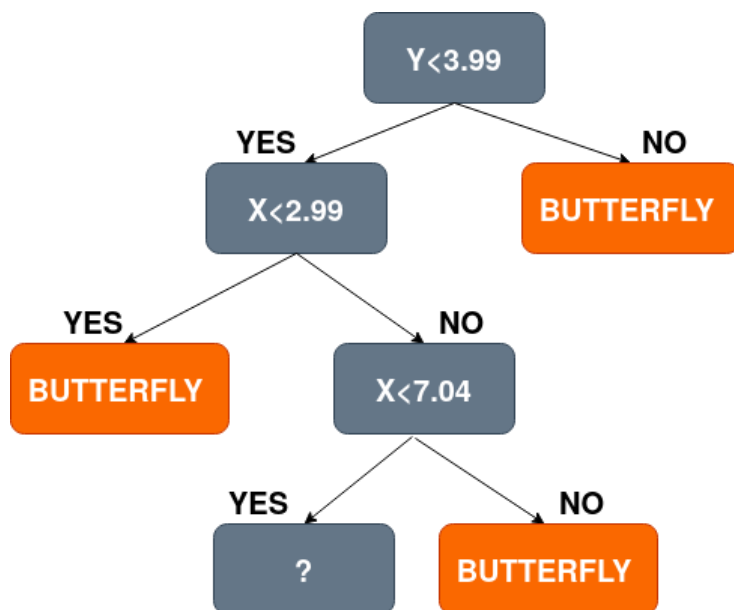


In first graph there are two types of dots. Purple dots represents butterflies and the orange dots represents birds in this case. With using information gain technique I find two highest information gain point as $Y(3.994)$ and $X(2.99)$. Because of, Y point has higher information gain (0.30 against 0.15), Y point chosen as first level point. Thus, $X(2.99)$ defined as second level point. With neglecting the values less than $X = 2.99$ I find third level point as $X(7.04)$ and the Y point stay same as (3.99). Below graph shows how I find first level point.



Here Y(3.994) point has the highest value as 0.315 and second highest point was X(2.99) as 0.155.

According to defined three level points I draw the below scheme.



To check my algorithm if it works properly or not. I used scikit-learn library's decision tree algorithm and compare the results. Results were very close. Almost my all three level values were same with scikit's level values. Below scheme represents scikit-learn decision tree algorithms result.

