Lab 4- Hagay Ringel

In general, since I came to New York 5 years ago I find myself flying often whether it's as part of work or as part of trips. The US is a huge country so in order to move from place to place the most effective tool is an airplane. Therefore, I chose the topic of flight satisfaction as a research topic, I am interested in what affects the satisfaction of the passenger (I'm asking myself, are these parameters relevant to me as well?), Which factors are more important to the passenger and which less, what is the weight of airplane cleaning versus flight service, Economy class passengers are necessarily less satisfied than Business class passengers? And is age a factor of satisfaction? All the above reasons (and there are more) made me intrigued and choose this topic for research and I would love to present the findings of this work. At this .model I used two algorithms, decision trees and random forest

First, according to the graphs in the model, age matters in terms of satisfaction level - usually, older passengers will be more satisfied than young passengers. Additional trend is that there is no affect of the distance of the flight on whether you are satisfying or not, in other words: A person who fly from NYC to Miami can be dissatisfy at the same level as someone who fly to .Australia

Accuracy- we calculate the accuracy to check the chances our classifier is right. In the DT algorithm, the average accuracy (after performing 5- fold cross validation) is 0.75 while in RF .it's 0.76. These values are not that high but still indicates that the classifier tends to be right

As for the classification report, I called the target values positive and negative: positive = 0= .satisfied, negative= 1= neutral or dissatisfied

Precision- we can see that in both RT and DT 3/4 of the cases the prediction is correct, and the answer is yes. That means that if we predict that a passenger will be satisfied, we have 75% in .DT and 76% in RF that our prediction will be correct, and the passenger will be satisfied

.Recall- aimed to predict how often the classifier will predict yes for our assuming
In 82%/81% in both algorithms the cases in the classifier will predict yes when we assume the passenger will be satisfied, while in only 67% in DT and 70% in RF the cases in the classifier will .predict yes when we assume a passenger will be neutral or dissatisfied

f1 score- Here in both algorithms the numbers are almost the same with a minor difference at the negative result (71%- DT, 72%-RF). This is the average of the precision and the recall and in .both cases the rate performance is high