**Business problem**

Create a model that tells what are the variables that influenced a person to buy or not a car after being exposed to an add on a social network.

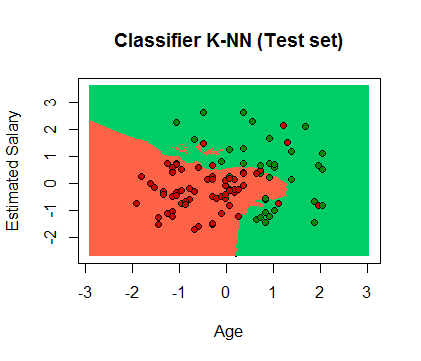
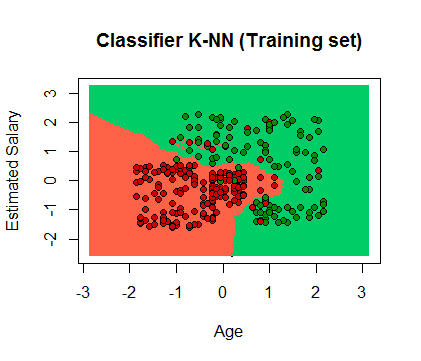
**Explaining the model**

1. Import the dataset, which is in a csv format.
2. We need to get rid of the columns that we won’t look at (gender and ID), so we need to subset the dataset.
3. After that, we will split the dataset into a training and test set (a 75% do the training set will be enough, as our dataset has 400 entries in total).
4. Before creating the classifier, we need to put all values on the same scale (after the sub-setting, the columns have new indexes).
5. After creating the classifier, we create the predictions. The first argument is the training set except the dependent variable (because we want to train our model without using the dependent), the second is the test set except the dependent variable (because we’re not “supposed” to know the real values), the third argument is the dependent variable and the last one is the number of neighbors (5 is the default value).
6. The confusing matrix allows us to compare the right vs wrong predictions.



As we can see, we got a total of 89 correct predictions (89%) and 11 wrong predictions (11%).

**Plotting the results**



**Evaluating the model’s performance**