#Tree model

--- As for this model, the data was divided into the natural groups X\_train, X\_test, y\_train, y\_test and after that the decision tree was correctly made. The confusion matrix was generated and was applied to both train and test sets.

As expected, the train set ofer a perfect distinction between the values of the targets, but it wasn't expected otherwise. In what concerns to the test set the results were a little diferent. Whereas the values "1" and "2" were significantly well classified, the same can not be said about "0", on which we see an almost equivalent distribution of the answers. This fact is displayed on the values of precision and recall, being respectively 0.65 and 0.63 for "1" and even better for "2" (0.73 and 0.65). Meanwhile we got humble results for value "0", only 0.37 and 0.47 respectively.

--- GraphViz could not be installed in the machine we worked on, so the visualization of the decision tree was made in GraphvizOnline, after the creation of a .dot file containing all the decision tree information. However, due to the extreme amounts of information in the dataset, the decision tree turned out to be massive, which turn out to shrunk then tree to a point where visualization was no longer possible.

#Rules model

--- After the workflow on Orange being done, following TP4 but using CN2 in trade for Tree, some print screens were made in order to better understand the results obtained. There were many rules discovered by the algorithm, the most significant being the ones with length 3. We can find some interesting discoveries and patterns,associated with probabilities for each result in the distributions, as can be seen in the image. To make this model, it was used CN2 as previously refered, with ordered rule ordering and exclusive as covering algorithm pattern. As for maximal rulelength was chosen the number 5 which preved to be suitable, dua to the fact that the rules discovered hat, at moste length=3.

The Data Table\_test and Data Table\_train, also displayed in the images, were then combined in a test & score, that led to the construction of a confusion matrix, for both train and test, and predictions