*Leo Toueg*

*20062982*

CISC 251 Exercise Sheet 5

1. I built the following workflow in KNIME, using the Support Vector Machine to predict the wine type:

Diagram

Description automatically generated

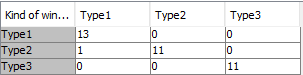
The partitioning node splits the data randomly into a 80/20 split, and tests the SVM on the 20%, which is used as testing data (Please let me know if this is generally what we are supposed to do with our models :^) ). Here are the results, using a Polynomial Kernel, with a C parameter of 1:

**Polynomial, C = 1:**

Table

Description automatically generated Perfect model, accuracy of 1.

**Polynomial, C = 3:**

Model accuracy: 97.222% 1 wrong classified. This makes sense as we are allowing for more error.

**Hyper Tangent, C = 1, C = 2, C = 3, C = 4, C = 5:**



We got the following error message. I am not entirely sure why, but it seems as though the hyper tangent method isn’t effective for our dataset.

**RBF, C = 1 AND C = 0.1:**

Table

Description automatically generated Model accuracy: 33.3%

This seemed to have labeled everything as a Type 2. I’m not sure why this is the case. Reducing the C parameter had no influence on the result.

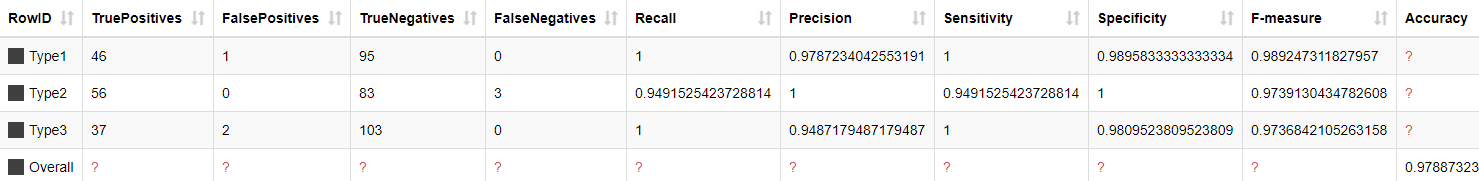
In conclusion, the Polynomial kernel function performed the best, using a C parameter of 1 or lower.

1. For the random forest model, I made the following workflow:

A picture containing timeline

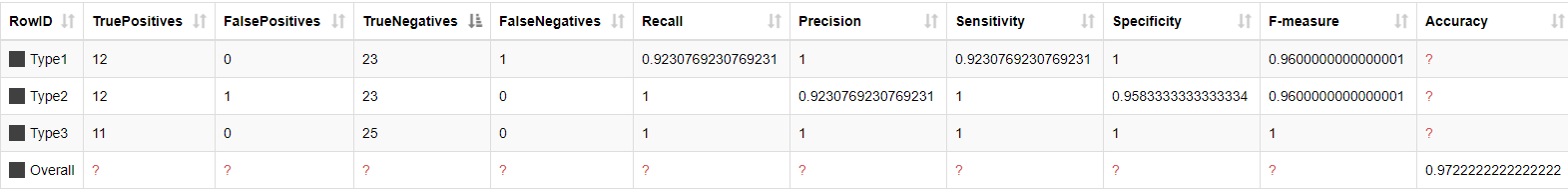
Description automatically generated

In order to get an idea of how good the random forest was, I also looked at the out-of-bag confusion matrix and model accuracy:



As we can see, we got an accuracy of 97.887%, this is really good.

Looking at the accuracy of the actual model, here are the results that were found:



With an accuracy of 97.2222%, this fits our out-of-bag rough estimate. This seems to imply that the random forest algorithm is a very good method.

Interestingly, increasing the number of trees from 100 to 1000 didn’t change any results. I hope this went into enough detail and I used the right modules, do let me know if there’s anything I should bother looking into! Thanks :^)