

## **Part 1- Impact of the different model parameters and network characteristics in the prediction performance using BMRF.**

Using human and yeast data, we choose values for the different model parameters (Appendix I, tables 1-3). Then, using chicken data, we defined a network for this species. The conclusions were as follows:

### **(i) Using human data**

- Using domain information increases the overall performance considerably (0.043 higher AUC). Thus, domain information will be used
- The filters of number of GO terms does not seem to have an effect. However, this may be the case only in those cases in which the number of GO terms remains high.
- Increasing the number of folds in the k-validation from 2 to 5 or 10 leads to an improvement of AUC of 0.026 and 0.037, respectively. K:10 will be used whenever possible.
- Due to time constraints, we will apply the filterminGOsize:20 and maxGOsize:0.1. However, the analysis could be extended to GO terms withminGOsize>9.

### **(ii) Using yeast data**

- The GO-size filters affect the overall prediction performance. We observed that as the filter form minGOsize becomes more strict and the filter for maxGOsize becomes less strict, the overall prediction performance increases.
- the GO-size filter seems to cause some alterations at the level of individual GO terms. In line with this, the prediction of one GO term are not entirely independent of the other GO terms considered in the analysis.
- We will perform the analysis with 20 replicates
- **Subsetting...**

### **(iii) Using chicken data**

- A Pearson correlation of 0.35 will be considered

### **(iv) Other conclusions.**

- Non-validated data seem to be more reliable for yeast than for humans.

Then, we investigated which characteristics in the network data have a larger impact of the prediction performance. For this, we used data from different species, since the network data differs considerably between these (Tables 7-9 in Appendix I). For a more detailed description of the differences in data between the species, as well as the data sources, see Tables 4-7 and Illustrations 1-3 in Appendix I.

The main conclusions from part 1b are:

- epp/tpepp seem to be the network parameter with a larger impact in the prediction performance.
- The epp/tpepp standirez is very similar for the species considered.
- Reducing Epn increases the performance
- Approaches like Positive-Unlabelled learnig can improve these parameter.

## Part 2 – BMRF results in poorly annotated species

Table mean, sd

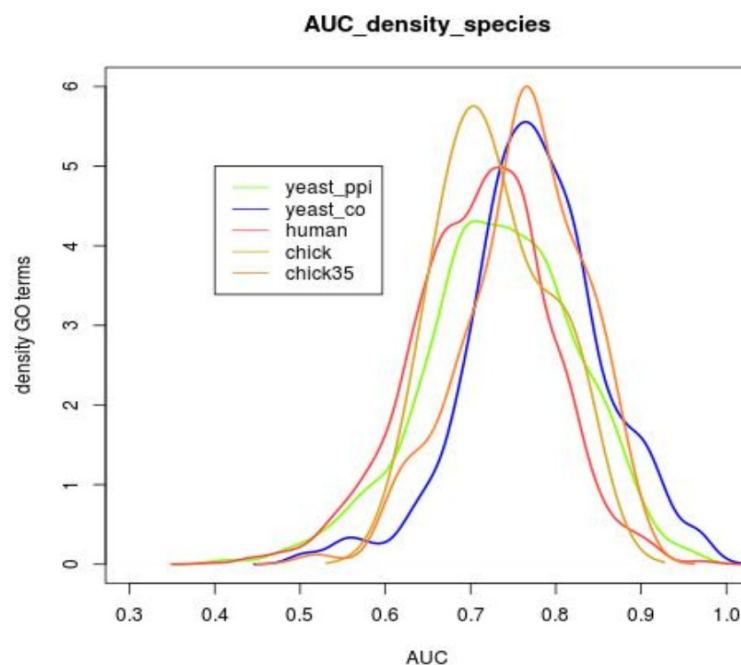


Illustration 1: AUC distribution for the different species.

Means are...

	yeast_ppi	yeast_co	humans	Chicken_07	Chicken_035
number of GOS with AUC>0.6	1013 (92%)	1155 (97.3%)	1819 (92%)	8 (89%)	140 (99.29)
mean depth	6.2	6.0	6.1	2.4	4.0
sd depth	1.5	1.6	1.5	0.5	1.3
>0.7	723 (66%)	1016 (85.6%)	1200 (60.06%)	5 (56%)	113(80.14%)
mean depth	6.5	6.0	6.2	2.4	4.1
sd depth	1.5	1.6	1.6	0.5	1.3
>0.8	281 (25,6%)	428 (36%)	384 (19.8%)	2	44(31.2)
mean depth	7.0	6.5	6.4	2.5	4.3
sd depth	1.4	1.4	1.7	0.7	1.4
>0.9	54 (5%)	83 (7%)	158 (8%)	0	0
mean depth	8.2	7.3	7.4	NaN	NaN
sd depth	1.0	1.4	2.1	NA	NA
>0.95	37 (3.4%)	21 (1.77%)	143 (7.2%)	0	0
mean depth	8.0	8.1	6.3	NaN	NaN
sd depth	0.9	1.7	1.5	NA	NA

*Table 1: Portion of GO terms above different AUC thresholds*

### **Part 3- Results using PU.**

Table differnces for chickens