

Project in Probabilistic Models

Eric Andrews

Tue 23, April 2013

High-level view of my approach

- ▶ **bnlearn** package for R
- ▶ Score-based local search methods
- ▶ Bayesian estimation of parameters
- ▶ Further in course: bootstrap aggregation, model averaging and 10-fold cross-validation.

My scoreboard progression

Submission #	1st	1st extra	2nd	3rd	2nd extra	4th
AUC	0.978	0.974	0.980	0.997	0.987	0.997
KL-divergence	3.060	2.338	2.341	1.321	1.564	1.204
AUC delta		-0.004	+0.002	+0.017	-0.010	0
KL delta		-0.722	+0.003	-1.017	+0.243	-0.117

- Learning arcs relatively easy; probabilities much harder.

1st submission

Submission #	1st	1st extra	2nd	3rd	2nd extra	4th
AUC	0.978	0.974	0.980	0.997	0.987	0.997
KL-divergence	3.060	2.338	2.341	1.321	1.564	1.204
AUC delta		-0.004	+0.002	+0.017	-0.010	0
KL delta		-0.722	+0.003	-1.017	+0.243	-0.117

- ▶ Learn 12 BNs with simulated annealing.
- ▶ Rank arcs by occurrence.
- ▶ Choose best BN by BDe score, learn parameters and predict probabilities.

1st extra submission

Submission #	1st	1st extra	2nd	3rd	2nd extra	4th
AUC	0.978	0.974	0.980	0.997	0.987	0.997
KL-divergence	3.060	2.338	2.341	1.321	1.564	1.204
AUC delta		-0.004	+0.002	+0.017	-0.010	0
KL delta		-0.722	+0.003	-1.017	+0.243	-0.117

- ▶ Previous idea pumped-up to 258 BNs.
- ▶ No improvement in arc ranking.
- ▶ Averaged network structure and lower ess \implies better KL-divergence.

2nd submission

Submission #	1st	1st extra	2nd	3rd	2nd extra	4th
AUC	0.978	0.974	0.980	0.997	0.987	0.997
KL-divergence	3.060	2.338	2.341	1.321	1.564	1.204
AUC delta		-0.004	+0.002	+0.017	-0.010	0
KL delta		-0.722	+0.003	-1.017	+0.243	-0.117

- ▶ 774 BNs; Try out Hill climbing and Tabu search as well.
- ▶ Arcs by occurrence in top 35 BNs.
- ▶ Test probs: Model averaging with top 35 BNs.

3rd submission

Submission #	1st	1st extra	2nd	3rd	2nd extra	4th
AUC	0.978	0.974	0.980	0.997	0.987	0.997
KL-divergence	3.060	2.338	2.341	1.321	1.564	1.204
AUC delta		-0.004	+0.002	+0.017	-0.010	0
KL delta		-0.722	+0.003	-1.017	+0.243	-0.117

- ▶ Finally, some progress!
- ▶ Learn BN structures using bootstrap aggregation.
- ▶ 150 with Tabu, 150 with SA
- ▶ Arc ranking: 50 best by hold-out likelihood using 10-fold cv.
- ▶ Test probs: model averaging over all 300 BNs.

2nd extra submission

Submission #	1st	1st extra	2nd	3rd	2nd extra	4th
AUC	0.978	0.974	0.980	0.997	0.987	0.997
KL-divergence	3.060	2.338	2.341	1.321	1.564	1.204
AUC delta		-0.004	+0.002	+0.017	-0.010	0
KL delta		-0.722	+0.003	-1.017	+0.243	-0.117

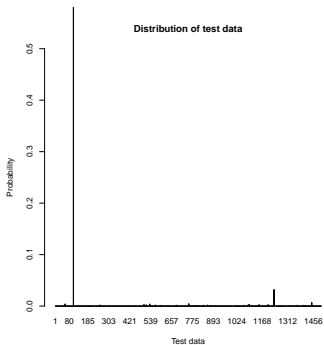
- ▶ Pump-up to 1400 BNs.
- ▶ Again, quantity doesn't help.
- ▶ Drop in AUC weird; Maybe because I used only Tabu?

Final submission

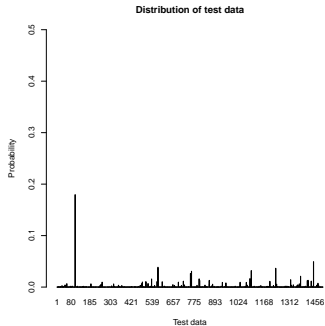
Submission #	1st	1st extra	2nd	3rd	2nd extra	4th
AUC	0.978	0.974	0.980	0.997	0.987	0.997
KL-divergence	3.060	2.338	2.341	1.321	1.564	1.204
AUC delta		-0.004	+0.002	+0.017	-0.010	0
KL delta		-0.722	+0.003	-1.017	+0.243	-0.117

- ▶ Discard BNs that give to any test point a probability mass greater than 0.3.
- ▶ Test probs: Model average with whats left.

Progression of probability predictions



First predictions. (worst)



Last predictions. (best)

Problematic line 104 in test data

1st	1st extra	2nd	3rd	2nd extra	4th
0.863	0.362	0.423	0.386	0.419	0.180

Done

Thank you for listening. Here is a picture of a Bayesian network:

