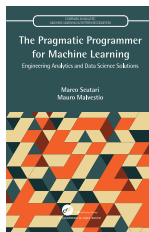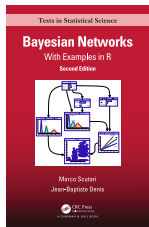- Home Page
- Documentation
- Examples
- Research Notes
- Small Simulation Studies
- Bayesian Network Repository

- About the Author

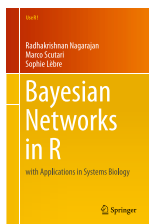The Pragmatic Programmer for Machine Learning
Engineering Analytics and Data Science Solutions
Marco Scutari
Mauro Malvestio

info & code

Texts in Statistical Science
Bayesian Networks
With Examples in R
Second Edition
Marco Scutari
Jean-Baptiste Denis
CRC Press

data & R code

UseR!
Radhakrishnan Nagarajan
Marco Scutari
Sophie Lèbre
Bayesian Networks in R
with Applications in Systems Biology
Springer

data & R code

Bayesian Network Repository

Several reference Bayesian networks are commonly used in literature as benchmarks. They are available in different formats from several sources, the most famous one being the Bayesian network repository hosted at the Hebrew University of Jerusalem. Others are shipped as examples of various Bayesian network-related software like Hugin or described in reference books such as Korb & Nicholson's "Bayesian Artificial Intelligence" or Koller & Friedman's "Probabilistic Graphical Models".

Even more networks are available from various papers that used Bayesian networks to analyze data from various domains. This is especially true for Gaussian networks and conditional linear Gaussian networks, since the original Bayesian network repository included only discrete Bayesian networks.

Here I collected all the networks that I produced in my work, that I used for various simulations and that I used to testing the implementations of `read.bif()`, `read.dsc()` and `read.net()` in **bnlearn**. All discrete networks are available in the BIF, DSC and NET formats and have been quality-checked and fixed as needed (*i.e.* all conditional probability distributions sum to one, no dummy nodes with a single level, no dangling dependencies on non-existent nodes, *etc.*). R objects with the `bn.fit` objects for all networks are provided both as RDA and RDS files. RDA files can be loaded with *e.g.* `load("asia.rda")`, which creates an object called `bn` in the current scope. RDS files can be loaded with *e.g.* `asia = readRDS("asia.rda")`, which returns the network and assigns it.

Discrete Bayesian Networks

Small Networks (<20 nodes)			
Name	**Nodes**	**Arcs**	**Parameters**
ASIA	8	8	18
CANCER	5	4	10
EARTHQUAKE	5	4	10
SACHS	11	17	178
SURVEY	6	6	21
Medium Networks (20–50 nodes)			
Name	**Nodes**	**Arcs**	**Parameters**
ALARM	37	46	509
BARLEY	48	84	114005
CHILD	20	25	230
INSURANCE	27	52	1008
MILDEW	35	46	540150
WATER	32	66	10083
Large Networks (50–100 nodes)			
Name	**Nodes**	**Arcs**	**Parameters**
HAILFINDER	56	66	2656
HEPAR2	70	123	1453
WIN95PTS	76	112	574
Very Large Networks (100–1000 nodes)			
Name	**Nodes**	**Arcs**	**Parameters**

Name	Nodes	Arcs	Parameters
ANDES	223	338	1157
DIABETES	413	602	429409
LINK	724	1125	14211
MUNIN (subnetwork #1)	186	273	15622
PATHFINDER	109	195	72079
PIGS	441	592	5618

Massive Networks (>1000 nodes)			
Name	Nodes	Arcs	Parameters
MUNIN (full network)	1041	1397	80592
MUNIN (subnetwork #2)	1003	1244	69431
MUNIN (subnetwork #3)	1041	1306	71059
MUNIN (subnetwork #4)	1038	1388	80352

Gaussian Bayesian Networks

Medium Networks (20–50 nodes)			
Name	Nodes	Arcs	Parameters
ECOLI70	46	70	162
MAGIC-NIAB	44	66	154

Large Networks (50–100 nodes)			
Name	Nodes	Arcs	Parameters
MAGIC-IRRI	64	102	230

Very Large Networks (101–1000 nodes)			
Name	Nodes	Arcs	Parameters
ARTH150	107	150	364

Conditional Linear Gaussian Bayesian Networks

Small Networks (<20 nodes)			
Name	Nodes	Arcs	Parameters
HEALTHCARE	7	9	42
SANGIOVESE	15	55	259

Medium Networks (20–50 nodes)			
Name	Nodes	Arcs	Parameters
MEHRA	24	71	324423

Last updated on `Tue Nov 29 13:14:40 2022` *with bnlearn* `4.9-20221107` *and* `R version 4.2.2 Patched (2022-11-10 r83330)` .