
A implementation of link strength measures for Bayesian networks

1 Description

Bayesian networks (BNs) [1, 2, 3] are probabilistic graphical models that serve to represent probabilistic dependencies between variables in uncertain domains. They are one of the most popular approaches in machine learning and have been applied to a large number of real-world problems [4, 5, 6, 7, 8].

Measures of connection strength (strength between any two nodes of the BN) and link strength (strength along a specific edge) in discrete BNs [9, 10, 11] are used to understand the importance of the relationships captured by the network. They are also used to visualize the connections in a BN learned from data.

2 Objectives

The goal of the project is to implement an algorithm to compute different connection and link strength measures in Python. The Python function will receive as input a BN and a given pair of nodes, and should output the link or connection strength (according to another input parameter specifying this choice).

The student should: 1) Implement the program to compute the measures. 2) Use different BN examples to validate the program. 3) Visualize a BN emphasizing the link or connection strength between the connected nodes.

As in other projects, a report should describe the characteristics of the design, implementation, and results. A Jupyter notebook should include calls to the implemented function that illustrate the way it works.

3 Suggestions

- Read the paper “Bayesian networks without tears”
<https://www.aaai.org/ojs/index.php/aimagazine/article/download/918/836>
- Try the pgmpy notebooks introducing BNs and methods to learn them from data:
https://github.com/pgmpy/pgmpy_notebook/tree/master/notebooks.
- Visit the Bayesian Network Repository for other examples of Bayesian networks:
<http://www.cs.huji.ac.il/~galel/Repository/#Networkformats>
- Read relevant papers on computation of link and connection strength [9, 10, 11].
- Implementations can use any other Python library.

References

- [1] David Heckerman, Dan Geiger, and David M. Chickering. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, 20:197–243, 1995.

- [2] K. Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.
- [3] J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2000.
- [4] S.K. Andersen, K.G. Olesen, F.V. Jensen, and F. Jensen. HUGIN— a shell for building Bayesian belief universes for expert systems. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, volume 2, pages 1080–1085, 1989.
- [5] Amir Hossein Ashouri, Giovanni Mariani, Gianluca Palermo, and Cristina Silvano. A Bayesian network approach for compiler auto-tuning for embedded processors. In *Embedded Systems for Real-time Multimedia (ESTIMedia), 2014 IEEE 12th Symposium on*, pages 90–97. IEEE, 2014.
- [6] C. Echegoyen, J. A. Lozano, R. Santana, and P. Larrañaga. Exact Bayesian network learning in estimation of distribution algorithms. In *Proceedings of the 2007 Congress on Evolutionary Computation CEC-2007*, pages 1051–1058. IEEE Press, 2007.
- [7] Ramon Etxeberria and Pedro Larrañaga. Global optimization using Bayesian networks. In A. Ochoa, M. R. Soto, and R. Santana, editors, *Proceedings of the Second Symposium on Artificial Intelligence (CIMAF-99)*, pages 332–339. Editorial Academia. Havana, Cuba, 1999.
- [8] P. H. Lee and H. Shatkay. Bntagger: improved tagging SNP selection using Bayesian networks. *Bioinformatics*, 22(14):e211–219, 2006.
- [9] B. Boerlage. Link strengths in Bayesian networks. Master’s thesis, Department of Computer Science, The University of British Columbia, 1992.
- [10] Imme Ebert-Uphoff. Measuring connection strengths and link strengths in discrete Bayesian networks. Research Report GIT-IIC-07-01, Georgia Institute of Technology, Interactive & Intelligent Computing Division, Georgia, US, 2007.
- [11] Imme Ebert-Uphoff. Tutorial on how to measure link strengths in discrete Bayesian networks. Georgia Tech Research Report GT-ME-2009-01, Georgia Institute of Technology, Interactive & Intelligent Computing Division, Georgia, US, 2007.