MLJ contribution information sheet (pls see the instructions)

***\*MLJ contribution information sheet*** *Each submission to the MLJ must be accompanied by a* ***1-2 page*** *information sheet that clearly answers the following questions. (Note that reviewers will expect these answers to be part of the introductory material of your paper; if this isn't currently the case, we strongly suggest you to revise your current manuscript before submission.) Please put the same care and effort in preparing this information sheet as you expect the reviewers to put in their reviews. Papers with incomplete or uninformative information sheets will be* ***returned without review****.   
  
1. What is the main claim of the paper? Why is this an important contribution to the machine learning literature?***Answer:** *We present a novel approach to learning dependency networks: first learn a Bayesian network, then perform a closed-form transformation of the Bayesian network to a dependency network. Learning RDNs via BNs scales much better to large datasets than with boosting, and provides competitive accuracy in predictions.*

*Many organizations store their information using relational structures (e.g., relational databases). Machine learning researchers have extensively investigated reasoning about such relational structures using formal logic. An important extension of this work is combining this formal logical reasoning with probabilistic reasoning about uncertainty, in particular to combine logic with Bayes nets, which have been widely used to represent uncertainty. We present a new method for making probabilistic predictions/inferences about relational structures using logic-based Bayes nets. We describe and evaluate fast algorithms for learning Bayes nets from relational data that perform well with our prediction method. Other researchers can use our results in two ways: (1) to perform inference for probabilistic queries about relational facts, given other facts in a relational structure. (2) To learn a Bayes net that represents statistical patterns in big relational data.*

*2. What is the evidence you provide to support your claim? Be precise.***Answer:** *We introduce a relational adaptation of the standard BN log-linear equation for the probability of a target node conditional on an assignment of values to its Markov blanket. The new log-linear equation uses a sum of expected values of BN log-conditional probabilities, with respect to a random instantiation of first-order variables. This is equivalent to using feature instantiation proportions as feature functions. We compare our approach to state-of-the-art functional gradient boosting methods on six benchmark datasets, using metrics of speed (time to learn) and accuracy (conditional log-likelihood and area under precision-recall curve). For most data sets, our method was much faster and of comparable accuracy to the functional gradient boosting methods.*

*3. What papers by other authors make the most closely related contributions, and how is your paper related to them?***Answer:**

*1). David Heckerman, David Maxwell Chickering, Christopher Meek, Robert Rounthwaite, Carl Kadie, and Pack Kaelbling. Dependency networks for inference, collaborative ltering, and data visualization. JMLR, 1:49-75, 2000.*

*2).* *Jennifer Neville and David Jensen. Relational dependency networks. JMLR, 8:653-692, 2007.*

*3). Sriraam Natarajan, Tushar Khot, Kristian Kersting, Bernd Gutmann, and Jude W. Shavlik. Gradient-based boosting for statistical relational learning: The relational dependency network case. Machine Learning, 86(1):25-56, 2012.*

*4).* *Tushar Khot, Sriraam Natarajan, Kristian Kersting, and Jude W. Shavlik. Learning Markov logic networks via functional gradient boosting. In ICDM, pages 320-329, 2011.*

*In Paper 1 the dependency networks were introduced by Heckerman et al.*

*In Paper 2 Neville and Jensen extend dependency networks to relational data.*

*In Paper 3 and Paper 4 the authors introduced the functional gradient boosting method for applying discriminative learning to build a generative graphical model (i.e. RDN, MLN).*

*The boosting approach to constructing a dependency network by learning a collection of discriminative models is very different from learning a Bayesian network. There are various options for hybrid approaches that combine the strengths of both. (1) Fast Bayesian network learning can be used to select features. Discriminative learning methods should work faster restricted to the BN Markov blanket of a target node. (2) The Bayesian network can provide an initial dependency network structure. Gradient boosting can then be used to fine-tune a discriminative model of a child node given parent nodes, replacing a flat conditional probability table.*

*4. Have you published parts of your paper before, for instance in a conference? If so, give details of your previous paper(s) and a precise statement detailing how your paper provides a significant contribution beyond the previous paper(s).*

**Answer:** *A preliminary version of this paper was presented at the StarAI 2012 workshop. A second version of this paper was presented at ILP 2014 but not published in the conference proceedings.*

*Main differences include: 1) a more extensive empirical evaluation, 2) relating our inference method to log-linear models, 3)* *discussion and proof of Consistency for RDNs.*

# Responses to ILP14 Reviews

|  |  |
| --- | --- |
| Review1 | Answers |
| Theoretical complexity study in section 3 | Explained in Section 5.1 |
| Pseudo code of Figure 2 | Updated figure 2 and added two algorithms |

|  |  |
| --- | --- |
| Review 2 | Answers |
| RDN consistency | Refer to Section 5.2 and the appendix for the details |
| Relational models v.s. Propositional models |
| Why our learning method is performing so well | Answer presented in Section 6.3 |
| only one large dataset | As indicated in Section 6.1, added even larger one IMDb |

|  |  |
| --- | --- |
| Review 3 | Answers |
| generate DN parameters given BN parameters | Addressed by adding Algorithm 2 and elaborating Table 1. |
| Minor typos | All fixed |