Dear Mrs. Riahi,

We have received the reports from our advisors on your manuscript, "Outlier Detection for Object-Relational Data Based on Graphical Models", which you submitted to Data Mining and Knowledge Discovery.

Based on the advice received, the Editor has assessed that your manuscript is not suitable for publication in this journal in its current form.  However, the Editor has assessed that your manuscript has significant merit and would like to encourage you to prepare a revised version for resubmission.  If you chose to do so, you are asked to carefully consider the reviewer comments which are attached, and submit a list of responses to them.  Your list of responses should be uploaded as a file in addition to your revised manuscript.

In order to submit your revised manuscript please log on and you will find your submission in the submissions needing revision box.  Click submissions needing revision, click edit submission, click attach files, and upload your revised version.  If you follow this process by 08 Jun 2018 we will endeavor to assign your manuscript to the same editor and reviewers, which greatly expedites the review process.  Otherwise it may be treated as a new submission.

Please note: When uploading your revised files, please make sure only to submit your editable source files (i. E. Word, tex).

Also, if you opt for a tex file, only one .tex file per submission should be used to guarantee proper building.  If it doesn't build, it is because the .sty files also have to be uploaded.  EM does support most new versions of LaTex, but the older ones are not.  More details are listed at the bottom of this email should you need it.

<https://dami.editorialmanager.com/>

Your username is: sriahi

If you have forgotten your password, kindly use the Send Login Details link on the login page.

Please click "Author Login" to submit your revision.

We look forward to receiving your revised manuscript.

Best regards,

Keerthana Govindarajan

Springer Journals Editorial Office

Data Mining and Knowledge Discovery

COMMENTS FOR THE AUTHOR:

Dear all,

Thank you for your submission to Data Mining and Knowledge Discovery.  I have received three reviews of your paper.  On the basis of these reviews I am unfortunately not able to accept your manuscript in its current form. In particular, while you managed to convince more or less two of the reviewers, reviewer 2 now become very negative. This is mainly due to the LR approach you now discuss again; it is good that you discuss it. However, the added discussion raised new issues raised by the Reviewer 2. Consequently,

I am asking you to provide a more in-depth discussion of the relation between ELD and LR. While I do no agree with the reviewer that a full theoretical justification is required, also because the reviewer was more positive in the previous round, I do think that the issues raised should be addressed as much as possible. For instance, would the divergence of the conditional distribution indeed work? How do we justify ELD given the rather strong theoretical

foundations for LR? With such a in-depth discussion and probably some more empirical evidence I would be happy to overrule Reviewer 2. So, if you revise the manuscript also paying careful attention to the other reviewers’ comments I would be pleased to receive the revision for another round of review.

Please note that addressing one issue sometimes raises new issues that were not previously apparent, so there is no guarantee that the paper will be accepted even if you do address all issues raised.

I look forward to receiving your revised manuscript should you decide to resubmit.

Yours sincerely,

Kristian Kersting,

DAMI Action Editor

Reviewer #1: The paper is definitely much improved.

Structurally, it is is still odd to me that section 5 comes before section Section 5. Section refers to Table 5, which has computations for two scores functions that are only defined in Section 6.  I'd fix this.

Minor points-per-game

p2 of the object linked to it → objects

p3 to baseline methods → to the baseline

p3 is independent → is an

p4 method(RIBL) → method (RIBL)

p5 a similarity measures → a similarity measured

p6 2 x Bayes network instead of Bayesian network as used in the rest of the paperwork

Section 3.1 twices refers to Table 1 as a summary.

P9: Fig 2 some of the probabilities seem to have extra )

p9 where v\_i resp. v\_pa\_i... → this sentence is difficult to parse as written (also it should have a \noindent)

p9 from context → from the context

p9 of a match only → reads funny, please reword

Section 4, uses subsubsection instead of subsections

p13 Interpretability, which... → this is not a complete sentence.  I think "Interpretability. This..." is better

p13 therefore an intuitively → and therefore is an intuitively

p13 class and object distribution → distributions

p13 F\_i,i=1,2 → F\_1 and F\_2 seems so much clearer

p15 it can be shown → needs a \noindent in front

p15 outier - outlier

p19 in literature → in the literature

p12 OutRank, recommended → OutRank as recommended

p22 "have high performance to detect outliers" → reword

p22 LR metrics → The LR metrics

p22 in NHL dataset → in the NHL datasets

p23 examine three top → examine the three top

p25 Bayesian network Learning → Bayesian network learning

p26 "we could found in" → that we could find in this dataset.

P30 season.Assists → season. Assists

Reviewer #2: From my previous review:

"SUMMARIZE ANY CHANGES THAT MUST BE MADE FOR A REVISED VERSION OF THIS PAPER TO BE ACCEPTABLE FOR PUBLICATION.

There must be a thorough comparison to likelihood-ratio tests (LR).  This comparison should include an explanation of when likelihood-ratio fails to detect outliers that ELD correctly detects, to help demonstrate the effective differences. A better theoretical explanation of ELD would also help, given the remaining weaknesses in the current explanation (discussed above)."

This revision adds LR to the experiments and discusses its relationship to ELD. However, it fails to provide a theoretical justification made to LR to obtain ELD. The stated motivations are avoiding the cancelation of differences and greater interpretability. There may be some interpretability advantage to ELD, although I think that LR could provide similar insights by looking at the divergence of each conditional distribution.

Avoiding cancelations, however, is not a strong justification for ELD, unless it can be shown that LR fails to detect clear outliers because of these cancelations. The toy examples demonstrate that LR obtains different numbers than ELD (as expected), but they do not show that LR ranks true outliers below normal instances.

From the text: "Table 5 illustrates the undesirable cancelling effects in LR."

Just because some of the terms in the sum are negative does NOT mean that the method is wrong. Entropy also includes positive and negative terms in its summation; however, taking the absolute value of each term in the entropy calculation would destroy its theoretical properties without providing any clear benefit. This seems to be the difference between LR and ELD as well. LR has a much stronger theoretical foundation than ELD; therefore, ELD requires strong empirical evidence to justify its use.

In Table 10, there is empirical evidence that LR works slightly worse on the high correlation and low correlation synthetic datasets (described in Figure 6(a) and (b)). This difference is small (AUC of 0.97 or 0.99 is quite close to 1.0). There is no discussion of statistical significance.

In Table 11 (real data), ELD outperforms LR in 4 out of 5 cases. This is somewhat interesting, but not strong enough support given the lack of theoretical justification or simple illuminating examples.

OTHER MINOR COMMENTS:

1. Regarding pseudo-likelihood and likelihood: It appears that pseudo-likelihood is being used as defined by Schulte (2011). However, the most common definition of pseudo-likelihood is from Besag (1975), and this pseudo-likelihood has already been used for learning probabilistic relational models (e.g., MLNs -- (Richardson & Domingos, 2006)). Besag's definition is also the one found on Wikipedia.

Please rename this function or at least clarify the difference from the standard, widely-known and widely-used pseudo-likelihood function.

2. "Khot et al. introduced a non-parametric relational one-class classification based on first-order trees. They proposed a tree-based distance metric to discover new relational features and to differentiate relational examples [23]."

Please explain why the work by Khot et al. isn't a suitable baseline. I imagine it could be justified, but it's worth including a sentence to make the reasoning explicit.

3. "log-likelihood distance (ELD)" -- you should explain why the initials "ELD" are an appropriate abbreviation for a log-likelihood distance, because it's not obvious.

4. In section 4, the subsections are listed as subsubsections (e.g., 4.0.1 instead of 4.1).

5. Section 6.1: The definition of LR\_i is missing parentheses around the difference of logs.

6. "The case studies illustrate that our outlier score is easy to interpret, because the Bayesian network provides a platform that makes the detected outliers very easy to interpret."  This is circular reasoning.

7. "statistical-relational learning, a recent field that combines AI and machine learning" -- I don't think this is a very good description of SRL, since it could apply just as well to Alpha-Go (which also combines AI and machine learning and is much more recent).

8. I suggest defining "i.i.d." before its first use.

9. Typos: "the the likelihood ratio" --> "the likelihood ratio"; "decomposing the log ration" --> "decomposing the log ratio"

Reviewer #3: Table 4 is not mentioned in text. Moreover, it seems to be out of place. It should probably appear in Sec 3.3, after the PBN figure has been introduced.

Pg 10, line 32: How are database frequencies different from data table frequencies? Aren't data tables also databases.

Page 11, line 26-28: It is not clear how are the probabilities in the equations computed. Can you please add the symbolic names (e.g P(shotEff = high...)) to equations.

I appreciate the re-ordering but now the descriptions and images are completely out of order. Table 5 refers to FD and ELD that have not yet been introduced. Also we don't know what the "high" and "low" correlation cases mean. Section 7.1 describing the synthetic dataset should be introduced in Section 5. It is a good running example to compare the issues with various scoring methods. Table 5 should follow the description of FD and ELD (maybe in section 6.2).

Table 7 is not mentioned anywhere in the text.

Page 22, line 30: Can you elaborate on how the center was computed for discrete features and why that results in the same distances for each individual.

Section 8.3: Am I correct in assuming that the model gave a higher outlier score to some object in the normal class than most of the actual outliers ? Even with few outlier feature values, I would at least have expected most of the outlier class objects to score higher than the normal class object.

Section 9: I would argue that an additional limitation would be the requirement that this approach needs sufficient data for an object to be able to reliably learn the object model.

Page 26, line 13: Which table contains these coefficients ranging from 0.45 and 0.82 ?

Scatter plots font size is too small and seems that there is enough space to use a larger font size.

Page 30, line 25: Seems a stretch to say that it can "identify future stars" given that the algorithm identifies them based on their successful performance. I would argue that these players are already stars, albeit relatively under-paid.

Section 10.3: Comparison to other metrics should be moved up before the sections that delve into the detailed analysis of the proposed method.

Minor issue:

Pg6, Line 43: The domain of predicate -> The range of predicate

Fig 2. P(Action) = T) -> P(Action) = T

properly relational ->  truly relational ?

Page 20, line 18: Table 7.3 -> Table 8

Page 22, line 45: ration -> ratio

Table 14: Bottom Team -> Teams

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

STYLE FILE INFORMATION:

For your reference, a sty file is:

STY is the file extension for a Style sheet file. A STYle template may be used by different publishers to define what should be bold/centered/italic in the paper.

The link below to Springer's own site can be used for latex references:

http://www.springer.com/sgw/cda/frontpage/0,,5-164-2-72376-0,00.html

We rely on the authors to create their sty files.  Authors need to create their sty files to govern their own work.

There is a link below we found in Google that may assist the author in completing their submission.

http://www.sci.usq.edu.au/staff/robertsa/LaTeX/latexintro.html

--

F.