You are being carbon copied ("cc:'d") on an e-mail "To" "Fatemeh Riahi" [sriahi@sfu.ca](mailto:sriahi@sfu.ca)   
CC: "Oliver Schulte" [oschulte@cs.sfu.ca](mailto:oschulte@cs.sfu.ca)   
  
Dear Mrs. Riahi,   
  
We are pleased to inform you that your manuscript, "Outlier Detection for Object-Relational Data Based on Graphical Models", has been accepted for publication in Data Mining and Knowledge Discovery subject to you making required changes by [28 Jul 2019](http://airmail.calendar/2019-07-28%2012:00:00%20PDT).   
  
When preparing your revised manuscript, you are asked to carefully consider the editor and reviewer comments which are attached, and submit a list of responses to the comments. 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 submissions in submissions needing revision box. Click submissions needing revision, click edit submission, click attach files, and upload your revised version. It is essential that you use this process and not submit the manuscript 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://www.editorialmanager.com/dami/>   
  
We look forward to receiving your revised manuscript.   
  
  
Best regards,   
  
Keerthana Govindarajan   
JEO Assistant   
Data Mining and Knowledge Discovery   
COMMENTS FOR THE AUTHOR:   
  
\*\*\*\*\*\*\*\*   
  
Dear Fatehmeh Riahi,   
  
Based on the assessment of the reviewers and action editor (see letter below), your paper has a good chance of being accepted for publication in Data Mining and Knowledge Discovery. Congratulations!   
  
However, before it can go into production, please take some time to make sure that your manuscript is ready for printing. In particular:   
  
- carefully read the author instructions that can be found at  <https://www.springer.com/computer/database+management+&+information+retrieval/journal/10618>, and make sure that your paper conforms to them. Your paper will be type-set by Springer, and not print in the way you turn it in. It is thus particularly important that you provide all information as required (e.g., citations and references).   
  
- carefully spell-check not only the text of the paper, but also the references (in particular with respect to capitalization of abbreviations etc.).   
  
- please fix the bibliography, and provide full publication information for all entries (all page numbers, avoid citations to TRs and arxiv if they have been later published at different venues, provide full conference names in addition to abbreviations, etc.). This will help the readers (and the production staff at Springer) to identify the right publications.   
  
- be aware that the production staff at Springer will change your citations from numbered citations to an author-year style. This does not always come out as expected, so I would recommend that you submit a final version with author-year citations where you have personally checked that they come out as you want it. See also previous publications in this journal and the instructions for authors.   
  
- be aware that the production staff at Springer will change your tables into a format that only uses horizontal lines. You may want to be pro-active on that and provide them with an appropriate format (e.g., using the LaTeX package ‘booktabs’).   
  
  
best regards,   
  
Johannes Fürnkranz   
Editor-in-Chief   
Data Mining and Knowlege Discovery   
  
  
Dear all,   
  
Thank you for your submission to DAMI. I am delighted to accept it for publication, subject to completion of rather minor revisions, which are stated in the corresponding reviews. As one of the reviewer states it, "this revision is a major rewrite and fixes   
the main issues with the previous versions."   
  
I look forward to receiving your revised manuscript.   
  
Yours sincerely,   
  
Kristian Kersting   
Action Editor   
  
Reviewer #1: I think that this paper presents a nice piece of work and deserves to be published. I commend the authors on their revisions.    
  
My only minor quibble is that in Section 10 for the sports examples, a number of the features are raw counts, which are clearly correlated with playing time. Hence some of the correlations may simply be explained by that. 

[added explanation addressing this point above table 14. See paragraph: “discussion of success metrics”]

Time played is reasonable to consider, but it also depends on things like injuries and what competitions a team plays in (e.g., resting a good player in the league ahead of a champions league game).    
  
I think it is worth commenting on this or at least adding this caveat to the discussion.

[acknowledged and added references to state-of-the-art player performance metrics that address the issue of isolating individual performance from context factors. See paragraph: “discussion of success metrics”]  
  
Minor points   
  
In the intro and conclusion you use 'log likelihood distances' but in rebuttal and the middle you talk about expected log ... to explain the ELD. I'd be consistent.

We went through all occurrences of log-likelihood distance and either added “expected” or rewrote to avoid mentioning the full “expected log-likelihood distance term” (which is a bit of a mouthful).

The following are smaller local which we fixed.

P3: relational data[21,50,13] -> needs a space between data and the cites

fixed  
  
P3 I think the 'quality measure' should be in \emph{} for consistency 

fixed

P3 fits fits

fixed  
  
P4 I think the 2 contributions would read better if written as complete sentences particularly given that they are structured grammatically that way

fixed  
  
P5 'from different data models,' -> from two different data models:

fixed  
  
P6 "the survey by [38] -> I found this sentence very long

broken up  
  
P6 The Distance-based -> The distance-based

fixed  
  
P6 "are far and isolated' -> do you mean far away?

Yes, rewritten  
  
P8 shows example -> shows an example

fixed  
  
Tables 2 & 3 are not particularly consistent (e.g., last row & last column do not have ..., last row of 3 has WA...

made consistent. The only difference now is that Table 3 has WA because all its entries have WA. This is also noted in the caption.  
  
p9 Fix a set... -> this sentence reads a bit oddly to me

rewritten (see Section 3.2)  
  
You often write Wigan Athletics, but I think it is Wigan Athletic

thank you Fixed  
  
P13 Therefore, Bayesian -> Therefore, the Bayesian

fixed  
  
P16 I am not sure I'd say strengths and weaknesses because it focuses on the other metrics, and does not illustrate weaknesses of your metric. Nonetheless I find this section very very interesting and illuminating. 

rewrote a bit to avoid misleading expectations

P19 to empirical -> to the empirical

fixed  
  
P19 defined a single -> a word seems to be missing here

changed to “defined over a single”  
  
P19 needs a \noindent after eq 9

fixed  
  
P19 by by

fixed  
  
P25 AUC is, 1 -> AUC is 1

rewritten to “The maximum value for $\auc$ is 1.0”  
  
P25 of normal community -> sentence is a bit unclear to me 

rewritten to “When comparing two individuals with only two binary features, $\it{RIBL}$ assigns to each individual, the same distance from the center of a normal population.  
P27 if you could cite to some article that states that Dzeko is a good dribbler that would be cool & strengthen your argument .”   
  
P30 than normal -> than the normal

fixed  
  
P30 in Premier -> in the Premier

fixed   
  
Figure 11, can you say what the salary and time scales are? E.g., is the salary per game, per year, per week?

Per year. This is now indicated in the paper.  
P34 in Forwarder -> in the Forwarder

fixed  
  
Figure 12 has no caption and same point about time units

added thank you  
  
In the appendix having some text before you jump into the proofs would be good

added transition text  
  
Cite 56 seems off

all citations were reviewed and edited, including that one  
  
Reviewer #2: This revision is a major rewrite and fixes the main issues with the previous versions. In particular, the paper now takes the angle of presenting a framework for exception mining in relational data. The ELD measure is an additional contribution, one which is evaluated alongside likelihood ratio (LR) and log-likelihood (LOG). The theoretical analysis showing the relationship between ELD and TVD is also nice, as is the discussion of when LOG works better for detecting success. I think it's interesting that LOG works well for detecting success but very poorly for differentiating classes.   
  
The remaining issues are minor, such as copy editing details and some clarifications.   
  
I think that some of the text is unnecessarily verbose, but it's probably safer to err on the side of too much detail.

We tightened up the writing to avoid repetitions, fill words, be more concise. At the same time the reviewers have asked us to add some extra material and clarifications. The net result is that the total number of pages is unchanged.  
  
The example in Section 5.2 shows that LR has both positive and negative terms, but it doesn't show meaningful consequences of that difference. Both of them report a positive value for the outlier, but they're on different scales, so it's hard to compare. It would help to show a pair of examples, x1 and x2, where LR ranks x1 as more unusual and ELD ranks x2 as more unusual.

Figure 8 gives an example where LR gives the same value for a normal and an outlier, but FD (and hence ELD) distinguishes them. See disussion around “While log-likelihood LR is a good baseline score for detecting outliers, it fails to detect some clear outliers, as shown in Figure 8” In general, the strength of ELD is not best shown in single cases, but that it provides scores that are further apart, making it easier to set a threshold to separate a set of outliers from a set of normals, rather than distinguish a particular case.

In Section 6, what does the second-order term look like for the Taylor expansion of ELD? It's interesting that TVD is a first-order approximation of ELD, but the second-order term could change its behavior substantially. Since LR is being analyzed via a second-order Taylor expansion, it seems appropriate to at least look at the second-order term for ELD.

Added an explicit formula for the second-order term (see Proposition 2)

In Section 8.2, I didn't quite follow the explanation of why RIBL fails in Table 10. This is minor, because I imagine that studying the details of RIBL from the original paper would clarify it.

We simplified and clarified the explanation.   
  
In Section 10, the justification for focusing on ELD is reasonable -- LR performs worse, and LOG only occasionally performs better. However, it would help to have some mention of this early on in the section, instead of just at the end.

Actually alternative metrics, such as LR and LOG, are mentioned in Section 10.1 at the beginning with a forward pointer to 10.3. We’ve expanded the forward pointer at the end of Section 10.1 so it’s not missed by the reader.  
  
TYPOS, GRAMMAR, AND OTHER SENTENCE-LEVEL ISSUES:   
  
(NOTE: This isn't necessarily a comprehensive list, but includes the ones I happened to notice and mark while reading through.)   
  
Abstract: "A object" --> "An object" .

fixed  
  
P.2: "each team comprises a set of players for that match" -- this sounds a bit odd; normally a team includes all players, even those that didn't play in a particular match. Is there a different word or phrasing that could be used here?

Changed to “fields a set of players”  
  
P.3: "relational data[21,50,13]" --> "relational data [21,50,13]"

fixed  
P.3: "model fits fits an input"

fixed  
P.5: "field, for a survey please see" --> "field; for a survey please see" (looks like a comma splice to me, but very minor)

fixed  
P.6: "The Distance-based" --> "The distance-based"

fixed  
P.8: "grounding count, divided" --> "grounding count divided" (I'd remove this comma, but that's very minor)

fixed  
P.12: "For the notations" --> "For the notation"

fixed  
P.13: "Therefore, Bayesian network" --> "Therefore, the Bayesian network"

fixed  
P.19: "can omit this section without loss of generality" -- I don't think the expression "without loss of generality" makes sense here

changed to “loss of continuity”  
P.19: "defined a single discrete random variable" --> "defined over a single discrete random variable"

fixed  
P.21: "wolrd" --> "world"

fixed  
P.21: "data set" "dataset" -- dataset can be one word or two, but it should be consistent

now always “data set”  
P.25: "are the only scores that achieves" --> "are the only scores that achieve"

fixed  
P.26: "works better than RIBL or equally good" --> "... equally well"

fixed   
P.26: "with precision@r% another metric" --> "with precision@r%, another metric"

fixed  
P.28: "belog to forward group" --> "belog to the forward group"

rewritte to be “players are forwards”  
  
  
  
OTHER PRESENTATION ISSUES:   
  
Capitalization is somewhat inconsistent. When discussing a method that also has an associated abbreviation (e.g., "Bayesian networks"), I recommend capitalizing the proper nouns (e.g., Bayes, Bayesian, Markov) and leaving all other letters lowercase. This paper uses a mix of capitalization strategies. Some words follow this approach ("Bayesian network", "Kullback-Leibler divergence") and others do not ("Exceptional Model Mining", "Markov Logic Networks"). I imagine the final style is determined by the journal, but it should be consistent throughout in any case.

We switched to lower case except for proper nouns, as you suggest  
  
Figure 4 and 5 have inconsistent typefaces.

Fixed (both Calibri now)  
  
Table 5 references Figure 7. Could Figure 7 come first?

It does (Figure 7 on p.17, Table 5 on p. 18)  
  
The text in the graphs is mostly readable, but could be even larger. (e.g., Fig.11)

We leave this to the final typesetting, also looking for advice from the journal.  
  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
  
STYLE FILE INFORMATION:   
  
Questions on submitting Tex files:   
  
What submission item should I use for a Tex file? The best submission item to use for a tex file is a "Manuscript" or similar worded submission item description.   
  
What if my Tex file(s) doesn't build? First, look in the PDF to see if an error message is available. If the error message suggests you are missing a sty file, then this may be the cause that the Tex submission will not build. If your submission still is not building after attempting to fix your Tex file from the suggested error message, you may want to verify the following:   
  
Are your images referenced correctly? Images cannot be referenced in subfolders, otherwise they will not appear in the PDF   
  
Are all your accompanying files referenced correctly in your Tex file?   
  
Why can't I submit Tex files and DVI files together? The System will accept either a Tex file(s) \_OR\_ a DVI file. If your DVI file does not build, then it is suggested you instead submit a Tex file. If you submit a DVI file and it does not build correctly, then you will need to provide the Tex file in order for the building problem to be researched.    
  
What if my figures are not appearing? It may be possible that your images are referenced in subfolders. Images cannot be referenced in subfolders, otherwise they will not appear in the PDF. One example of a correctly referenced image is: epsfig{figure=alld.eps,width=.5 extwidth} . An example of an incorrectly referenced image is: epsfig{figure=images/alld.eps,width=.5 extwidth}   
  
What if I view large or cut off EPS images? If you view large or cut off images, then you will need to resize the image to fit on one 8.5 x 11 page. When there is a problem caused by not resizing PostScript files (the images get cut off…) the author will either need to resize the images or save the files in a format that EM can recognize as an image.( Perhaps as a TIFF or a JPEG). Please note that EPS files are the best choice for image files in Tex submissions.   
  
Do I need to use my STY files too? Yes. The sty files will be needed to properly build your submission.   
What if I see (?) Question marks in my PDF? If you see question marks in the references of your PDF, most likely your Tex file(s) are in subdirectories. Tex submissions cannot include subdirectories for your submission to properly build. All associated files must be in one directory for the submission to build.   
  
Please note:   
1) The best image/figure file types for Tex submissions are EPS files. The author should attempt to submit EPS images for all their figure files. The author should verify that their EPS images are not in subfolders, otherwise the files may not build into the submission   
  
2) The author should look at the error generated at the end of their Tex submission. The author should verify the format of their Tex submission, as the error message suggests the Tex file may be corrupt.   
  
3) The author should re-verify that the submission and all of their accompanying Tex related files (STY files, etc) are uploaded to the submission   
  
4) The author should then re-build their submission.   
  
5) The author verify their images are not in subfolders, otherwise the images may not appear.   
  
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>   
  
TeX Web Resources   
  
There has been a significant increase in the number of TeX submissions to journals using Editorial Manager. While we do not offer direct technical support for TeX, just as we don't offer direct technical support for Microsoft Word, we have compiled a list of TeX-related web sites for journals to use, but please do feel free to distribute this information to your authors if you deem it helpful.   
  
Helpful TeX Links:   
  
Beginners guide to TeX   
This excellent introduction to TeX contains links to a basic explanation of TeX, a more thorough overview, and FAQs. You'll also find user help, documentation, sample documents, and a list of recommended reference books.   
<http://www.tug.org/begin.html>   
  
The Comprehensive TeX Archive Network   
If you know absolutely nothing about TeX and would like to learn about what TeX is and where it came from, be sure to take a look at the article entitle "What it TeX?" There is a search function for files and documentation on the site as well as links to sign up for TeX users groups and announcements lists.   
<http://www.ctan.org/>   
  
TeX Guides   
An excellent resource offering a variety of TeX guides including guides for Mathematical Symbols in TeX and TeX for Word Processor Users.   
<http://www.mcs.vuw.ac.nz/~david/latex/>   
LaTeX Encyclopedia   
An online LaTeX "encyclopedia". The site contains a table of contents with links to information on documentation, installation, typography, and a Navigator for the site.   
<http://tex.loria.fr/>   
LaTeX Math Guide   
The American Mathematical Society's Short Math Guide for LaTeX.   
<ftp://ftp.ams.org/pub/tex/doc/amsmath/short-math-guide.pdf>   
  
  
  
  
Recipients of this email are registered users within the Editorial Manager database for this journal. We will keep your information on file to use in the process of submitting, evaluating and publishing a manuscript. For more information on how we use your personal details please see our privacy policy at <https://www.springernature.com/production-privacy-policy>. If you no longer wish to receive messages from this journal or you have questions regarding database management, please contact the Publication Office at the link below.   
  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   
In compliance with data protection regulations, you may request that we remove your personal registration details at any time. (Use the following URL: <https://www.editorialmanager.com/dami/login.asp?a=r>). Please contact the publication office if you have any questions.