Author’s response to reviews

Title: Mining Adverse Drug Reactions from Online Healthcare Forums using Hidden Markov Model

Authors:

Hariprasad Sampathkumar (hariprsd@ku.edu)
Xue-wen Chen (xwchen@wayne.edu)
Bo Luo (bluo@ku.edu)

Version: 5  Date: 4 August 2014

Author’s response to reviews: see over
From: Hariprasad Sampathkumar, Xue-wen Chen and Bo Luo

Date: 3 Aug 2014

To:

Editor

BMC Medical Informatics and Decision Making

RE: Revision of manuscript for 'Research' article (MS: 1098530775116882) titled “Mining Adverse Drug Reactions from Online Healthcare Forums using Hidden Markov Model”

We would like to thank the Associate Editor and the Reviewers for reviewing our manuscript and giving us another opportunity to revise it. In this revised manuscript, we have updated the paper to address the minor essential revisions in response to the Reviewer's comments. Through this revision we have updated the abstract to include a brief summary of the component analysis results along with varying dictionary sizes for the HMM classifier. We have also tried to address display issues with some of the figures. We provide detailed responses to the each of the individual Reviewers' comments as a follow up.

We thank you again for giving us an opportunity to contribute a part of the BMC Medical Informatics and Decision Making Journal and look forward to hearing from you regarding the review process.

Sincerely,

Hariprasad Sampathkumar
Xue-wen Chen
Bo Luo
Detailed Responses to Reviewer's comments

Comments from Reviewer 1:
Reviewer: Samir Abdelrahman

Reviewer's report:

Minor essential revision..

Thanks authors for this piece of work. Really, I am very satisfied and appreciated your efforts.

Authors' Response: We are glad that we have been able to address all the concerns and would like to thank the Reviewer for giving us the opportunity and guidance to help present our work in a better way.

However, I suggest to put a couple of sentences in the Abstract results subsection. These sentences emphasis the component analysis results; which component with size is the most effective and which is the least. These debates may be only for HMM classifier "no need for baseline if the abstract maximum length will be a concern".

Authors' Response: As suggested by the Reviewer we have updated the Results subsection of the Abstract to include the results of the component analysis and the variations in the dictionary sizes. To keep in line with the word limit of the Abstract, we have included text only in reference to the HMM classifier, also, as suggested by the Reviewer. The following is the updated text in the Results subsection of the Abstract:

Results: A 10-fold cross-validation on the manually annotated dataset yielded on average an F-Score of 0.76 from the HMM Classifier, in comparison to 0.575 from the Baseline classifier. Without the Plain Text Filter component as a part of the Text Processing module, the F-Score of the HMM Classifier was reduced to 0.378 on average, while absence of the HTML Filter component was found to have no impact. Reducing the Drug names dictionary size by half, on average reduced the F-Score of the HMM Classifier to 0.359, while a similar reduction to the side-effects dictionary yielded an F-Score of 0.651 on average. Adverse side-effects mined from www.medications.com and www.steadyhealth.com were found to match the Adverse Drug Reactions on the Drug Package Labels of several drugs. In addition, some novel adverse side-effects, which can be potential Adverse Drug Reactions, were also identified.

Please also check the figures through the pdf. Some figures were not seen well--Figure 10 as an example.

Authors' Response: We have updated the Figures 1, 8 and 10 to make them display better in the pdf.

Thanks again

Authors' Response: We would also like to thank the Reviewer again for giving us an opportunity to revise our manuscript.