

Review Form 1.7

Journal Name:	Asian Journal of Research in Computer Science
Manuscript Number:	Ms_AJRCOS_110817
Title of the Manuscript:	Application Research on Semantic Analysis using Latent Dirichlet Allocation and Collapsed Gibbs Sampling for Topic Discovery
Type of the Article	

PART 1: Review Comments

	Reviewer's comment	Author's comment (if agreed with reviewer, correct the manuscript and highlight that part in the manuscript. It is mandatory that authors should write his/her feedback here)
Compulsory REVISION comments 1. Is the manuscript important for scientific community? (Please write few sentences on this manuscript) 2. Is the title of the article suitable? (If not please suggest an alternative title) 3. Is the abstract of the article comprehensive? 4. Are subsections and structure of the manuscript appropriate? 5. Do you think the manuscript is scientifically correct? 6. Are the references sufficient and recent? If you have suggestion of additional references, please mention in the review form. <u>(Apart from above mentioned 6 points, reviewers are free to provide additional suggestions/comments)</u>	<p>1. The manuscript uses LDA with Gibbs sampling on Nigerian newspapers' sports, entertainment, politics, and health sections for discovery and semantic analysis (knowing which topics are being discussed and also discover word counts and important topic keywords. Manuscript may serve interesting to people wanting to get an idea of what type of topic emerge in the Nigerian news.</p> <p>2. Abstract states “One of the main problems of LDA is that the topics extracted are of poor quality if the document does not coherently belong to a single topic. However, Gibbs sampling operates on a word-by-word basis, which allows it to be used on documents with a variety of topics and modifies the topic assignment of a single word.” Is only partially true. Gibbs sampling is just an inference procedure (e.g., you can use variational inference instead of Gibbs sampling to obtain model parameters). Gibbs sampling or any other inference e.g., variational inference doesn't help with noisy document topics.</p> <p>3. Abstract summarizes the work.</p> <p>4. Subsections are ok.</p> <p>5. Manuscript doesn't seem to have any incorrect information; however the work should provide some details, e.g., which convergence criteria were ultimately used by the authors for producing the final topics. III.C.iii</p> <p>6. References are broad and generic. Since this work is focused on employing LDA for topic discovery on a different dataset, maybe works (e.g., Dahal et al., 2019; Huang et al.,2014)) can be more relevant and may be added.</p> <p>Dahal, Biraj, Sathish AP Kumar, and Zhenlong Li. "Topic modeling and sentiment analysis of global climate change tweets." <i>Social network analysis and mining</i> 9 (2019): 1-20.</p> <p>Huang, Siqi, Yitao Yang, Huakang Li, and Guozi Sun. "Topic detection from microblog based on text clustering and topic model analysis." In <i>2014 Asia-Pacific Services Computing Conference</i>, pp. 88-92. IEEE, 2014.</p>	
Minor REVISION comments 1. Is language/English quality of the article suitable for scholarly communications?	English is correct.	
Optional/General comments		

[Review Form 1.7](#)

PART 2:

	Reviewer's comment	Author's comment <i>(if agreed with reviewer, correct the manuscript and highlight that part in the manuscript. It is mandatory that authors should write his/her feedback here)</i>
Are there ethical issues in this manuscript?	<i>(If yes, Kindly please write down the ethical issues here in details)</i>	

Reviewer Details:

Name:	Arjun Mukherjee
Department, University & Country	University of Houston, USA