

# Trends on Health in Social Media: Analysis using Twitter Topic Modeling

Mohsen Asghari, Daniel Sierra-Sosa, Adel Elmaghraby

**Abstract—** There is a growing interest on social networks for topics related to Healthcare. In particular, on Twitter, millions of tweets related to healthcare can be found. These posts contain public opinions on health, and allow to understand how is the popular perception on topics such as medical diagnosis, medicines, facilities, and claims. In this paper we present an adaptive system designed using 5 layers. The system contains a combination of unsupervised and supervised algorithms to track the trends of health social media. As it is based on a word2vec model, it also captures the correlation of words based on the context, improving over time, enhancing the accuracy of predictions and tweet tracking. In this work we focused on United States data and use it to detect the trending topics of each state. These topics are followed including new social network contributions. The supervised algorithm implemented is a Convolutional Neural Network (CNN) in conjunction with the Word2Vect model to classify and label new tweets, assigning a feedback to the topic models. The results of this algorithm present an accuracy of 83.34%, precision of 83%, recall 84% and F-Score of 83.8% when evaluated. Our results are compared with two state of the art techniques demonstrating an advantage that can be leveraged for further improvements.

**For the published version of record document, go to:**  
**<http://dx.doi.org/10.1109/ISSPIT.2018.8642679>**









