



VCU

Virginia Commonwealth University
VCU Scholars Compass

Graduate Research Posters

Graduate School

2020

Relevance of Health-Related Hashtags on Twitter: A Text Mining Approach

Mauli Dalal

Virginia Commonwealth University

Kweku-Muata Osei-Bryson

Virginia Commonwealth University

Follow this and additional works at: <https://scholarscompass.vcu.edu/gradposters>



Part of the [Management Information Systems Commons](#)

Downloaded from

Dalal, Mauli and Osei-Bryson, Kweku-Muata, "Relevance of Health-Related Hashtags on Twitter: A Text Mining Approach" (2020). *Graduate Research Posters*. Poster 35.

<https://scholarscompass.vcu.edu/gradposters/35>

This Poster is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Graduate Research Posters by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

Relevance of Health-Related Hashtags on Twitter: A Text Mining Approach



Mauli Dalal, M.S., PhD Candidate, Kweku-Muata Osei-Bryson, PhD

Department of Information Systems, School of Business, Virginia Commonwealth University, Richmond, VA

ABSTRACT

Intro:

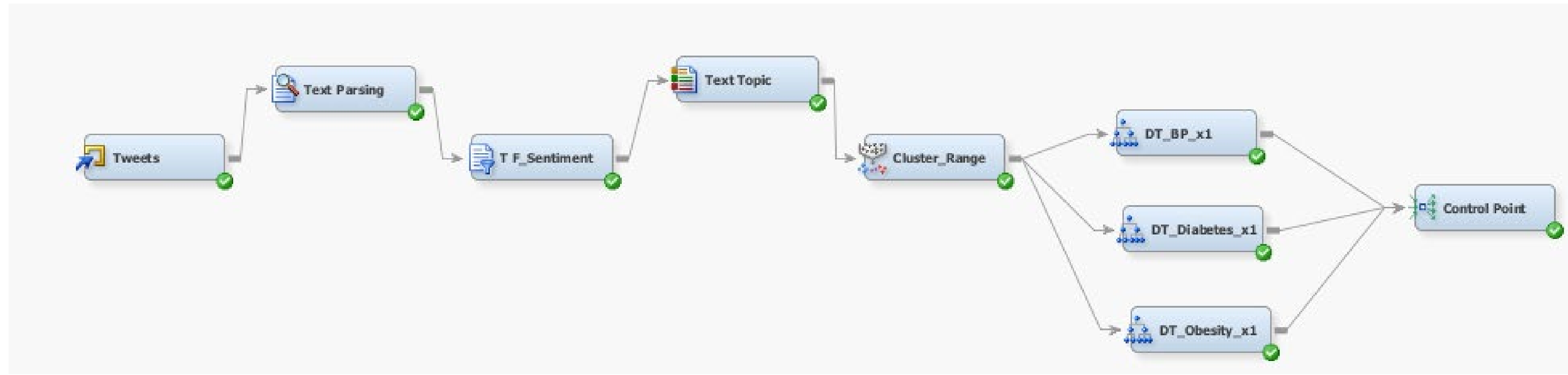
- Social media platforms facilitate user interaction and impact decision making.
- Users prefer to use hashtags while sharing posts.
- Knowing the sentiment towards **diabetes, bloodpressure, and obesity** is fundamental to understanding the impact of these information on patients and their families.
- The study seeks to determine the relevance of health-related hashtags on Twitter and analyze sentiments about diabetes, obesity, blood pressure.

Methods:

- Tweets retrieval using #Hashtags
- Overall 10,881 tweets retrieved
- 3 Analytical Approaches:
 - Text Topic Modeling
 - Clustering
 - Decision Trees (DTs)

Results:

- The important topics identified vary across clustering and DTs.
- Using '10 crossfold validation' for each DT with 'misclassification rate' demonstrated good accuracy



Topic Modeling identified multi-term topics of relevance to:

- Blood Pressure (BP)
- Diabetes
- Obesity

Clustering node captured broader range of topics as compared to each disease specific DT.

Health related information - checkups, remedies, prevention, surgical procedures are actively discussed on such platforms.

Important Topics for Clustering & DT

Clustering_Range	DT_BP_x1	DT_Diabetes_x1	DT_Obesity_x1
Diabetes-Nutrition	BP-Remedy	Diabetes-Nutrition	Healthcare-Burden
Diabetes-Risk-Woman	BP-Prevention	Different-Diabetes-Type	BP-Remedy
Health-Awareness	Lung-Accumulation	Diabetes-Checkup	Food-Prevent
Bariatric-Surgery-Woman	Food-Prevent	BP-Obesity	Lung-Accumulation

CONCLUSION

- Combination of different data mining and text mining approaches can enhance our understanding about hashtag relevance on social media platforms.
- It can thus increase our understanding about user engagement on such platforms and potentially help improve managing public health strategically.

Contact Information: Mauli Dalal

Email: dalalmb@mymail.vcu.edu