Mining HIV trends in Social Media Data

Patrick Breen
The Institute of Bioinformatics
The University of Georgia
Athens, Georgia
pbreen@uga.edu

Shannon Quinn*
Institute of Computer Science
University of Georgia
Athens, Georgia
squinn@cs.uga.edu

Table 1: Cosine Similarity to 'prep'

Related Word	Cosine Similarity to 'prep' heighttruvada
0.836507 heightbenegative	0.809185
charliesheen	0.805214
hivtestweek	0.797361
sexwork	0.774867
gettested	0.765772
hiv	0.765389
icasa2015	0.749948
doingit	0.749230
aids	0.738572

Table 2: Cosine Similarity to 'truvada'

Related Word	Cosine Similarity to 'truvada'
truvada	0.836507 heightbenegative
0.809185	
charliesheen	0.805214
hivtestweek	0.797361
sexwork	0.774867
gettested	0.765772
hiv	0.765389
icasa2015	0.749948
doingit	0.749230
aids	0.738572

ABSTRACT

Here is the abstract.

Keywords

Social Media, Topic Modelling, Document Embedding

1. INTRODUCTION

Introduce and literature review of: 1) PrEP, HIV, Twitter
2) Word2Vec, Doc2Vec
3) Dynamic Topic models, Latent Dirichlet allocation

2. RESULTS

High level, but more specific than intro, introduce the approach and methods used in this paper here.

2.1 Word and Document Similarity

2.2 Time Domain

2.3 User Timeline Analysis

2.3.1 Sentiment Classification

Figure 1: Time domain plot.

3. CONCLUSIONS

Conclusions go here.
Example citation (needed right now to compile):[1]

4. ACKNOWLEDGMENTS

Acknowledgements go here.

5. REFERENCES

[1] L. Lamport. *LaTeX User's Guide and Document Reference Manual*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1986.

5.1 References

*Corresponding author.

WOODSTOCK '97 El Paso, Texas USA

DOI: 10.475/123_4