Enhancing Transparency in Predicting Private Traits from Digital Records of Human Behavior

Daizhuo Chen
Columbia Business School

Samuel Fraiberger
NYU

Robert Moakler
NYU Stern School of Business

Foster Provost
NYU Stern School of Business

Wally Wang
NYU Stern School of Business

December 9, 2013

Abstract

Blah

1 Preliminary Results

Blah

1.1 Data Description

Blah

1.2 Feature Space

Blah

References

[1] Sundararajan A., Provost F., Oestreicher-Singer G., Aral S., *Information in digital, economic and social networks*, Information Systems Research, 2013

[2] Bachrach Y., Kohli P., Kosinski M., Stillwell D., Graepel T. *Personality and Patterns of Facebook Usage*, Web Science'12, June 22-24, 2012

[3] Butler D., *Data Sharing Threatens Privacy*, Nature, 449(7163):644, 2007

[4] Kosinski M., Stillwell D., Graepel T., *Private Traits and Attributes are Predictable from Digital Records of Human Behavior*, Proceedings of the National Academy of Sciences, 2013

[5] Macskassy S.A. and Provost, F.J. *Classication in Networked Data: A toolkit and a univariate case study* The Journal of Machine Learning Research 8, 935-983, 2007

[6] Schwartz HA, Eichstaedt JC, Kern ML, Dziurzynski L, Ramones SM, et al. *Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach* PLoS ONE 8(9): e73791. doi:10.1371/journal.pone.0073791, 2013