

The Evolution of Accounting Skills: An Introduction to Topic Modeling

Overview:

“Big Data” and “Data Analytics” are relatively new themes that seem poised to revolutionize the fundamental accounting tasks of adding structure to data and organizing information. In the near future, accountants without data analytics skills risk being shut out of the labor market. This workshop session introduces some data analytics concepts – in a kind and gentle manner – and uses an interactive exercise to build a topic model to complete a classification task. Ultimately, participants will gain insight into how computer-aided techniques can be used to support accountants while being encouraged to acquire the underlying skills themselves.

Workshop Session Facilitator:

Andrew M. Bauer

University of Waterloo

ambauer@uwaterloo.ca

Requirements:

To participate, **a tablet or computer is strongly encouraged**. In particular, “readings” are provided for download (only a few copies will be available on paper). No other advanced knowledge or software is required to participate, however.

Please see the workshop link or participant email for downloadable materials used in hands-on exercises. Other materials used during the session can be provided also, after completion.

Pre-Workshop Instructions:

- a) Read through the 15 articles provided. Each one is no more than a page or two.
- b) While reading, make brief notes about what you think each one is about.
- c) The entire task should take 20-30 minutes, depending on how quickly you read. Please bring your electronic readings and notes to the workshop.

Additional (Voluntary) Materials:

If you’re interested in learning more about data analytics, computer science, or computer programming – or interested in more data-driven readings – consider the following links or publications.

Online courses/specializations:

[Coursera](#)

University of Illinois: [Data Mining Specialization](#)

University of Michigan: [Applied Data Science with Python Specialization](#)

John Hopkins University: [Data Science Specialization](#)

[Udacity](#) Various: [Data Analytics](#)

Various: [Data Science](#)

[Codecademy](#) Learn to code (self-paced) in various languages (Python, Ruby, Javascript, etc.)

More options are available with a simple online search!

Accounting academic publications/reviews:

Li, F. 2010. Textual analysis of corporate disclosures: A survey of the literature. *Journal of Accounting Literature* 29: 143-165.

Loughran, T., and B. McDonald. 2016. Textual Analysis in accounting and finance: A survey. *Journal of Accounting Research* 54(4): 1187-1230.

Accounting practitioner publications:

Bauer, A. M. 2017. Data analytics: A high-level introduction for accounting practitioners. *The Tax Adviser* 48(5): 366-369.

Ng, C., and M. Pan. 2017. Dynamic decision making through data visualization. *Pennsylvania CPA Journal* 88(1): 22-25.

Stippich, W. W., and B. J. Preber. 2016. *Data Analytics: Elevating Internal Audit's Value*. Altamonte Springs: Institute of Internal Auditors Research Foundation.

Tschakert, N., J. Kokina, S. Kozlowski, and M. Vasarhelyi. 2016. The next frontier in data analytics. *Journal of Accountancy* 222(2): 58-63.

Computer science publications:

Crossno, P., A. Wilson, T. Shead, and D. Dunlavy. 2011. Topicview: Visually comparing topic models of text collections. In *Tools with Artificial Intelligence (ICTAI), 1022 23rd IEEE International Conference*: 936-943.

Han, J., M. Kamber, and J. Pei. 2012. *Data mining: concepts and techniques*, 3rd ed. Watham: Elsevier.

Lu, Y., Q. Mei, and C. Zhai. 2011. Investigating task performance of probabilistic topic models: An empirical study of PLSA and LDA. *Information Retrieval* 14(2): 178-203.

Miller, T. W. 2014. *Modeling techniques in predictive analytics with Python and R: A guide to data science*, 1st ed. Upper Saddle River: Pearson.

General interest reading and websites:

<https://fivethirtyeight.com/features/technology-is-biased-too-how-do-we-fix-it/>

<http://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/#5ffbf9b2f0c8>

<http://www.floatingssheep.org/>