

July 18, 2022

Dissertation "Prediction methods for networked bibliographic data" by Rajmund Kleminski.

Review by Professor Nitesh Chawla, University of Notre Dame, USA

The dissertation by Kleminski focuses on science of science to delve into the questions of how scientific fields evolve, quality of scientific output and predicting citations and interest in a scientific topic that may not have been of interest in the past. This is a fascinating topic as the notion of how research develops, how certain topics stand the test of time, how and why certain topics have low interest in the beginning and then pick later, and how to predict such evolution of topics, including citation. The dissertation while addressing aspects of science of science makes several technical contributions, including an extension to loopy belief propagation for discrete data, evaluation of the quality of networked scientific papers, impact of citation network structures on topic identification, a method for predicting future citation of a publication, and preliminary work on recommending venues for a publication. This work has led to three journal publications and one conference. It is noteworthy that one journal publication in 2022 has already led to 14 citations.

The dissertation is very well written and structured. Kleminski through the various chapters clearly spells out his contributions, situates his work against the related work, and empirically and fundamentally demonstrates the scholarly impact of his work. In my opinion, I find his chapters 4 and 5 shed new interesting insights through intuitive methods about predicting potential topics for researchers to focus on early in their careers — this is compelling as it offers a possible pathway for early researchers to explore as they develop their respective careers. Chapter 5 is absolutely a delight not only for the methods that span text mining and network science, but also the expert evaluation of the results. The fact that there is bibliographic coupling to retrieve publication is a notable contribution in itself.

Kleminski has demonstrated strong scientific mindset, deep understanding of network science and machine learning, and wonderful exposition of research ideas. This dissertation contains a set of original solutions to interesting array of scientific problems, and through his writing Kleminski has demonstrated a command of methodologies and algorithms, as well as a strong grasp of related work. It is apparent from the writing and publications that the Kleminski can conduct independent scientific work.

Nitesh Chawla

Nitesh Chawla, PhD
Frank M. Freimann Professor of Computer Science & Engineering
Director, *Lucy Family Institute for Data and Society*
University of Notre Dame



574 631 1090
nchawla@nd.edu

Lucy Family Institute for Data & Society
384E Nieuwland Science Hall
Notre Dame, IN 46556 USA
+1 574-631-7095
Lucyinstitute.nd.edu