HYUKJUN GWEON, University of Waterloo
Automated Occupational Coding Using Machine Learning

Occupational coding refers to categorizing a survey respondent's text answer into one of hundreds of occupation codes. Instead of manual coding, we propose automated coding using machine learning. Automated coding is a challenging problem because answers usually consist of only a few words while there are hundreds of possible categories. Nevertheless, the use of machine learning approaches may be helpful for a fraction of records. We include simulation results using occupational data from a German agency.