Learning, Information Extraction and the Web^{*}

Tom M. Mitchell

Machine Learning Department
Carnegie Mellon University, USA
tom.mitchell@cs.cmu.edu

Abstract. Significant progress has been made recently in semi-supervised learning algorithms that require less labeled training data by utilizing unlabeled data. Much of this progress has been made in the context of natural language analysis (e.g., semi-supervised learning for named entity recognition and for relation extraction). This talk will overview progress in this area, present some of our own recent research, and explore the possibility that now is the right time to mount a community-wide effort to develop a never-ending natural language learning system.

^{*} Invited speakers at ECML/PKDD are supported by the PASCAL European network of excellence.