

# Machine Learning for Natural Language Processing (and Vice Versa?)

Claire Cardie

Department of Computer Science, Cornell University, USA  
cardie@cs.cornell.edu  
<http://www.cs.cornell.edu/home/cardie/>

Over the past 10-15 years, the influence of methods from machine learning has transformed the way that research is done in the field of natural language processing. This talk will begin by covering the history of this transformation. In particular, learning methods have proved successful in producing stand-alone text-processing components to handle a number of linguistic tasks. Moreover, these components can be combined to produce systems that exhibit shallow text-understanding capabilities: they can, for example, extract key facts from unrestricted documents in limited domains or find answers to general-purpose questions from open-domain document collections. I will briefly describe the state of the art for these practical text-processing applications, focusing on the important role that machine learning methods have played in their development.

The second part of the talk will explore the role that natural language processing might play in machine learning research. Here, I will explain the kinds of text-based features that are relatively easy to incorporate into machine learning data sets. In addition, I'll outline some problems from natural language processing that require, or could at least benefit from, new machine learning algorithms.