

Crowdsourcing using Mechanical Turk: Quality Management and Scalability

Panos Ipeirotis
Stern School of Business, New York University
New York, NY, USA
panos@stern.nyu.edu

Abstract

I will discuss the repeated acquisition of “labels” for data items when the labeling is imperfect. Labels are values provided by humans for specified variables on data items, such as “PG-13” for “Adult Content Rating on this Web Page.” With the increasing popularity of micro-outsourcing systems, such as Amazon’s Mechanical Turk, it often is possible to obtain less-than-expert labeling at low cost. We examine the improvement (or lack thereof) in data quality via repeated labeling, and focus especially on the improvement of training labels for supervised induction. We present repeated-labeling strategies of increasing complexity, and show several main results: (i) Repeated-labeling can improve label quality and model quality (per unit data-acquisition cost), but not always. (ii) Simple strategies can give considerable advantage, and carefully selecting a chosen set of points for labeling does even better (we present and evaluate several techniques). (iii) Labeler (worker) quality can be estimated on the fly (e.g., to determine compensation, control quality or eliminate Mechanical Turk spammers) and systematic biases can be corrected. I illustrate the results with a real-life application from on-line advertising: using Mechanical Turk to help classify web pages as being objectionable to advertisers. Time permitting, I will also discuss our latest results showing that mice and Mechanical Turk workers are not that different after all.

Biography

Panos Ipeirotis is an Associate Professor at the Department of Information, Operations, and Management Sciences at Leonard N. Stern School of Business of New York University. His recent research interests focus on crowdsourcing and on mining user-generated content on the Internet. He received his Ph.D. degree in Computer Science from Columbia University in 2004, with distinction. He has received two “Best Paper” awards (IEEE ICDE 2005, ACM SIGMOD 2006), two “Best Paper Runner Up” awards (JCDL 2002, ACM KDD 2008), and is also a recipient of a CAREER award from the National Science Foundation. He also blogs about crowdsourcing and Mechanical Turk from time to time, an activity that seems to generate more interest and recognition than any of the above.