

## **CROWDSOURCING: QUALITY MANAGEMENT AND SCALABILITY**

Panagiotis Ipeirotis

New York University  
44 West 4<sup>th</sup> Street, Ste 8-84  
New York, NY 10012  
e-mail: panos@stern.nyu.edu

### **ABSTRACT**

I will discuss the use of crowdsourcing for building machine learning models quickly and under budget constraints, with a focus on the case where humans are noisy and the of "labels" provided by humans for data items are imperfect. I will present strategies of managing quality in a crowdsourcing environment, showing in parallel how to integrate data acquisition with the process of learning machine learning models. I illustrate the results using real- life applications drawn from the field of online advertising. Time permitting, I will also discuss our latest results showing that mice and Mechanical Turk workers are not that different after all.