

The Smarter Crowd: Active Learning, Knowledge Corroboration, and Collective IQs

Thore Graepel
Microsoft Research
Cambridge, UK
thoreg@microsoft.com

Abstract

Crowdsourcing mechanisms such as Amazon Mechanical Turk (AMT) or the ESP game are now routinely being used for labelling data for machine learning and other computational intelligence applications. I will discuss three important aspects of crowdsourcing which can help us tap into this powerful new resource in a more efficient way.

When obtaining training data from a crowdsourcing system for the purpose of machine learning we can either collect all the training data in one batch or proceed sequentially and decide which labels to obtain based on the model learnt from the data labelled so far, a method often referred to as active learning. I will discuss which criteria can be used for selecting new examples to be labelled and demonstrate how this approach has been used in the FUSE/MSRC news recommender system projecttemporia.com to categorise news stories in a cost-efficient way.

Data obtained from crowdsourcing systems is typically plentiful and cheap, but noisy. The redundancy in the data can be used to improve the quality of the inferred labels based on models that take into account the reliability and expertise of the workers as well as the nature and difficulty of the tasks. I will present an algorithm for such a corroboration process based on graphical models, and show its application on the example of verifying the truth values of facts in the entity-relationship knowledge base Yago.

Finally, I will talk about some very recent results on the effects of parameters of crowdsourcing marketplaces (such as price and required track record for participation) on the quality of results. This work is based on methods from psychometrics, effectively measuring the IQ of the Mechanical Turk when viewed as a form of collective intelligence.

This is joint work with Ralf Herbrich, Ulrich Paquet, David Stern, Jurgen Van Gael, Gjergji Kasneci, and Michal Kosinksi.

Biography

Thore Graepel is a senior researcher at Microsoft Research Cambridge (MSRC), UK, and heads the Online Services and Advertising (OSA) research group. The OSA group conducts research in the area of applied machine learning with applications to online advertising,

online gaming, and web search. A common theme of Thores research is the idea of using the tremendous wealth of data Microsoft is generating through its services to improve those services based on data-driven computational intelligence. Thores research has impacted Microsoft products and services, e.g. through the TrueSkill ranking and matchmaking system in Xbox Live and Halo 3, as well as through his work with adCenter on the prediction of user behaviour. Before starting the OSA group together with Ralf Herbrich, Thore co-founded the Applied Games group at MSRC whose research is aimed at bringing adaptable artificial intelligence to computer games. Thore has a strong academic track record with over 50 peer-reviewed publications in machine learning and probabilistic modelling. His basic research focuses on inference in large scale probabilistic models and knowledge bases. He is also interested in classification, recommender systems, and crowdsourcing for machine learning.