The Knowledge Engineering Review, Vol. 24:4, 409–413. © Cambridge University Press, 2009

Abstracts of Recent PhDs

Qualified Predictions for Large Data Sets

Candidate: Harris Papadopoulos
Institution: Computer Science Department of Royal Holloway, University of London, England, UK
Supervisors: Alex Gammerman and Volodya Vovk
Year awarded: 2004
doi:10.1017/S0269888909990221

Abstract

The Inductive Confidence Machine (ICM) provides an alternative method to that of the Transductive Confidence Machine (TCM) for complementing the bare predictions produced by traditional machine-learning algorithms with measures of confidence. These measures give an indication of how 'good' each prediction is, which is highly desirable in risk-sensitive applications. The motivation behind the introduction of the ICM was to produce algorithms that overcome the computational inefficiency problems suffered by TCMs.

In this thesis, we study the ICM method, describing how it works and how it can be applied to different traditional machine-learning algorithms. More specifically we detail how we implemented the Nearest Neighbours and Neural Networks ICMs for pattern recognition, and the Ridge Regression and Nearest Neighbours Regression ICMs. The results obtained by our methods demonstrate that the accuracy of ICMs in terms of error percentage is comparable to both traditional methods and TCMs and that the confidence measures they produce are useful in practice. In addition, our time efficiency comparisons exhibit their huge advantage in this sector over TCMs. These properties make ICMs the most suitable choice for obtaining qualified predictions when dealing with large data sets.

Monitoring Large-Scale Multi-Agent Systems using Overhearing

Candidate: Gery Gutnik
Institution: Computer Science Department, Bar-Ilan University, Israel
Supervisor: Gal A. Kaminka
Year awarded: 2006
URL: `http://www.cs.biu.ac.il/~gutnikg`
doi:10.1017/S0269888909990233

Abstract

Overhearing is fast gaining attention as a generic method for monitoring open, distributed multi-agent systems. In such settings, agents' internal structure is not generally known to a monitoring agent, but overhearing does not require such knowledge. Instead, the monitoring agent uses the overheard routine communications as a basis for inference about the other agents.

Previous work on overhearing investigated an extensive set of techniques and implementations of overhearing. However, focusing mainly on its potential applications, those investigations often rely on assumptions related to the fundamentals of overhearing. In contrast, we dedicate our research to a comprehensive study of the fundamental building blocks that allow overhearing in the first place tackling those problematic assumptions. In particular, our study focuses on overhearing in large-scale multi-agent systems and addresses the specific challenges and limitations that characterize such settings.

The first overhearing building block, addressed by our research, is the representation of multi-agent conversations. Here, building on the insights gained from analyzing the strengths and weaknesses of the rather radical Petri net approaches introduced by previous work, we propose a novel representation technique especially suitable for overhearing. Furthermore, we show this representation to be more scalable than previous representations, and thus more appropriate for monitoring conversations in large-scale settings.

Next, we addressed the building block of conversation recognition—the process of identifying the actual conversation based on a sequence of overheard messages. Although conversation recognition is a key step in overhearing prior to any possible inference, it is often discarded by previous investigations. Our work addresses the challenges related to conversation recognition by first introducing a formal model of overhearing. Then, based on this model, we provide a skeleton algorithm for conversation recognition, and provide instantiations of it for lossless and lossy settings. Since in large-scale settings overhearing agent has to process large quantities of intercepted messages, we also analyze the efficiency of those algorithms in terms of their run-time complexity.

The final building block addressed is selective overhearing, that is, overhearing under the restriction of