

Title

Data Mining Human Reasoning

Supervisor

John Lawrence

Project description

Successes in text mining in the past decade have delivered opinion mining and sentiment analysis which have become enormously commercially successful in PR and in predicting financial markets. The next step is to understand not just what opinions people hold, but why they hold the opinions they do – that is, to automatically extract the structure of human reasoning from text. This new area is known as argument mining and the group at Dundee is a world-leader in the field.

The student will work as a part of the Centre for Argument Technology, an internationally recognised team of interdisciplinary researchers. ARG-tech has built a reputation on balancing high quality theoretical research with robust and reliable software engineering. Its software is in use in thousands of schools, colleges and universities in over 80 countries around the world. It works with lawyers, magistrates, intelligence analysts, mediators, educators, programme-makers at the BBC, and data mining experts in IBM to explore deep challenges in artificial intelligence framed by the way in which real and very sophisticated arguments and debates are developed in demanding domains. Students with ARG-tech are treated as part of the team, and are invited to join both research activities, including weekly reading groups and discussions, and also to contribute to the groups general activities including demonstrations both at Dundee and elsewhere, blog posts, videos, and so on. The group has recently been contributing to an international effort to build the infrastructure for the Argument Web, a bold new vision of how debate and discussion will be conducted online in an environment of semantically rich, machine processable concepts.

The project will focus on techniques for automatically extracting arguments, debates, reasons, claims, conclusions, supports, counters and other argument-related units from free text using a variety of existing and novel computational linguistics techniques. The task is computationally very demanding and the student will be involved in cutting edge research. The prototype will be developed in the context of a large research project in collaboration with IBM, and may use data from the International Argument Mining Workshop *Unshared Task*.

Skills the student will develop

Experience with argumentation structures and analysis, with semantic web representation and processing technologies, including XML, RDF, OWL, scalable and federated database technologies.

In addition, experience techniques in computational linguistics, text mining and information extraction.

Industrial Relevance

We are working with users and major players in domains such as law, media and education, and the tools we are building are regularly demonstrated to groups within these domains. There are strong possibilities for the student to be exposed to potential employers through the work conducted on this project. Argument mining as a field is predicted to form a market worth over \$10bn by 2020.