Department	SCEBE
Supervisors	Dr Simon Wells
Project Title	Stereotypical Patterns of Human & Machine Reasoning for Explainable AI

PROJECT DESCRIPTION

There have been significant recent advances in the effectiveness of machine intelligence systems based upon neural networks, deep learning, and machine learning techniques. However questions have been raised about the nature of the reasoning occuring in these systems, how opaque that reasoning is, and how that impacts people's ability to trust the resulting systems. However patterns of human reasoning has been well studied over the last two millenia and a particularly useful approach has been that of argumentation schemes which capture stereotypical patterns of presumptive reasoning as used by people in the form of defeasible arguments.

This project will involve a detailed study of both existing machine intelligence systems and of argumentation schemes. The novel contribution will be an executable model of argumentation schemes and the development of a correspondence theory relating argumentation schemes to machine reasoning. This theory will be used to demonstrate how human and machine reasoning patterns are related, and how the symbolic and numeric approaches inherent to each can be augmented to facillitate the construction of explainable, justifiable, and composable AI systems.

Prospective applicants are encouraged to contact the Supervisor before submitting their applications. Applications should make it clear the project you are applying for and the name of the supervisor(s).

Academic qualifications

A first degree (at least a 2.1) ideally in Computer Science or Machine Learning. with a good fundamental knowledge of either Argumentation Theory or Machine Learning.

English language requirement

IELTS score must be at least 6.5 (with not less than 6.0 in each of the four components). Other, equivalent qualifications will be accepted. Full details of the University's policy are available online.

Essential attributes:
- Experience of fundamental Machine Learning techniques
- Competence in ML toolkits, e.g. Scikit-learn, Tensorflow, PyTorch, etc.
- Knowledge of Argumentation Theory
- Good written and oral communication skills
- Strong motivation, with evidence of independent research skills relevant to the project
- Good time management

Desirable attributes:

Understanding of topics in machine learning, computational argumentation, and defeasible reasoning would be advantageous.

Indicative Bibliography	Walton (2012) "Fundamentals of Critical Argumentation" Walton, Reed, and Macagno (2012) "Argumentation Schemes" Bishop (2007) "Pattern Recognition and Machine Learning" Goodfellow & Bengio (2017) "Deep Learning" Murphy (2012) "Machine Learning"
Enquiries	For informal enquiries about this PhD project, please contact s.wells@napier.ac.uk
Web page	https://www.napier.ac.uk/research-and-innovation/research-degrees/application-process