Hadoop Ecosystem for Machine Learning

Dr. Dengya (Simon) Zhu

DATE: Thursday, 29 November 2018
TIME: Starting at 10:30am
VENUE: Murdoch University, Building 245 Room 3.039
COST: Free, please RSVP to k.wong@murdoch.edu.au

This seminar is organised by the Joint IEEE Computational Intelligence Society and IEEE Robotics & Automation Society (WA Chapter).

Abstract:

Hadoop Ecosystem is now the de facto platform for Organisations and Government departments for big data applications. Components in the platform contains a variety of tools/packages/applications that facilitate the whole life of big data processing, such as data Extract Transform and Load, data wrangling, organise and store, machine learning, data mining, and data visualisation. While most of ecosystems are available from big cloud suppliers such Microsoft Azure, Amazon AWS, Google cloud, the cost of the Hadoop service is also unaffordable for most academic research groups. In this talk, Dr Simon Zhu will first briefly introduce a list of major components in a small scale Hadoop ecosystem constructed by his team. Then in the second part, he will show how the Hadoop cluster is used for distributed machine learning applications.

About the speaker:

Dr Dengya (Simon) Zhu is an honorary research fellow at the School of Engineering and Information Technology at Murdoch University and an adjunct research fellow at the School of Management, Curtin Business School, Curtin University. He is a data scientist at an Australia Government agent. He has a broad range of experience in government, industry, and academic research. Dr Zhu's research interests include big data, information retrieval, data mining, machine learning, natural language processing, sentiment analysis, open source software and software development. His research projects are usually practically oriented to address real world issues. He has published a number of papers in his research areas.