

Title: Artificial Intelligence for Data Analytics
Prof Chris Williams FRSE Emeritus Professor of Machine Learning
School of Informatics, University of Edinburgh

Abs: The practical work of deploying a machine learning system is often dominated by issues outside of training a model; this includes understanding the data set, data preparation and cleaning, model interpretation and explanation etc. The goal of the Artificial Intelligence for Data Analytics (AIDA) project at the Alan Turing Institute (UK) was to help to automate the whole data analytics process by drawing on advances in AI and machine learning. I will describe tools developed in AIDA to address such tasks, including identifying syntactic and semantic data types, data integration, and identifying and repairing missing and anomalous data.

More generally, automation in data science aims to facilitate and transform the work of data scientists, not to replace them. Important parts of data science are already being automated, especially in the modelling stages, where techniques such as automated machine learning (AutoML) are gaining traction. Other aspects are harder to automate, not only because of technological challenges, but because open-ended and context-dependent tasks require human interaction.

Refs:

AIDA: Artificial intelligence for data analytics

<https://www.turing.ac.uk/research/research-projects/artificial-intelligence-data-analytics>

Tijl De Bie, Luc De Raedt, Jose Hernandez-Orallo, Holger H. Hoos, Padhraic Smyth, Automating data science, Communications of the ACM 65(3) 76-87 (2022).