

## **Towards human-like communication in machine learning applications**

Peter Flach, University of Bristol

One question I consider in my work is how to design machine learning (ML) applications that people will trust and want to use. One aspect of trustworthiness is that the ML system needs to clearly communicate its operation and outputs to the user. When it comes to explaining outputs of ML systems to end-users the literature offers two dominant approaches: to present the user with a trace of the relevant features that have led to a particular target value or classification (such as the sequence of conditions leading from a decision tree's root to the relevant leaf); or to use weights to quantify and visualise the correlation (positive or negative) of the most important features with the target. While both approaches have value, neither can be expected to be fully successful in all circumstances or with all users, as neither takes the recipient of the information into consideration.

In contrast, I am proposing the use of storytelling techniques and narratives to improve communication from the ML system to humans. This requires constructing a pathway from mathematical structures representing the operation of the ML system, via logical and grammatical structures representing the intended meaning in an abstract sense, to concrete narratives expressed in natural language. The use of narratives is not limited to explaining the predictions of the ML system; I suggest that storytelling techniques can also be useful for machine learning researchers to make sense of empirical results obtained in experiments.

+-----+  
Peter A. Flach, Professor of Artificial Intelligence ph +44 117 9545162  
Intelligent Systems Lab, University of Bristol fax +44 117 9545208  
Editor-in-Chief, Machine Learning journal office +44 117 3315201  
Merchant Venturers Building, Woodland Road Peter.Flach@bristol.ac.uk  
Bristol BS8 1UB, United Kingdom http://www.cs.bris.ac.uk/~flach/