How can humans effectively team up with Machine Learning?

Machine Learning is advancing at a higher pace. It has been a powerful tool to humans in the sense that it can help us detect and predict events and situations, and consequently allows us to prepare for these events and situations. As powerful as it may be, several questions have been raised regarding the other side of the coin. These questions usually involve several issues such as bias, ethics, explainability, transparency, trust, etc.

While, as humans, it may be hard to be rigorous, accurate and make use of all data and adequate methods when predicting certain events, this is usually an easy and fast task for machines. On the other hand, humans are often better able to deal with questions related to bias and ethics. So why not putting humans in the loop of a process when it comes to using Machine Learning to solve some problems?
Teaming up with Machine Learning may as well be the best option we have. Can we find ways of expressing to the machine what we want? Will the machine be able to help us with solutions within the context we are considering? How can we, as humans, help the machine take into account issues such as bias or discrimination when making predictions, while being in the loop of the process? Which parts can the human and the machine do and what are the advantages and disadvantages of such chosen roles? In this Master topic, we aim at studying possible team mental models that allows us to have hybrid teams of both humans and Machine Learning algorithms.

Let's consider, as an example, a search and rescue scenario, where artificial agents may use Machine Learning techniques to create mental models of their human teammates and, consequently, predict their actions. Which Machine Learning methods can we use for this task? Which personal characteristics do we use to generate predictions (e.g., age, sex, race)? What possible ethical problems can this lead to? How can a human be integrated into the prediction process as to make it more accurate, less biased, more transparent, etc? Which parts will the agent have to explain? How does this change from method to method?

If you are interested in one or more of these topics, reach out to us! We'll set up a meeting to discuss your interests and what specific research topic would fit.

Recommended courses:

Machine Learning (I/II), Socio-Cognitive Engineering, AI Techniques

Supervisory team

This project will be coached by Myrthe Tielman (an assistant-professor at the Interactive Intelligence group), Carolina Jorge and/or Ruben Verhagen (both PhD students at the Interactive Intelligence group).

Contact

Contact M.L.Tielman@tudelft.nl (Myrthe Tielman)