Promises and Perils of Artificial Intelligence and Machine Learning: Disruption, Adoption, Dehumanization, Governance, Risk and Compliance

Valeria Sadovykh
University of Auckland
valeriasadovykh@gmail.com

Johnny Chan
University of Auckland
jh.chan@auckland.ac.nz

David Sundaram
University of Auckland
d.sundaram@auckland.ac.nz

In the last decade, Artificial Intelligence (AI) and Machine Learning (ML) have developed from peripheral technologies to dominant drivers of innovation. They are routinely used to recognize images; parse speech; respond to questions; make decisions; and replace humans.

Given that AI and ML tools are becoming a part of our everyday lives, it is critical that researchers and practitioners understand their state of art, adoption and influence. Improperly deployed AI and ML tools can violate privacy, threaten safety, and take questionable decisions that can affect individuals, organizations and ultimately society.

This mini track focus is on the promises and perils of AI and ML with a particular emphasize on (a) adoption, (b) disruption, (c) potential dehumanization, and (c) governance, risk and compliance mechanisms required to protect and enhance human wellbeing. We welcomed wide-ranging papers with qualitative and quantitative orientations; with theoretical and practical contributions; from personal, organizational and societal perspectives.

This mini track has been running for more than 2 years and continues to attract interesting and futuristic papers. This year we had a number of submissions with a diverse range of topics. The three selected papers investigate key issues that exist in the research area.

The first paper, *Machine Learning Systems in Clinics – How Mature Is the Adoption Process in Medical Diagnostics?* This paper touches on an important question of aging population and the ML tools maturity to provide a better diagnostic. Presently, with machine learning (ML), a technological change is taking place that could provide high quality healthcare and especially, improve efficiency of medical diagnostics in clinics. Specifically, the authors questioning whether ML needs to be deeply integrated in clinical routines which highly differs from the integration of previous health IT given the specific characteristics of ML. This paper describes an explorative qualitative study based on a conceptual basis consisting of the technological organizational-environmental framework (TOE) and the healthcare specific framework of non-adoption, abandonment, scale-up, spread, and sustainability (NASSS). By interviewing experts from clinics and their suppliers, the authors were able to integrate both frameworks and identify influencing factors specific to the adoption process of ML in medical diagnostics.

The second paper, *Acceptance of AI for delegating emotional intelligence: Results from an experiment* looks at AI based mechanisms for detecting human emotions.

URI: https://hdl.handle.net/10125/71380
978-0-9981331-4-0

Detecting emotions of other humans is challenging even for us humans, so how much more difficult it can be for technology? However, it is important in many social contexts and many individuals need and seek help in this regard. As technology is evolving, more and more AI based options emerge that promise to detect human emotions and support decision making. The research focuses on the idea of the full delegation of detecting emotions to AI. This study conducted an online scenario-based experiment in which participants have the choice to delegate emotion detection to another human in one group and to an AI tool in the other group. The results showed that the delegation rates are higher for a human, but surprisingly high for AI. The results provided insights that should be considered when designing AI-based emotion-detection tools to build trustworthy and acceptable designs.

The third paper, *Stakeholder-dependent views on biases of human- and machine-based judging systems* investigates a vital and highly concerning question of AI and ML biases. The authors discuss how talent management is an important business strategy, but inherently expensive due to the unique, subjective, and developing nature of each talent. Applying artificial intelligence to analyze large-scale data, talent intelligence management system (TIMS) is intended to address the talent management problems of organizations. While TIMS has greatly improved the efficiency of talent management, especially in the processes of talent selection and matching, high-potential talent discovery and talent turnover prediction, it also brings new challenges. Ethical issues, such as how to maintain fairness when designing and using TIMS, are typical examples. Through a Delphi study in a leading global AI company, this paper proposes eight rules to avoid fairness risks when designing TIMS.