

Abstract:

Added Norwegian abstract

Chapter 2: Reinforcement Learning, Explainability, and Interpretability

Fixed several typographic errors pertaining to differences in American vs British spelling, e.g., hierarchization -> hierarchisation, characterizing -> characterising, etc.

Chapter 3: Principles of Affinity-Based Reinforcement Learning

Added a missing full stop to the last sentence in Section 3.2

Chapter 5: Principles of Affinity-Based Reinforcement Learning

Fixed a typographic error in the caption of Figure 5.2: "a subset off" -> "a subset of"

References:

Updated status of paper "Towards Artificial Virtuous Agents: Games, Dilemmas and Machine Learning" to "publised"

Appended Papers:

Replaced the appended manuscript with the publised paper.