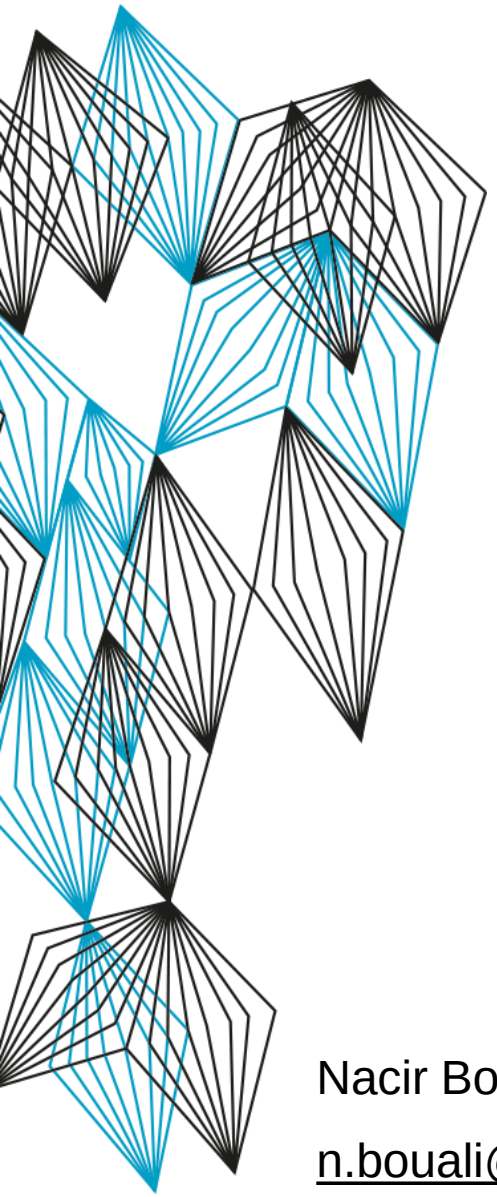

ARTIFICIAL INTELLIGENCE & CYBER SECURITY

REINFORCEMENT LEARNING MOTIVATION

Nacir Bouali

n.bouali@utwente.nl

UNIVERSITY OF TWENTE.

OBJECTIVES

- Motivation behind reinforcement learning
- Markov Decision Process
- Value Iteration
- Policy Iteration
- Direct Utility Estimation
- Adaptive Dynamic Programming
- Temporal Difference Learning
- Passive and Active Learning
- Exploration and Exploitation

REINFORCEMENT LEARNING

- Teaching an agent how to play a game such as chess requires an infinite amount of pairs (configuration, move) labelled by a chess master.
- 10^{40} configurations, each requiring a label.
- Not cost-effective (if at all possible).
- A different learning paradigm is needed in which the agent learns from the outcomes of the games
  - Win=good game policy
  - Lose= bad game policy

REINFORCEMENT LEARNING

- In reinforcement learning we want the agent to learn from rewards or reinforcements.

The question we wish to answer

Can an agent learn from punishment and reward to its actions?