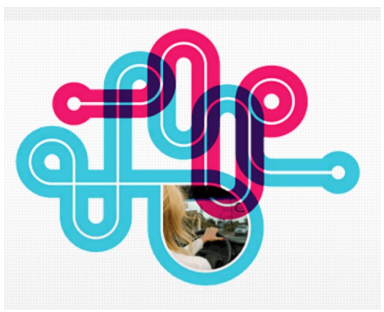
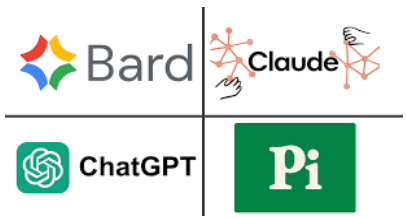
Online and Reinforcement Learning (OReL) 2026: Course Introduction

Mohammad Sadegh Talebi
sadegh.talebi@di.ku.dk
Department of Computer Science

AI Applications

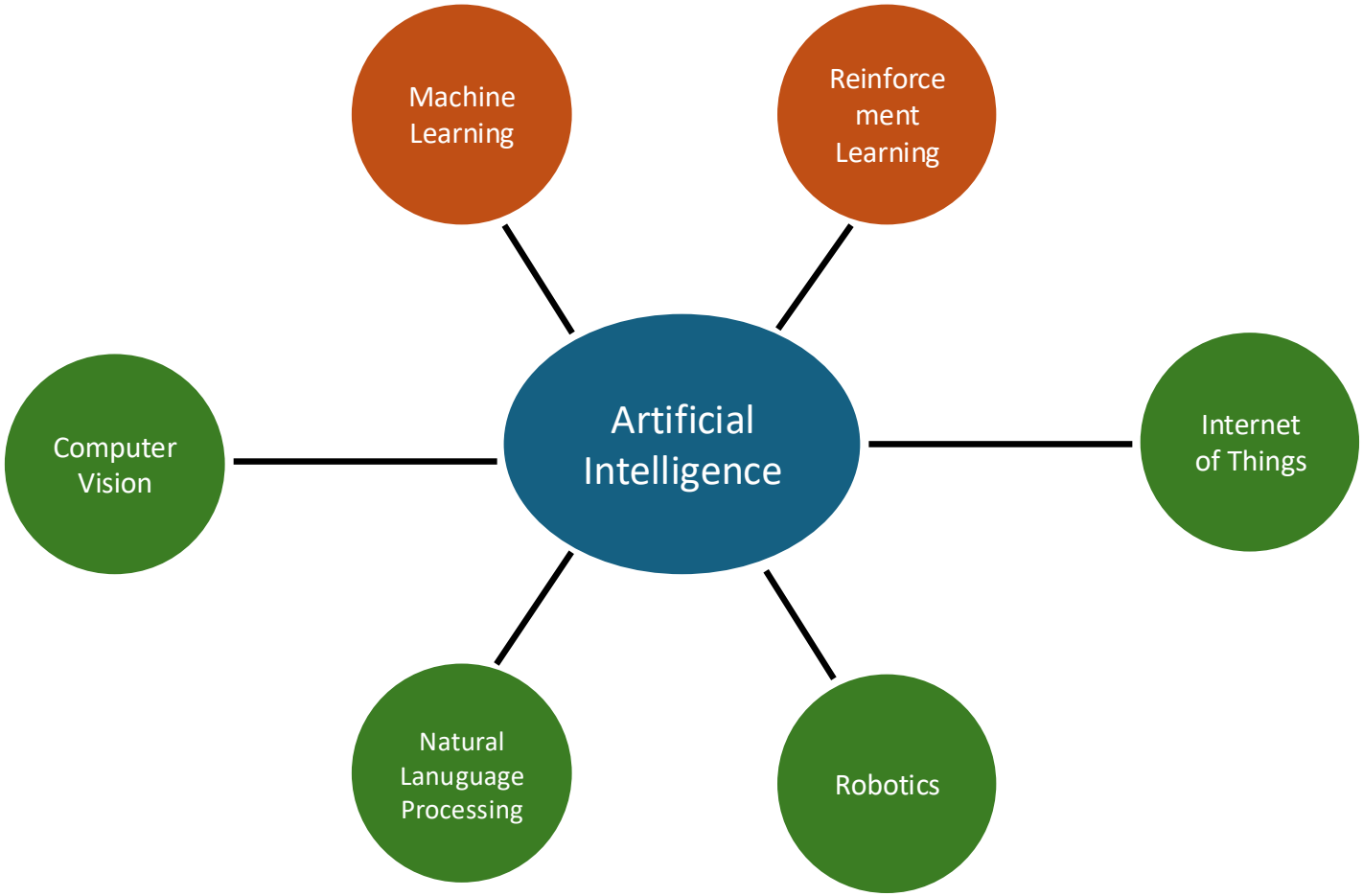
AI is everywhere in our real life:



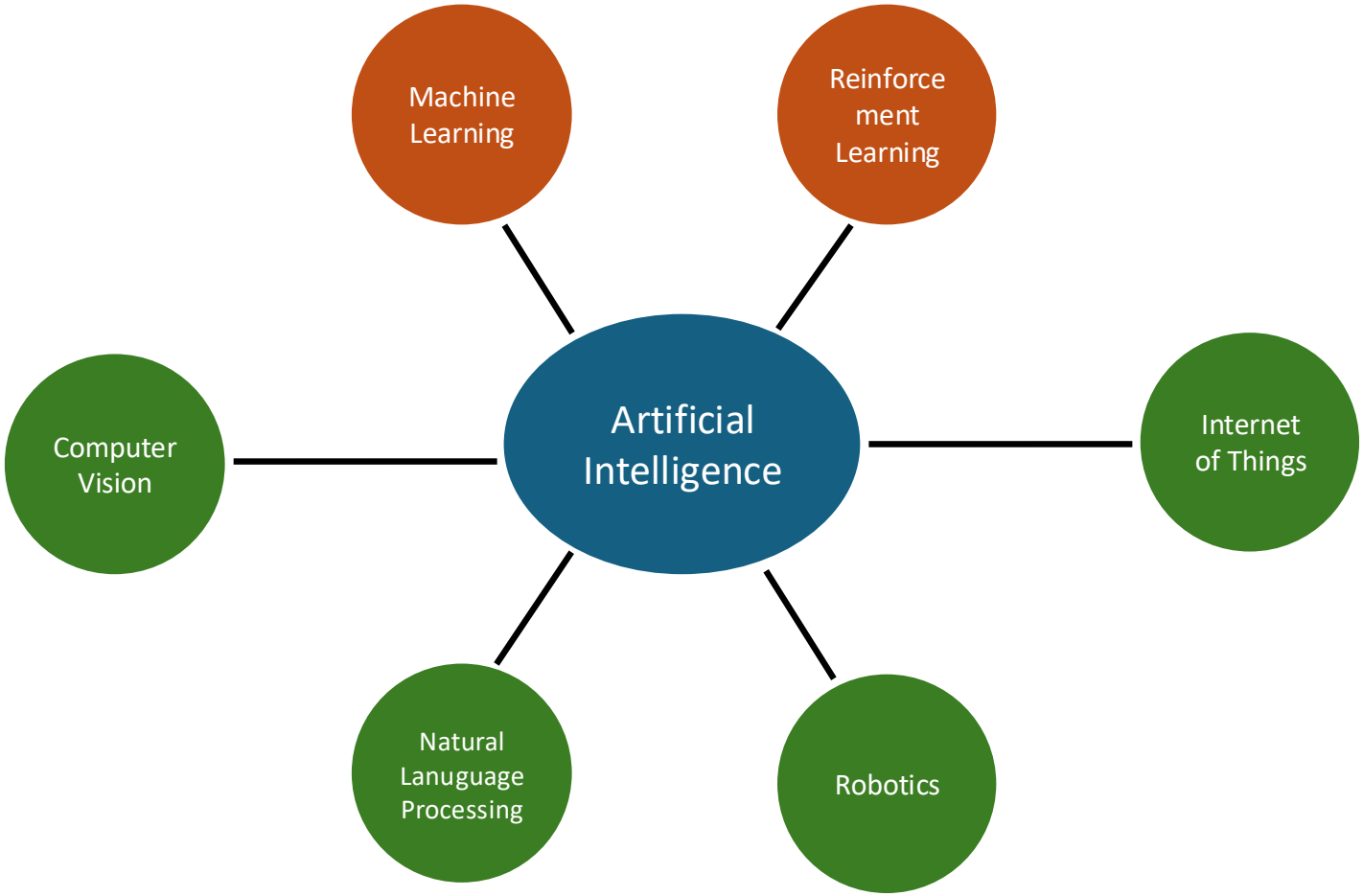


- Playing computer games
- Portfolio optimization
- Robotics
- Inventory management
- Query optimization
- Precision agriculture and farming
- LLMs

Areas of AI


Machine Learning
Reinforcement Learning
Computer Vision
Artificial Intelligence
Internet of Things
Natural Lanuguage Processing
Robotics

Areas of AI



- For Reinforcement Learning, at DIKU we offer OReL (B3) and SiRL (B4)
- A course dealing with basic theory and algorithms of sequential decision making problems tackled via online learning and reinforcement learning.

OReL 2025

- This will be the 5th round.
- 143 students — 47% increase relative to ‘OReL 2025’.

OReL 2025

- This will be the 5th round.
- 143 students — 47% increase relative to 'OReL 2025'.
- OReL is a restricted elective course in:
  - MSc Programme in Computer Science
  - MSc Programme in Statistics
  - MSc Programme in Mathematics-Economics

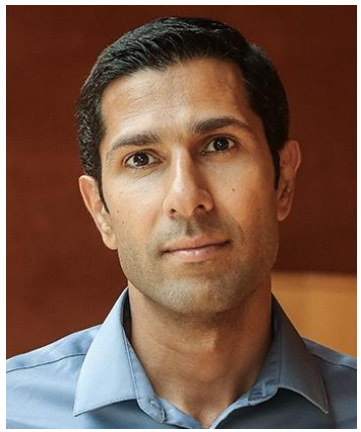
  In addition, we have students from other disciplines.

Course Team

Course Team: Instructors

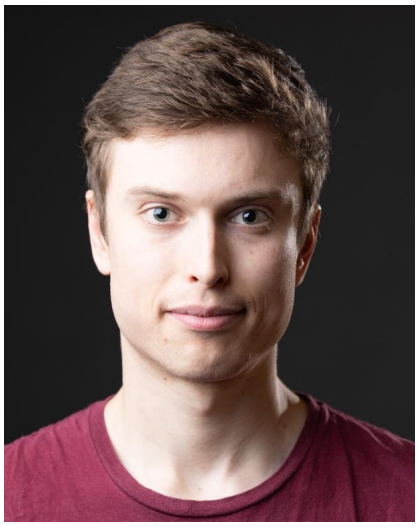
Christian Igel
Professor at ML Section, DIKU
Head of AI Center
igel@di.ku.dk
(4 lectures)

Yevgeny Seldin
Professor and head of
ML Section, DIKU
seldin@di.ku.dk
(5 lectures)



Sadegh Talebi
Assistant Professor at ML Section, DIKU
sadegh.talebi@di.ku.dk
(7 lectures & course management)

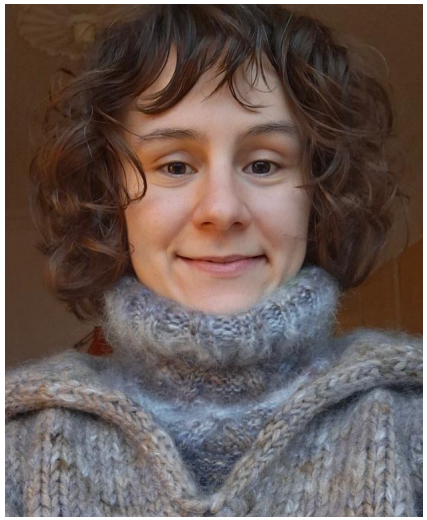
Course Team: Teaching Assistants



Alexander Meislich
MSc Student at DIKU

Hubert Drazkowski
PhD Candidate at DIKU

Laura Skovbæk
PhD Candidate at DIKU



Sarah Clusiau
Research Assistant at DIKU

Course Plan

Weekly Plan

	Mon	Tue	Wed	Thu	Fri
9:15 – 10		TA (x3)		Lecture 1	
10 – 11		TA (x3)		Lecture 1	
11 – 12		TA (x3) + OTA		Lecture 1 + Q&A	
12 – 13:15					
13:15 – 14				Lecture 2	TA (x1)
14 – 15				Lecture 2	TA (x1)
15 – 16				Lecture 2 + Q&A	TA (x1)
20:59 (sharp)			HA DEADLINE		

- 'TA' denotes *physical* exercise sessions.
- 'OTA' denotes the *online* TA session (held over Zoom).

Weekly Plan

	Mon	Tue	Wed	Thu	Fri
9:15 – 10		TA (x3)		Lecture 1	
10 – 11		TA (x3)		Lecture 1	
11 – 12		TA (x3) + OTA		Lecture 1 + Q&A	
12 – 13:15					
13:15 – 14				Lecture 2	TA (x1)
14 – 15				Lecture 2	TA (x1)
15 – 16				Lecture 2 + Q&A	TA (x1)
20:59 (sharp)			HA DEADLINE		

- 'TA' denotes *physical* exercise sessions.
- 'OTA' denotes the *online* TA session (held over Zoom).
- You can attend any and as many sessions you like.
- Note: A slightly different plan in Week 6.
- For locations and potential updates, check "Where & When" on Absalon.

Tentative Lecture Plan

Week 1	9:15-12:00			13:15-16:00
Tuesday		Course Introduction	Overview of Online and Reinforcement Learning	-
Thursday	Theory of MDPs (Part 1)			Stochastic Bandits (Part 1)

Week	Thursday, 9:15-12:00	Thursday, 13:15-16:00
2	Stochastid Bandits (Part 2) + Adversarial Full Information (Part 1)	Theory of MDPs (Part 2)
3	Off-policy Evaluation	Adversarial Full Information (Part 2) + Adversarial Bandits
4	Monte Carlo Methods and Direct Policy Search	Offline Evaluation of Bandits. Advanced Learning Settings (Contextual, Linear, ...)
5	Off-policy Optimization and Tabular Q-Learning	Policy Gradient Methods
6	Deep RL I	Sample Complexity under Simulators
7	Deep RL II	Regret Minimization in Episodic MDPs
8	Regret Minimization in Average-Reward MDPs	Course Summary

Yevgeny	Christian	Sadegh

Subject to minor changes; check Absalon for the latest info.

Tentative Lecture Plan

Week 1	9:15-12:00			13:15-16:00
Tuesday		Course Introduction	Overview of Online and Reinforcement Learning	-
Thursday	Theory of MDPs (Part 1) (RL)			Stochastic Bandits (Part 1) (Online Learning)

Week	Thursday, 9:15-12:00	Thursday, 13:15-16:00
2	Stochastid Bandits (Part 2) + Adversarial Full Information (Part 1) (Online Learning)	Theory of MDPs (Part 2) (RL)
3	Off-policy Evaluation (RL)	Adversarial Full Information (Part 2) + Adversarial Bandits (Online Learning)
4	Monte Carlo Methods and Direct Policy Search (RL)	Offline Evaluation of Bandits. Advanced Learning Settings (Contextual, Linear, ...) (Online Learning)
5	Off-policy Optimization and Tabular Q-Learning (RL)	Policy Gradient Methods (RL)
6	Deep RL I (RL)	Sample Complexity under Simulators (RL)
7	Deep RL II (RL)	Regret Minimization in Episodic MDPs (RL + Online Learning)
8	Regret Minimization in Average-Reward MDPs (RL + Online Learning)	Course Summary

Yevgeny	Christian	Sadegh

ubject to minor changes; check Absalon for the latest info.

Course Material

- **Main material:** Lecture notes, slides, some papers, . . . and blackboard

Lecture Notes:

- [YS] Machine Learning -- The Science of Selection under Uncertainty
- [CI] Concepts of Deep Reinforcement Learning: A Short Course
- [ST1] Notes on Theory of Discounted MDPs
- [ST2] Notes on Theory of Episodic MDPs
- [ST3] Notes on Theory of Average-Reward MDPs

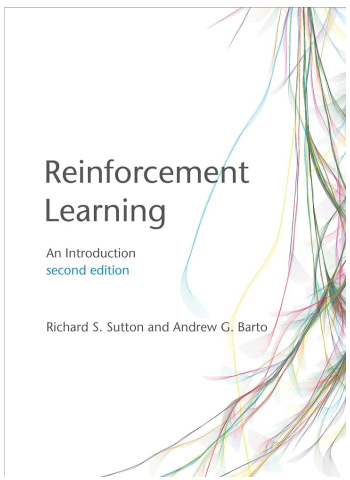
Course Material

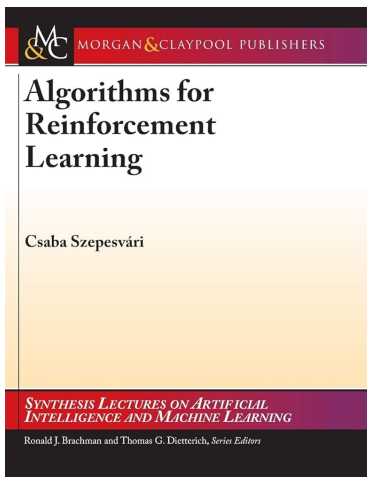
- **Main material:** Lecture notes, slides, some papers, . . . and blackboard

Lecture Notes:

- [YS] Machine Learning -- The Science of Selection under Uncertainty
- [CI] Concepts of Deep Reinforcement Learning: A Short Course
- [ST1] Notes on Theory of Discounted MDPs
- [ST2] Notes on Theory of Episodic MDPs
- [ST3] Notes on Theory of Average-Reward MDPs

- **Supplementary material (optional):**
  - R. Sutton & A. Barto. Reinforcement Learning: An Introduction.
  - Cs. Szepesvári. Algorithms for Reinforcement Learning.





Home Assignments

Home Assignments

Weekly home assignments

- Due dates: **Wednesdays at 20:59**.
- No resubmissions (except for re-exam, if necessary)
- Individual submission:
  - You are welcome to study together, but should write the solution **individually**.
  - Use of LLMs (e.g., ChatGPT, DeepSeek) is deemed cheating.
- Final grade $=$ the average grade of $n-1$ 'best' HAs
  - $n=8$ planned HAs
  - Submitting all HAs could help when your average is close to the boundaries.

Late Submissions

Late submissions will not be graded
... irrespective of the reason

Late submissions:

- Will not be graded ... irrespective of the reason
- But do submit, because we will consider (i) the number of submissions and (ii) their content when giving the final grade (i.e., it may help in borderline cases)

Sickness:

- Do NOT notify us about sickness/late submissions/etc.
- The final grade formula covers one potential emergency you may have during the course (e.g., sickness).
- If more than one emergency during the course, please, inform us at the end of the course and we will look into it.

Home Assignments: Feedback and Ref Solutions

- TAs provide feedback on your submissions:
  - *Short comments* regarding what was wrong when they take points.
  - Due to resource constraint, they are not expected to provide written feedback on how to fix your mistakes.
  - You can get further feedback or help from the TAs at a TA session (or a Q&A hour).

Home Assignments: Feedback and Ref Solutions

- TAs provide feedback on your submissions:
  - *Short comments* regarding what was wrong when they take points.
  - Due to resource constraint, they are not expected to provide written feedback on how to fix your mistakes.
  - You can get further feedback or help from the TAs at a TA session (or a Q&A hour).
- In case of complaints:
  - Any questions regarding the feedback? Ask the responsible TA first.
  - Generally, we will not consider complaints below 10 points per assignment. Contact the relevant TA first.

The Discussions Forum

In OReL, we use Discussions extensively.

Dos:

- Questions regarding course content
- Q/A regarding HAs
- Missing information
- Error reporting
- ...

The Discussions Forum

In OReL, we use Discussions extensively.

Dos:

- Questions regarding course content
- Q/A regarding HAs
- Missing information
- Error reporting
- ...

Donts:

- Individual matters (e.g., requests, point deductions),
- Complaints
- ...

For individual matters or complaints, contact the course coordinator (Sadegh).