

АННОТАЦИЯ рабочей программы дисциплины

Б1.В.08 «Методы обучения с подкреплением»

Объем трудоемкости: 2 з.е.

Цели дисциплины:

- познакомить студентов с основными разделами, концепциями и принципами обучения с подкреплением, включая ключевые термины, модели и алгоритмы.;
- обеспечить глубокое понимание теоретических основ, таких как марковские процессы принятия решений (MDP), функции ценности, политики и стратегии;
- научить студентов применять методы обучения с подкреплением на практике, включая разработку, реализацию и тестирование алгоритмов на реальных и синтетических задачах.

Задачи дисциплины: в соответствии с поставленной целью состоят в следующем:

- изучить основные задачи, решаемые при помощи методов обучения с подкреплением;
- развить навыки анализа и оценки эффективности различных алгоритмов обучения с подкреплением, а также умение выбирать подходящий метод в зависимости от задач;
- изучить библиотеки, необходимые при работе с методами обучения с подкреплением на Python (TF-Agents/ PyTorch RL/ OpenAi Gym)

Место дисциплины в структуре ООП ВО

Дисциплина «Методы обучения с подкреплением» относится к части, формируемой участниками образовательных отношений Блока 1 «Дисциплины (модули)» учебного плана. Данная дисциплина в качестве пререквизитов требует освоения следующих дисциплин: «Математические модели нейронных сетей» (основные алгоритмы обучения, функции потерь, оптимизаторы, начальные знания об интерпретируемости моделей, работа в jupyter notebook и google colab), «Глубокое обучение» (глубокие архитектуры, основы фреймворка pytorch), «Математический анализ» (основы дифференциального исчисления, интегральное исчисление и ряды, теория функций нескольких переменных), «Курс теории вероятностей» (матожидание, дисперсия, случайные величины и распределения, условная вероятность и Байесовский вывод), «Алгебра и геометрия» (операции с векторами и матрицами, собственные значения и собственные векторы), «Фундаментальные дискретные модели» (теория графов и теория игр), «Алгоритмы и структуры данных» (умение писать эффективный код). Постреквизитов в рамках учебного плана данная дисциплина не имеет. Смежными дисциплинами, изучаемыми параллельно или логически близко, являются дисциплины по выбору «Мультиагентные системы» и «Инженерия интеллектуальных агентов» (внедрение некоторых методов обучения с подкреплением на практике, в частности в робототехнике).

Требования к уровню освоения дисциплины

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих компетенций:

Код, уровень и формулировка компетенции	Индикаторы	Уровни освоения индикаторов компетенции
ПК-2 Способен участвовать в исследовании новых математических моделей в прикладных областях	ПК-2.1 Умеет анализировать и адаптировать существующие математические модели для решения прикладных задач в конкретной предметной области	Выполняет анализ и адаптацию существующих математических моделей в задачах обучения с подкреплением
	ПК-2.2 Способен предлагать и обосновывать новые математические подходы для моделирования процессов в прикладных исследованиях	Предлагает новые математические подходы для моделирования процессов в обучении с подкреплением
ML-6	ML-6.1 Обосновывает способы и варианты применения алгоритмов	(П) Применяет алгоритмы обучения на основе временных разностей и Q-обучения;

Способен применять алгоритмы обучения с подкреплением	обучения с подкреплением в задачах ИИ, включая их преобразование и адаптацию к специфике задачи	разрабатывает процедуру планирования поведения агента с помощью поиска по дереву Монте-Карло. Адаптирует под задачу архитектуры DQN, DDPG, A3C, TRPO, PPO и SAC. Понимает принципы архитектур AlphaGo, AlphaZero и MuZero
О-2 Способен применять и (или) разрабатывать мультиагентные алгоритмы	О-2.4 Оценивает результативность применения мультиагентных алгоритмов в задачах ИИ на основе сопоставления с аналогами	(П) Создает метрики качества решения задач ИИ, в которых учитывается эффект самоорганизации агентов
LLM-4 Проектирует, разрабатывает и интегрирует интеллектуальных агентов на базе генеративных моделей	LLM-4.1 Умеет применять и разрабатывать интеллектуальных агентов	(П) Проектирует архитектуру мультиагентных систем с адаптацией и памятью

Содержание дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.

№	Наименование разделов (тем)	Всего	Количество часов			
			Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	СРС
1	2	3	4	5	6	7
1.	Введение в обучение с подкреплением (RL)	11	4		2	5
2.	Табличные методы обучения с подкреплением	8,8	2		2	4,8
3.	Приближённые методы RL	12	4		2	6
4.	Политико-ориентированное обучение	10	2		2	6
5.	Усовершенствованные методы обучения с подкреплением	10	2		2	6
6.	Применения глубокого RL	14	2		6	6
7.	Тренды в RL	6	2			4
ИТОГО по разделам дисциплины		69,8	18		16	35.8
Контроль самостоятельной работы (КСР)		2				
Промежуточная аттестация (ИКР)		0.2				
Подготовка к текущему контролю						
Общая трудоемкость по дисциплине		72				

Курсовые работы: не предусмотрены

Форма проведения аттестации по дисциплине: зачет.

Авторы: Казаковцева Е.В. – к. ф.-м. н., доцент КАДИИ, Кадури А.А. – доцент КАДИИ.