

Аннотация дисциплины Б.1.2.5 Дисциплина. Математические основы обучения с подкреплением

Дисциплина "Математические основы обучения с подкреплением" изучается обучающимися по основной профессиональной образовательной программе "Интеллектуальные информационные системы и технологии" направления подготовки "09.03.02 Информационные системы и технологии".

Дисциплина изучается в 5 семестре. Общая трудоемкость дисциплины составляет 144/4 часов/з.ед. Самостоятельная работа заключается в выполнении работ, указанных в разделе 4.

В ходе изучения дисциплины осуществляется текущий контроль в форме технологии рейтингового контроля в соответствии с технологической карты дисциплины, размещенной на электронном курсе, а также промежуточный контроль в форме балльно-рейтинговый контроль.

Целью изучения дисциплины является формирование следующих компетенций:

1. MF-1 Способен применять современную теоретическую математику для разработки новых алгоритмов и формулирования перспективных задач искусственного интеллекта
2. MF-4 Способен применять статистические методы для анализа данных, валидации моделей машинного обучения и проведения экспериментов в области искусственного интеллекта

В ходе изучения дисциплины последовательно рассматриваются темы:

1. Лекция 1 (4ч) Введение в обучение с подкреплением. Постановка задачи и связь с другими подходами ML.
2. Лекция 2 (2ч) Марковские процессы принятия решений (MDP). Состояния, действия, вознаграждения, переходы.
3. Лекция 3 (3ч) Свойства марковских процессов, вероятностные переходы, стохастические политики.
4. Лекция 4 (3ч) Динамическое программирование: итерации по ценности и политике.
5. Лекция 5 (3ч) Методы Монте-Карло. Оценка функции ценности и оптимизация политик.
6. Лекция 6 (3ч) TD (Temporal Difference Learning) - обучение (TD(0), TD(?)). Сходимость, статистические свойства TD-ошибки.
7. Лекция 7 (2ч) Q-learning и SARSA. Анализ стохастических градиентов и статистических оценок.
8. Лекция 8 (4ч) Исследование–эксплуатация: ϵ -жадные, Softmax и байесовские стратегии.
9. Лекция 9 (3ч) Аппроксимация функции ценности: линейные и нейросетевые подходы.
10. Лекция 10 (3ч) Actor-Critic и стохастические градиентные методы.
11. Лекция 11 (3ч) Современные методы RL: DQN, Double DQN, PPO, DDPG, A3C
12. Лекция 12 (3ч) Валидация RL-агентов: A/B тестирование, доверительные интервалы, байесовские тесты.

Основными стратегическими образовательными технологиями являются: лекционные занятия, практические занятия.

В рамках указанных технологий применяются тактические образовательные технологии: задания, игровое проектирование, информационные, мини-проекты.