

**Федеральное государственное образовательное бюджетное
учреждение высшего образования
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ
ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)**

**Краснодарский филиал Финуниверситета
Кафедра «Математика и информатика»**

СОГЛАСОВАНО

ООО «Портал-Юг»
Генеральный директор



Е.В. Мостовой

«20» февраля 2024 г.

УТВЕРЖДАЮ

Краснодарский филиал
Финуниверситета
Директор



Э.В.Соболев

«20» февраля 2024 г.

Кирий В.А.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕМ**

для студентов, обучающихся по направлению подготовки
01.04.02 Прикладная математика и информатика,

Направленность программы магистратуры:
Анализ больших данных и машинное обучение
в экономике и финансах

*Рекомендовано Ученым советом Краснодарского филиала Финуниверситета
(протокол № 12 от 20.02.2024)*

*Одобрено кафедрой «Математика и информатика»
(протокол № 13 от 13.02.2024)*

Краснодар 2024

УДК 004.622

ББК 32.972

К30

Рецензент: Калайдин Е.Н., профессор кафедры «Математика и информатика», д-р физ.-мат. наук, Коренева О.В., доцент кафедры «Математика и информатика», канд. техн. наук

Кирий В.А. «Обучение с подкреплением». Рабочая программа дисциплины для студентов, обучающихся по направлению подготовки 01.04.02 Прикладная математика и информатика, направленность программы магистратуры: Анализ больших данных и машинное обучение в экономике и финансах. - Краснодар: Краснодарский филиал Финансового университета, Кафедра математики и информатики 2024.

Дисциплина **«Обучение с подкреплением»** относится к Модулю общепрофессиональных дисциплин направления подготовки 01.04.02 - Прикладная математика и информатика, направленности программы магистратуры: «Анализ больших данных и машинное обучение в экономике и финансах».

Рабочая программа дисциплины содержит цели и задачи дисциплины, требования к результатам освоения дисциплины, содержание дисциплины, тематику практических занятий и технологии их проведения, формы самостоятельной работы, систему оценивания, учебно-методическое и информационное обеспечение дисциплины.

Обучение с подкреплением

Рабочая программа дисциплины

*Формат 60*90/16. Гарнитура Times New Roman*

Усл. п.л. 2,0. Изд. № _от.

Тираж 100 экз.

Заказ № _____.

Отпечатано в Краснодарском филиале Финуниверситета

© Кирий В.А., 2024

© Краснодарский филиал Финансового университета, 2024

СОДЕРЖАНИЕ

1. Наименование дисциплины⁴
2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине.⁴
3. Место дисциплины в структуре образовательных программ⁵
4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся⁵
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий⁶
6. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по данной дисциплине⁹
7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины⁹
8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины¹³
9. Методические указания для обучающихся по освоению дисциплины¹³
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем¹⁴
11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине¹⁴

1. Наименование дисциплины

Дисциплина «Обучение с подкреплением»

2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине.

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
ПКН - 4	Способность разрабатывать информационные системы и алгоритмы на основе математических методов и моделей, в том числе из области искусственного интеллекта, в решении профессиональных задач	1. Демонстрирует знания в области оригинальных алгоритмов и программной реализации	Знать: основные инструменты в области оригинальных алгоритмов и программной реализации. Уметь: демонстрировать знания в области оригинальных алгоритмов и программной реализации.
		2. Создает оригинальные алгоритмические и программные средства в решении профессиональных задач	Знать: теоретические основы создания оригинальных алгоритмических и программных средств в решении профессиональных задач. Уметь: создавать оригинальные алгоритмические и программные средства в решении профессиональных задач.
		3. Демонстрирует знания в области современных интеллектуальных технологий.	Знать: теоретические основы в области современных интеллектуальных технологий. Уметь: демонстрировать знания в области современных интеллектуальных технологий.

		4. Выбирает интеллектуальные технологии и решает профессиональные задачи с их использованием.	Знать: интеллектуальные технологии и теорию решения профессиональных задач с их использованием. Уметь: выбирать интеллектуальные технологии и решать профессиональные задачи с их использованием.
--	--	---	--

3. Место дисциплины в структуре образовательных программ

Дисциплина «Обучение с подкреплением» относится к Модулю общепрофессиональных дисциплин направления подготовки 01.04.02 - Прикладная математика и информатика, направленности программы магистратуры: «Анализ больших данных и машинное обучение в экономике и финансах».

Дисциплина «Обучение с подкреплением» базируется на знаниях, полученных при изучении дисциплин «Проектирование архитектуры информационных систем» и «Построение и оценка моделей машинного обучения».

4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Модуль 4 (в часах)
Общая трудоемкость дисциплины	4 з/е, 144 ч.	144
<i>Контактная работа - Аудиторные занятия</i>	40	40
<i>Лекции</i>	8	8
<i>Семинары, практические занятия</i>	32	32
Самостоятельная работа	104	104
В семестре	68	68
Контроль (подготовка к зачету/экзамену)	36	36
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	Экзамен	Экзамен

5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

5.1. Содержание дисциплины

1. Задачи обучения с подкреплением. Оценочная обратная связь

Элементы обучения с подкреплением. Жадные и эpsilon-жадные стратегии. Метод UCB (upper confidence bound). Стратегия Softmax.

Взаимосвязь агент - окружающая среда. Цели и вознаграждения. Выгода. Марковские процессы принятия решений. Оптимальные функции ценности.

2. Методы Монте-Карло. Обучение по временной разнице

Прямое обучение от опыта агента. Обучение по полным эпизодам. Обучение по средней награде. Управление с разделяющей оценкой ценности стратегии.

Метод временных разностей TD. Метод Q-обучения. Градиентная оптимизация стратегии (policy gradient). Связь с максимизацией log-правдоподобия. Игры, после состояния и другие особые случаи.

3. Планирование и обучение. Оценка функции

Обучение на основе моделей. Понятие модели в обучении с подкреплением. Использование данных о модели в обучении с подкреплением. Эвристический поиск.

Задачи обучения с подкреплением, в которых требуется оценка функции. Характеристики состояния системы. Алгоритмы машинного обучения для оценки функции цены.

4. Важнейшие аспекты обучения с подкреплением. Конкретные примеры

Единый подход. Иерархическое обучение с подкреплением. Определение субкомпонент. Алгоритмы MAXQ и ALISP. Мульти-агентное обучение с подкреплением. Методы полной координации. Методы непрямо́й координации.

5.2. Учебно-тематический план

№ п/п	Наименование тем (разделов) дисциплины	Трудоёмкость в часах					Формы текущего контроля успеваемости
		Всего	Контактная работа - Аудиторная работа			Самостоятельная работа	
			Общая в т.ч.:	Лекции	Семинары, практические занятия		
1.	Задачи обучения с подкреплением. Оценочная обратная связь	36	10	2	8	26	Самостоятельные работы. Участие в решении задач на практических занятиях. Собеседования по домашним заданиям.
2.	Методы Монте- Карло. Обучение по временной разнице	36	10	2	8	26	
3.	Планирование и обучение. Оценка функции	36	10	2	8	26	
4.	Важнейшие аспекты обучения с подкреплением. Конкретные примеры	36	10	2	8	26	
	В целом по дисциплине	144	40	8	32	104	Согласно учебному плану: контрольная работа

5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8,9 (указывается раздел и порядковый номер источника)	Формы проведения занятий
Задачи обучения с подкреплением. Оценочная обратная связь	Введение в обучение с подкреплением: от многорукого бандита до полноценного RL агента. Задача выбора распознающего детектора. Алгоритм жадного выбора детектора <i>Рекомендуемые источники: п.8, [1]-[3]; п.9, [11]</i>	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Методы Монте-Карло. Обучение по временной разнице	Обучающие методы Монте-Карло. Прямое обучение от опыта агента. Обучение по полным эпизодам. Обучение по средней награде. Обучение по временной разнице. Бутстрэп по оценке функции ценности. TD-Lambda алгоритм Рекомендуемые источники: п. 8, [1]-[3].	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Планирование и обучение. Оценка функции	Обучение на основе моделей. Понятие модели в обучении с подкреплением. Использование данных о модели в обучении с подкреплением. Оценка функции. Задачи обучения с подкреплением, в которых требуется оценка функции. Характеристики состояния системы. Алгоритмы машинного обучения для оценки функции цены. Рекомендуемые источники: п.8, [1]-[3]; п.9, [11]	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Важнейшие аспекты обучения с подкреплением. Конкретные примеры	Иерархическое обучение с подкреплением. Назначение иерархических методов обучения с подкреплением. Определение субкомпонент. Алгоритмы MAXQ и ALISP. Мультиагентное обучение с подкреплением. Постановка задачи мультиагентного обучения с подкреплением. Методы полной координации. Методы непрямо́й координации. Рекомендуемые источники: п.8,[1]-[3]; п.9, [8]-[11]	Интерактивная форма, Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений

6. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по данной дисциплине

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Обучение с подкреплением».

Фонд оценочных средств включает контрольные материалы для проведения текущего контроля и промежуточной аттестации.

Вопросы для оценки знаний и умений, характеризующих формирование компетенций

Шифр компетенции	Вопросы	Правильный ответ
ПКН-4	1. Что такое функция ценности в обучении с подкреплением?	Оценка долгосрочной выгоды
	2. Что такое политика в контексте обучения с подкреплением?	Стратегия выбора действий агентом
	3. Как называется процесс принятия решений в обучении с подкреплением?	Процесс управления
	4. Что такое Q-функция в обучении с подкреплением?	Оценка долгосрочной выгоды для конкретного действия
	5. Что такое "исследование и эксплуатация" в обучении с подкреплением?	Баланс между исследованием новых действий и использованием известных
	6. Оценку чего выполняет функция ценности действия в обучении с подкреплением?	Выгоды
	7. Какие типы обратной связи используются в обучении с подкреплением?	Положительная и отрицательная
	8. Что такое "временная разница" (Temporal Difference) в обучении с подкреплением?	Метод оценки ценности состояний и действий
	9. Какие компоненты включает в себя модель среды в обучении с подкреплением?	Состояния, действия, вознаграждения
	10. Какие компоненты включает в себя модель обучения с подкреплением?	Агент, среда, награда, стратегия
	11. Какой метод используется для управления исследованием и использованием в обучении с подкреплением?	Epsilon-greedy
	12. Какой метод используется для оценки функции ценности в обучении с подкреплением?	Метод временных различий
	13. Какие виды обучения с подкреплением существуют на основе модели среды?	Модельное обучение с подкреплением
	14. Что такое "система наград" в контексте обучения с подкреплением?	Поощрение или наказание
	15. Какие методы используются для	Динамическое

обучения с подкреплением с моделью среды?	программирование, метод Монте-Карло
16. Какие типы стратегий можно использовать в обучении с подкреплением?	Детерминированные и стохастические стратегии
17. Что такое "задача оптимальности" в обучении с подкреплением?	Поиск стратегии, максимизирующей общую награду в среде
18. Что такое обучение с подкреплением?	Отношение агента к окружающей среде
19. Какие основные компоненты обучения с подкреплением?	Агент, среда, награда, стратегия
20. Какие типы обратной связи используются в обучении с подкреплением?	Положительная и отрицательная
21. Что такое функция ценности в обучении с подкреплением?	Оценка долгосрочной выгоды
22. Какие методы обучения с подкреплением вы знаете?	Q-обучение, SARSA, глубокое обучение
23. Что такое Q-обучение?	Метод оценки ценности действий
24. Какие проблемы могут возникнуть при обучении с подкреплением?	Проблема исследования и эксплуатации
25. Что такое функция ценности состояния-действия?	Оценка ожидаемого вознаграждения для конкретного действия в конкретном состоянии
26. Какие алгоритмы используются для управления обучением с подкреплением?	Алгоритмы SARSA и Q-обучения
27. Что такое функция вознаграждения?	Оценка полезности состояния или действия
28. Какие типы обучения с подкреплением можно выделить?	Обучение со случайными действиями и обучение с эксплорацией
29. Что такое стратегия в контексте обучения с подкреплением?	Правила выбора действий агентом
30. Какие методы используются для балансировки исследования и эксплуатации?	Epsilon-жадная стратегия и softmax-стратегия
31. Что такое модель в контексте обучения с подкреплением?	Представление окружающей среды агентом
32. Какие типы задач можно решать с помощью обучения с подкреплением?	Управление и принятие решений
33. Что такое функция ценности состояния?	Оценка ожидаемого вознаграждения для каждого состояния
34. Какие методы используются для обучения с подкреплением на основе нейронных сетей?	Глубокое Q-обучение и глубокое SARSA

Тесты

Шифр компетенции	Тестовые задания	Правильный ответ
ПКН-4	1. Обучение с подкреплением отличается от классического машинного обучения: а) наличием явной целевой функции б) отсутствием явной целевой функции с) использованием подкрепления для обучения д) использованием меток для обучения	b
	2. Алгоритм Q-learning оценивает качество каждого возможного действия с помощью: а) функции ценности б) функции вероятности с) функции ошибки д) функции подкрепления	a
	3. Марковский процесс принятия решений предполагает, что: а) будущее состояние системы зависит от всех предыдущих состояний б) будущее состояние системы не зависит от текущих состояний с) будущее состояние системы зависит только от предыдущего состояния д) будущее состояние системы определяется случайно	c
	4. Метод Policy Gradients: а) использует функции вероятности для выбора действий б) оптимизирует функцию политики с) использует градиентный спуск для оптимизации д) использует стохастический градиентный спуск	b
	5. Обучение с подкреплением — это ... а) Метод обучения без учителя б) Метод обучения с учителем с) Метод обучения на основе отклика окружающей среды д) Метод обучения на основе рандомизации	c
	6. Выберите основную цель алгоритмов обучения с подкреплением. а) Максимизация функции потерь б) Минимизация стоимости обучения с) Максимизация накопленной награды д) Минимизация времени обучения	c
	7. Определите название метода обучения с подкреплением, в котором агент обучается путем проб и ошибок, основанный на принципе проб и ошибок. а) Поиск с адаптивным шагом б) Метод осуществления с) Метод проб и ошибок (trial and error) д) Алгоритмическая демонстрация	c
	8. "Политика" в контексте обучения с подкреплением — это ...	b

	a) Задачи и цели, стоящие перед агентом b) Совокупность доступных действий и вероятностей их выбора c) Стратегия обучения агента d) Все варианты верны	
	9. Определите методы, которые могут использоваться для решения проблемы "значимость долгосрочной награды" в обучении с подкреплением. a) Дисконт-фактор b) Метод конечных разностей c) Метод кратчайших путей d) Все варианты верны	a
	10. Определите методы, которые могут использоваться для решения проблемы "исследование-использование" в обучении с подкреплением. a) Е-жадный метод b) Метод случайных действий c) Метод градиентного спуска d) Все варианты верны	d
	11. Определите роль, которая играет функция награды в обучении с подкреплением. a) Определяет степень уровня стресса агента b) Подталкивает к выбору наиболее выгодных действий c) Задает временной лимит для обучения d) Определяет вероятность ошибок при принятии решений	b
	12. Методология "обучение с подкреплением" в области создания информационных систем и алгоритмов — это ... a) Метод обучения на основе случайных исходов b) Метод обучения на основе правил вывода c) Метод обучения с использованием обучающих выборов d) Метод обучения на основе взаимодействия с окружающей средой	d

7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

7.1 Основная литература

1. Кондрашов, Ю. Н. Анализ данных и машинное обучение на платформе MS SQL Server: учебное пособие / Ю.Н. Кондрашов. — Москва: Русайнс, 2020. — 303 с. — Текст : непосредственный.— То же 2024.— ЭБС BOOK.ru. — URL: <https://book.ru/book/947076>. — Текст : электронный.

7.2. Дополнительная литература

1. Ростовцев, В. С. Искусственные нейронные сети : учебник / В. С. Ростовцев. — Санкт-Петербург : Лань, 2021. — 216 с. — ЭБС Лань. — URL: <https://e.lanbook.com/reader/book/160142> . — Текст: электронный.
2. Нестеров, С. А. Основы интеллектуального анализа данных. Лабораторный практикум : учебное пособие / С. А. Нестеров. — Санкт-Петербург : Лань, 2020. — 40 с. — ЭБС Лань. — URL: <https://e.lanbook.com/book/130181>. — Текст : электронный.
3. Михайлов, Г. А. Статистическое моделирование. Методы Монте-Карло : учебное пособие для вузов / Г. А. Михайлов, А. В. Войтишек. — Москва : Издательство Юрайт, 2024. — 323 с. — (Высшее образование). — ISBN 978-5-534-11518-5. — Текст : электронный // Образовательная платформа Юрайт [сайт]. URL: <https://urait.ru/bcode/516333>

8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Информационно-образовательный портал Финансового университета при Правительстве Российской Федерации <http://portal.ufrf.ru/>
2. Сайт департамента анализа данных, принятия решений и финансовых технологий.
3. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/> (<http://library.fa.ru/files/elibfa.pdf>)
4. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
5. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» <http://biblioclub.ru/>
6. Электронно-библиотечная система Znanium <http://www.znaniy.com>
7. «Деловая онлайн библиотека» издательства «Альпина Паблишер» <http://lib.alpinadigital.ru/en/library>
8. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com/>
9. Электронно-библиотечная система издательства «ЮРАЙТ» <https://www.biblio-online.ru/>
10. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>
11. Электронная библиотека (www.bibliotekar.ru).

9. Методические указания для обучающихся по освоению дисциплины

Самостоятельная работа студентов проходит аудиторно и внеаудиторно. Организации самостоятельной работы служит учебно-тематический план изучения дисциплины. В этом плане указана тематика лекций, практических занятий, вопросы и задания для самостоятельного изучения.

Контрольные работы проводятся в форме письменного экспресс-опроса, открытых и закрытых тестовых заданий, решения ситуационных и практических задач по темам учебной дисциплины

Домашние задания следует выполнять регулярно при подготовке к практическим занятиям. Контроль выполнения домашних заданий осуществляется в ходе практических занятий в процессе выборочного собеседования.

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем

10.1. Комплект лицензионного программного обеспечения:

- ОС Astra Linux,
- Libre Office,
- Антивирус Kaspersky.

10.2. Современные профессиональные базы данных и информационные справочные системы:

- Информационно-правовая система «Консультант Плюс»,
- Информационно-правовая система «Гарант»,
- Система комплексного раскрытия информации «СКРИН» - <http://www.skrin.ru/>

10.3. Сертифицированные программные и аппаратные средства защиты информации:

- не предусмотрены.

10.4. Свободный open source дистрибутив для языков программирования Python — ANACONDA

10.5. Эконометрический пакет R и интерфейс Rstudio.

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Материально-техническая база Краснодарского филиала Финансового университета соответствует действующим противопожарным правилам и нормам, обеспечивает проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Образовательный процесс обеспечивается специальными помещениями, которые представляют собой аудитории для проведения занятий лекционного типа, занятий семинарского типа, выполнения курсовых работ, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, помещения для самостоятельной работы студентов и помещения для хранения и профилактического обслуживания учебного оборудования. Специальные помещения укомплектованы специализированной мебелью и техническими средствами обучения, обеспечивающими представление учебной информации большой аудитории, демонстрационным оборудованием.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой, обеспечивающей доступ к сети Интернет и электронной информационно-образовательной среде Краснодарского филиала Финансового университета.