ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ

Сертификат: 025053D20009B2A5B043583E315FEF31AD
Владелец: Ивашкевич Евгений Васильевич
Действителен: с 14.10.2024 до 14.01.2026

УТВЕРЖДЕНА

Приказом Ректора АНО ВО
«Центральный университет»
Е.В. Ивашкевич
от «26» июня 2025 г. № 0626.32

Рабочая программа дисциплины (модуля) «AI Beyond Fit-Predict (Искусственный интеллект в действии)» дополнительной профессиональной программы – программы профессиональной переподготовки «Академия data science»

Траектория: Backend-разработка

Москва
2025

Электронный документ

Содержание

1. Краткая характеристика дисциплины (модуля)

Изучение дисциплины (модуля) «AI Beyond Fit-Predict (Искусственный интеллект в действии)» позволяет понять полный жизненный цикл AI-проектов, включая подготовку данных, разработку, развертывание и мониторинг моделей, что критично для успешной интеграции искусственного интеллекта в бизнес-процессы.

Цель изучения дисциплины (модуля): освоение методов и алгоритмов искусственного интеллекта для эффективного решения задач оптимизации, рекомендаций и обеспечения безопасности ML-моделей.

Задачи изучения дисциплины (модуля):

- изучить алгоритмы поиска и оптимизации для построения эффективных маршрутов и решений на графах;
- освоить методы принятия решений в условиях неопределенности на основе моделей многоруких бандитов;
- исследовать применение контекстуальных моделей для улучшения качества рекомендательных систем;
- ознакомиться с техниками выявления и противодействия враждебным атакам на модели машинного обучения;
- развить навыки планирования и реализации комплексных ML-проектов с учетом безопасности и устойчивости моделей.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- принципы работы и применение алгоритма поиска A* в оптимизации маршрутов и задачах на графах;
- модель многоруких бандитов;
- модель контекстуальных многоруких бандитов и их применение к рекомендательным системам;
- принципы и применение методов враждебного машинного обучения;
- принципы защиты нейросетей от враждебных атак.

уметь:

- применять алгоритм A* для разработки решений в задачах оптимизации и поиска путей;
- применять алгоритм UCB-1;
- находить атаки на незащищенные модели машинного обучения;
- защищать модель логистической регрессии от атак.

владеть:

- навыками планирования и проведения своей работы по разработке ML-решения.

2. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		Очная форма				
		Аудиторная работа		Контроль	Самостоятельная работа	
		Лекции	Семинары (практические занятия)			
1	Поиск и планирование в пространстве состояний	3	9		48	Подготовка к семинару, Домашние задания
2	Стохастическое принятие решений и обучение с подкреплением (Bandits)	3	10		50	Подготовка к семинару, Домашние задания, Контрольная работа
3	Безопасность и уязвимости алгоритмов машинного обучения	3	10		50	Подготовка к семинару, Домашние задания
	Зачет с оценкой			4		
	Итого:	***9***	***29***	***4***	***148***	
Объем дисциплины (модуля) (в ак. ч.)		***190***				

3. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Поиск и планирование в пространстве состояний	A* на графе. A* в непрерывном пространстве: приложение к видеоиграм.
2	Стохастическое принятие решений и обучение с подкреплением (Bandits)	Бандиты 1: постановка задачи; Бандиты 2: эксперты; Бандиты 3: контекстуальные бандиты.
3	Безопасность и уязвимости алгоритмов машинного обучения	Атаки на методы машинного обучения; Защита методов машинного обучения.

4. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый слушатель в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Слушателям обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Новиков, Ф. А. Символический искусственный интеллект: математические основы представления знаний : учебник для вузов / Ф. А. Новиков. — Москва : Издательство Юрайт, 2025. — 278 с. — (Высшее образование). — ISBN 978-5-534-00734-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/561410.

Дополнительная литература:

1. Russell, Stuart, Norvig, Peter. Artificial Intelligence: A Modern Approach. 3 : Prentice Hall, 2010.

5. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

— столами и стульями;

— компьютерной техникой;

— механическими калькуляторами;

— специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование"	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое

Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

6. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «AI Beyond Fit-Predict (Искусственный интеллект в действии)» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, практические занятия, домашние задания, контрольная работа, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удается, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Участие в семинаре (практическом занятии) – активная работа слушателя на семинаре, его ответы на вопросы преподавателя и участие в дискуссии.

Для успешного участия в семинаре слушателям рекомендуется заранее ознакомиться с темой обсуждения, прочитать необходимые материалы и подготовить вопросы. Важно активно слушать и вовлекаться в дискуссию, высказывая свои мнения и аргументируя их. При ответах на вопросы преподавателя стоит быть уверенным, четким и логичным, опираясь на изученный материал. Также полезно поддерживать диалог с однокурсниками, чтобы обогатить обсуждение и расширить свои знания.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал. Использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Контрольная работа – письменная работа с набором задач, которые нужно решить за ограниченное время.

Цель контрольной работы – получить специальные знания по одной или нескольким темам дисциплины и продемонстрировать навыки их практического применения.

Самостоятельная работа – работа слушателей, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы слушатели взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи

слушателя включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Оценивание уровня учебных достижений обучающихся по дисциплине (модулю) осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме ***зачета с оценкой***.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом:

Десятибалльная оценка	**Пятибалльная оценка**	**Оценка за зачет**	**Общая характеристика результата обучения по дисциплине (модулю)**
10	Отлично	Зачтено	Слушатель полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину (модуль). Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Слушатель хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
9	Отлично	Зачтено	
8	Отлично	Зачтено	
7	Хорошо	Зачтено	Слушатель обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи.
6	Хорошо	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Слушатель хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
5	Удовлетворительно	Зачтено	Слушатель обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Слушатель способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	Зачтено	
3	Не сдан	Не зачтено	Слушатель не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	Не зачтено	
1	Не сдан	Не зачтено	

Дисциплина (модуль) «AI Beyond Fit-Predict (Искусственный интеллект в действии)» оценивается следующим образом:

Активность	Вес	Описание
Аудиторная работа	10%	Активная работа слушателя на семинаре
Домашние задания	60%	Набор задач по темам недели
Контрольная работа	30%	Письменная работа с набором задач, которые нужно решить за ограниченное время

Формула расчёта итоговой оценки по дисциплине (модулю) «AI Beyond Fit-Predict (Искусственный интеллект в действии)»: «0,1 × аудиторная работа + 0,6 × среднее за домашние задания + 0,3 × контрольная работа».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные вопросы для подготовки к семинарам

Подготовка к семинару 1.

1. Что такое пространство состояний и как оно используется в задачах поиска и планирования?
2. Как работает алгоритм A* на графе: объясните роль эвристической функции и стоимости пути?
3. В чем разница между A* и другими алгоритмами поиска, такими как Dijkstra?
4. Как применяется A* в непрерывном пространстве, например, в видео-играх для планирования маршрутов персонажей?

5. Какие преимущества и ограничения имеет A* при решении задач оптимизации в динамических средах?

Подготовка к семинару 2.

1. Что такое постановка задачи многоруких бандитов и как она моделирует выбор действий в условиях неопределенности?
2. Как работает алгоритм UCB (Upper Confidence Bound) в контексте бандитов и почему он эффективен?
3. Что такое "эксперты" в моделях бандитов и как они используются для улучшения решений?
4. В чем суть контекстуальных бандитов и как они применяются в рекомендательных системах?
5. Как сравниваются стратегии в многоруких бандитах, такие как ε-greedy и UCB, по эффективности в долгосрочной перспективе?

Подготовка к семинару 3.

1. Какие типы атак на методы машинного обучения вы знаете и как они работают?
2. Что такое adversarial examples и как они могут обмануть модель, например, в классификации изображений?
3. Как можно обнаруживать атаки на незащищенные модели машинного обучения?
4. Какие методы защиты от враждебных атак существуют для нейросетей?
5. Как защитить модель логистической регрессии от adversarial inputs и какие инструменты для этого используются?

Примерные домашние задания

Домашнее задание 1.

1. Опишите основные шаги алгоритма A*. Как он отличается от других алгоритмов поиска, таких как Dijkstra?
2. Приведите пример графа и выполните поиск пути с использованием алгоритма A*. Укажите значения функции оценки (f(n)) для каждого узла.
3. Объясните, как алгоритм A* может быть адаптирован для работы в непрерывном пространстве. Какие методы используются для дискретизации пространства?
4. Исследуйте применение алгоритма A* в видеоиграх. Приведите примеры игр, где используется этот алгоритм, и объясните, как он улучшает игровую механику.
5. Разработайте собственный алгоритм A* для решения задачи поиска пути на двумерной сетке. Опишите его реализацию и протестируйте на простом примере.

Домашнее задание 2.

1. Определите, что такое задача многоруких бандитов и в чем ее основная сложность.
2. Объясните метод ε-жадного алгоритма. Как он помогает в решении задачи многоруких бандитов?
3. Рассмотрите ситуацию, когда у вас есть несколько экспертов, каждый из которых делает свои прогнозы. Как можно использовать модели многоруких бандитов для выбора лучшего эксперта?
4. Опишите, что такое контекстуальные бандиты. Как они отличаются от классических моделей многоруких бандитов?
5. Приведите пример реальной задачи, где можно применить модели многоруких бандитов, и опишите, как вы бы ее решили.

Домашнее задание 3.

1. Опишите основные типы атак на методы машинного обучения. Как они могут повлиять на производительность модели?
2. Разработайте пример атаки на модель машинного обучения, используя методы, описанные в литературе. Объясните, как эта атака может быть осуществлена.
3. Исследуйте методы защиты от атак на машинное обучение. Как можно улучшить устойчивость модели к атакам?
4. Обсудите важность этики в контексте машинного обучения и защиты данных. Как это может повлиять на разработку и развертывание моделей?
5. Проведите анализ уязвимостей в одной из популярных моделей машинного обучения. Какие меры можно предпринять для их устранения?

Примерные задания для контрольной работы

1. Объясните принцип работы алгоритма A* на графе. Приведите пример простого графа с 5 узлами и укажите, как A* выберет кратчайший путь от начального узла к целевому, используя эвристическую функцию (например, манхэттенское расстояние).
2. В чем разница между алгоритмами Dijkstra и A*? Приведите сценарий, где A* будет эффективнее Dijkstra, и объясните почему.
3. Опишите применение A* в непрерывном пространстве на примере планирования маршрута персонажа в видео-игре. Какие дополнительные вызовы возникают по сравнению с дискретным графом?
4. Предложите модификацию A* для обработки препятствий в непрерывном пространстве видео-игры. Объясните, как это повлияет на вычислительную сложность.
5. Что такое многорукие бандиты? Опишите классическую постановку задачи с примерами действий и наград.
6. Как работает алгоритм ε-greedy в контексте бандитов? Приведите пример с 3 руками и объясните, как он балансирует исследование и эксплуатацию.
7. Что такое модель бандитов с экспертами? Сравните ее с базовой моделью бандитов и приведите пример применения.
8. Опишите алгоритм Hedge для бандитов с экспертами. Как он учитывает веса экспертов при выборе действия?
9. В чем суть контекстуальных бандитов? Приведите пример их применения в рекомендательных системах.
10. Как контекстуальные бандиты улучшают базовую модель? Объясните на примере, где контекст влияет на выбор действия, и укажите потенциальные преимущества.