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»
Ивашкевич Е.В.
от «19» января 2024 г. № 0119.37

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

Траектория: Машинное обучение

Москва
2024

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 практических методов внедрения и масштабирования AI-решений для эффективного решения реальных задач.

Задачи изучения дисциплины (модуля):

— формирование знания о принципах работы и применении алгоритма поиска A* в оптимизации маршрутов и задачах на графах;

— формирование знания о модели многоруких бандитов;

— формирование знания о модели контекстуальных многоруких бандитов и их применении к рекомендательным системам;

— формирование знания о принципах и применении методов враждебного машинного обучения;

— формирование знания о принципах защиты нейросетей от враждебных атак;

— формирование умения применять алгоритм A* для разработки решений в задачах оптимизации и поиска путей;

— формирование умения применять алгоритм UCB-1;

— формирование умения находить атаки на незащищенные модели машинного обучения;

— формирование умения защищать модель логистической регрессии от атак;

— формирование навыка 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	Алгоритм A*	3	9		48	Домашние задания
2	Модели многоруких бандитов	3	10		50	Домашние задания Тест
3	Безопасность в машинном обучении	3	10		50	Домашние задания Кейс
	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	Алгоритм A*	A* на графе. A* в непрерывном пространстве: приложение к видеоиграм.
2	Модели многоруких бандитов	Бандиты 1: постановка задачи; Бандиты 2: эксперты; Бандиты 3: контекстуальные бандиты.
3	Безопасность в машинном обучении	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 к 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, тест, кейс задания, 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.

Решение кейса – практическая работа слушателей над реальными или смоделированными задачами, что позволяет слушателю применять теоретические знания на практике.

Слушатель самостоятельно разрабатывает стратегию решения поставленной задачи, что способствует развитию навыков критического мышления и самостоятельного принятия решений. Такой подход помогает подготовить будущих специалистов к реальным вызовам в их профессиональной деятельности.

Тест – особая форма проверки знаний. Проводится после освоения одной или

нескольких тем и свидетельствует о качестве понимания основных понятий изучаемого материала. Тестовые задания составлены к ключевым понятиям, основным разделам, важным терминологическим категориям изучаемой дисциплины (модуля).

Для подготовки к тесту необходимо знать терминологический аппарат дисциплины (модуля), понимать смысл научных категорий и уметь их использовать в профессиональной лексике. Владение понятийным аппаратом, включённым в тестовые задания, позволяет преподавателю быстро проверить уровень понимания слушателем важных методологических категорий.

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	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
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. 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.
6	Хорошо	Зачтено	
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
Домашние задания	50%	Набор задач по темам недели
Тест	20%	Набор заданий по теме на проверку знаний
Кейсы	30%	Практическая работа слушателей над реальными или смоделированными задачами

Формула расчёта итоговой оценки по дисциплине (модулю) «AI Beyond Fit-Predict (Искусственный интеллект в действии)»: «0,5 × среднее за домашние задания + 0,3 × среднее за кейсы + 0,2 × среднее за тесты».

Электронный документ

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание: Алгоритм A*

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: Модели многоруких бандитов

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: Применение и защита методов машинного обучения

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 кейсов

Кейс-задача: Атаки на методы машинного обучения

Ситуация:
Вы работаете в команде, которая разрабатывает модель машинного обучения для классификации изображений с целью автоматического распознавания лиц в системе безопасности. Ваша модель демонстрирует высокую точность на тестовых данных, но вы подозреваете, что она может быть уязвима к атакам.

Задание:

1. Определите, какие типы атак могут быть применены к вашей модели (например, атаки с использованием adversarial examples).

2. Проведите анализ возможных последствий успешной атаки на вашу модель. Как это может повлиять на безопасность системы?

3. Предложите методы тестирования вашей модели на устойчивость к атакам. Какие подходы вы бы использовали для обнаружения уязвимостей?

4. Разработайте стратегию для улучшения устойчивости модели к атакам. Какие изменения в архитектуре модели или в процессе обучения вы бы предложили?

5. Обсудите, как можно информировать пользователей о возможных рисках и ограничениях вашей модели.

Кейс-задача: Защита методов машинного обучения

Ситуация:

Ваша компания разрабатывает систему рекомендаций для онлайн-магазина. Модель машинного обучения анализирует поведение пользователей и предлагает товары на основе их предпочтений. Однако вы обеспокоены возможностью манипуляции моделью со стороны злоумышленников, которые могут пытаться исказить результаты рекомендаций.

Задание:

1. Определите основные уязвимости вашей системы рекомендаций и возможные способы атаки (например, манипуляция данными пользователей).

2. Исследуйте методы защиты, которые можно применить для повышения устойчивости вашей модели к таким атакам. Как можно использовать методы анонимизации или фильтрации данных?

3. Разработайте план по мониторингу и обнаружению аномалий в поведении пользователей, который поможет выявить потенциальные атаки на вашу модель.

4. Обсудите, как можно улучшить прозрачность и объяснимость вашей модели для пользователей. Почему это важно для повышения доверия к системе?

5. Подготовьте рекомендации для команды разработки о том, как интегрировать защитные меры в процесс разработки модели машинного обучения.

Примерные тестовые задания

Тест по теме "Бандиты"

1. **Что такое задача многоруких бандитов?**

   A) Задача, где необходимо выбрать один из нескольких вариантов действий для максимизации прибыли.
   B) Задача, где необходимо классифицировать объекты.
   C) Задача, где требуется предсказать будущее состояние системы.
   D) Задача, связанная с обработкой изображений.

2. **Какой из следующих методов является примером стратегии выбора действия в задаче многоруких бандитов?**

   A) Метод градиентного спуска
   B) ε-жадный алгоритм
   C) Метод опорных векторов
   D) Алгоритм K-средних

3. **Какой из следующих терминов описывает ситуацию, когда агент выбирает действие, основываясь на предыдущем опыте?**

A) Исследование
B) Эксплуатация
C) Обучение
D) Адаптация

4. **Что такое "эксперт" в контексте многоруких бандитов?**

   A) Модель, которая всегда дает идеальные прогнозы.
   B) Стратегия, которая выбирает действия на основе случайного выбора.
   C) Система, которая использует знания для оптимизации выбора действий.
   D) Способ оценки качества данных.

5. **Какой из следующих методов относится к контекстуальным бандитам?**

   A) Упрощенный ε-жадный алгоритм
   B) Алгоритм UCB (Upper Confidence Bound)
   C) Метод линейной регрессии
   D) Алгоритм Thompson Sampling

6. **Какова основная цель контекстуальных бандитов?**

   A) Минимизировать количество действий.
   B) Максимизировать ожидаемую награду с учетом контекста.
   C) Обучить модель без использования данных.
   D) Оптимизировать структуру данных.

7. **Что означает термин "долгосрочная награда" в контексте многоруких бандитов?**

   A) Награда, полученная за одно действие.
   B) Награда, полученная за последовательность действий.
   C) Награда, которая не зависит от выбора действий.
   D) Награда, полученная от случайного выбора.

8. **Какой из следующих подходов можно использовать для оценки эффективности стратегии в задаче многоруких бандитов?**

   A) Кросс-валидация
   B) A/B тестирование
   C) Метод главных компонент
   D) Линейная регрессия

9. **Какой из следующих факторов не влияет на выбор действия в контекстуальных бандитах?**

   A) Предыдущие действия
   B) Контекстные данные
   C) Случайные числа
   D) Награды за действия

10. **Что такое "exploration-exploitation trade-off"?**

    A) Баланс между обучением и тестированием модели.
    B) Баланс между исследованием новых действий и использованием известных.
    C) Баланс между количеством данных и качеством данных.
    D) Баланс между различными моделями машинного обучения.

11. **Какой из следующих алгоритмов является примером контекстуального бандита?**

    A) ε-жадный алгоритм

B) UCB1
C) LinUCB
D) Алгоритм KNN

12. **Что происходит, если агент всегда выбирает действие, которое уже принесло наибольшую награду?**

    A) Агент всегда будет успешным.
    B) Агент может пропустить более выгодные действия.
    C) Агент не сможет обучиться.
    D) Агент будет получать случайные награды.

13. **Какой из следующих методов используется для адаптации стратегии в контекстуальных бандитах?**

    A) Метод наименьших квадратов
    B) Линейная регрессия
    C) Метод максимизации правдоподобия
    D) Thompson Sampling

14. **Какой из следующих подходов может быть использован для уменьшения влияния "шумных" данных в задаче многоруких бандитов?**

    A) Увеличение размера выборки
    B) Использование фильтров для очистки данных
    C) Применение регуляризации
    D) Все вышеперечисленное

15. **В чем основное преимущество контекстуальных бандитов по сравнению с классическими многорукими бандитами?**

    A) Они требуют меньше данных.
    B) Они учитывают дополнительную информацию о контексте.
    C) Они проще в реализации.
    D) Они не требуют наград для обучения.