ЦЕНТРАЛЬНЫЙ УНИВЕРСИТЕТ

УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«24» июня 2025 г.
Протокол № 2

Рабочая программа дисциплины (модуля)
«AI Beyond Fit-Predict (Искусственный интеллект в действии)»

Направление подготовки: *02.04.01 Математика и компьютерные науки*

Направленность (профиль) подготовки: *Машинное обучение*

Квалификация (степень) выпускника: *магистр*

Форма обучения: *очная*

Срок освоения программы: *2 года*

Год набора: *2025*

ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ

Сертификат: 0215B0C000D1B3FEB44A37EE631E0C27D4
Владелец: Ивашкевич Евгений Васильевич
Действителен: с 13.01.2026 до 13.04.2027

Электронный документ

Москва
2025

Содержание

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «AI Beyond Fit-Predict (Искусственный интеллект в действии)» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – магистратура по специальности 02.04.01 Математика и компьютерные науки, профиль Машинное обучение, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 810 от 23.08.2017 года.

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.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки магистратуры по направлению 02.04.01 Математика и компьютерные науки, профиль Машинное обучение и входит в вариативную часть Блока 1, формируемую участниками образовательных отношений, как дисциплина по выбору.

Дисциплина (модуль) изучается на 2 курсе в 3 семестре, доступна для прохождения при условии успешного завершения дисциплины (модуля) «Machine Learning (Машинное обучение)».

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. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
УК-6.	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	УК-6.1.	Знает основные методы самооценки и анализа своей деятельности, а также принципы управления временем и целеполагания
		УК-6.2.	Умеет ставить реалистичные и достижимые цели, определять приоритеты в своей деятельности, а также разрабатывать и внедрять планы по совершенствованию своих навыков и компетенций на основе полученной самооценки
		УК-6.3.	Имеет практический опыт применения методов самооценки в своей профессиональной деятельности, включая участие в тренингах, семинарах и проектах, направленных на развитие личной эффективности и профессионального роста
ПК-1.	Способен определять общие формы и закономерности области машинного обучения	ПК-1.1.	Знает основные теоретические концепции и принципы, относящиеся к области машинного обучения, а также ключевые закономерности и модели, которые помогают в анализе и интерпретации данных
		ПК-1.2.	Умеет проводить систематический анализ области разработки, выявлять и формулировать общие закономерности и тенденции, а также применять методы исследования для получения новых знаний и понимания
		ПК-1.3.	Имеет практический опыт работы в области машинного обучения, включая участие в научных проектах, исследованиях или практических заданиях, где были выявлены и описаны общие формы и

			закономерности
ПК-4.	Способен публично представлять собственные и известные научные результаты	ПК-4.1.	Знает основные принципы эффективного публичного выступления, методы визуализации данных и основные требования к научным презентациям, включая структуру и содержание
		ПК-4.2.	Умеет четко и логично формулировать свои научные результаты, адаптируя их для различных аудиторий, а также использовать визуальные средства для улучшения восприятия информации
		ПК-4.3.	Имеет практический опыт участия в научных конференциях, семинарах или других мероприятиях, где успешно представлял свои и известные научные результаты, получая обратную связь и взаимодействуя с аудиторией
ПК-5.	Способен передавать результат решенных прикладных задач в виде конкретных рекомендаций, выраженных в терминах области машинного обучения	ПК-5.1.	Знает основные методы и подходы к формулированию рекомендаций на основе результатов решения прикладных задач, а также термины и концепции, специфичные для области машинного обучения
		ПК-5.2.	Умеет анализировать результаты решенных задач и формулировать четкие, конкретные рекомендации, адаптируя их к требованиям и ожиданиям целевой аудитории
		ПК-5.3.	Имеет практический опыт в разработке и представлении рекомендаций на основе анализа прикладных задач, включая участие в проектах, где результаты были успешно применены и оценены в контексте области машинного обучения

3. 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)	Консультации			
1	Поиск и планирование в пространстве состояний	3	2	2			Подготовка к семинару, Домашние задания
2	Стохастическое принятие решений и обучение с подкреплением (Bandits)	3	3	3			Подготовка к семинару, Домашние задания, Контрольная работа
3	Безопасность и уязвимости алгоритмов машинного обучения	2	3	3			Подготовка к семинару, Домашние задания
	Зачет с оценкой				4		
	Итого:	***8***	***8***	***8***	***4***	***86***	
Объем дисциплины (модуля) (в ак. ч.)		***114***					
Объем дисциплины (модуля) (в зач. ед.)		***3***					

4. 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 видео-играм
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

5. Учебно-методическое обеспечение

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

Каждый студент 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.

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

Основная литература:

1. Барретт, С. Ф. Arduino: искусственный интеллект и машинное обучение : практическое руководство / С. Ф. Барретт ; с англ. Ю. В. Ревича. – Москва : ДМК Пресс, 2024. - 244 с. – ISBN 978-5-93700-276-1. - Текст : электронный. - URL: https://znanium.ru/catalog/product/2205065.

2. Харбанс, Р. Грокаем алгоритмы искусственного интеллекта : практическое руководство / Р. Харбанс. - Санкт-Петербург : Питер, 2023. - 368 с. - (Серия «Библиотека программиста»). - ISBN 978-5-4461-2924-9. - Текст : электронный. - URL: https://znanium.com/catalog/product/2123366.

3. Мишра, П. Объяснимые модели искусственного интеллекта на Python. Модель искусственного интеллекта. Объяснения с использованием библиотек, расширений и фреймворков на основе языка Python : практическое руководство / П. Мишра ; пер. с англ. С. В. Минца. - Москва : ДМК Пресс, 2022. - 298 с. - ISBN 978-5-93700-124-5. - Текст : электронный. - URL: https://znanium.com/catalog/product/2109490.

Дополнительная литература:

1. Постолит, А. В. Основы искусственного интеллекта в примерах на Python : самоучитель / А. В. Постолит. - Санкт-Петербург : БХВ-Петербург, 2021. - 448 с. - (Самоучитель). - ISBN 978-5-9775-6765-7. - Текст : электронный. - URL: https://znanium.com/catalog/product/2123395.

2. Уорр, К. Надежность нейронных сетей: укрепляем устойчивость ИИ к обману : практическое руководство / К. Уорр. - Санкт-Петербург : Питер, 2021. - 272 с. - (Серия «Бестселлеры O'Reilly»). - ISBN 978-5-4461-1676-8. - Текст : электронный. - URL: https://znanium.com/catalog/product/1739681.

6. 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

7. 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, семинары, консультации, аудиторная работа, 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.

Семинар — это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Аудиторная работа – активная работа студента на семинаре, его ответы на вопросы преподавателя и участие в дискуссии.

Для успешного участия в семинаре студентам 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.

Консультации – структурированные встречи, на которых преподаватели предоставляют индивидуальную или групповую помощь в освоении учебного материала, обсуждении вопросов и решении проблем, возникающих в процессе обучения.

Консультации могут включать разъяснение сложных тем, подготовку к экзаменам и помощь в выполнении проектных работ, что способствует более глубокому пониманию предмета и улучшению академической успеваемости.

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 студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента 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)

Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «AI Beyond Fit-Predict (Искусственный интеллект в действии)»

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,*** при этом проводится оценка компетенций, сформированных по дисциплине.

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	Студент полностью владеет знаниями,

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
9	Отлично	Зачтено	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. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
8	Отлично	Зачтено	
7	Хорошо	Зачтено	Студент 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. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
6	Хорошо	Зачтено	
5	Удовлетворительно	Зачтено	Студент 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. Студент
4	Удовлетворительно	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
3	Не сдан	Не зачтено	Студент 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 студента 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.

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Укажите приоритет в самооценке навыков для изучения A* на графе.	Планирование / Оптимизация / Поиск	УК-6
2.	Укажите название алгоритма поиска пути, который использует эвристическую функцию для оптимизации в пространстве состояний.	A*	ПК-1
3.	Укажите рекомендацию по стратегии выбора действий в задаче бандитов для баланса исследования и эксплуатации.	Upper Confidence Bound / UCB	ПК-5
4.	Назовите способ совершенствования навыков в A* для видео-игр на основе самооценки.	Практика / Анализ / Эксперименты	УК-6
5.	Назовите способ демонстрации методов защиты ML.	Тесты / Модели / Рекомендации	ПК-4
6.	Укажите тип задачи обучения с подкреплением, где агент выбирает действия в условиях неопределенности без полной модели среды.	Multi-armed bandits / Бандиты	ПК-1
7.	Укажите ключевой приоритет в самооценке для решения задач бандитов.	Эксперименты / Моделирование / Оценка	УК-6
8.	Укажите рекомендацию по публичному представлению атак на ML.	Примеры / Доказательства / Защита	ПК-4

9.	Укажите общий тип уязвимостей алгоритмов машинного обучения, связанных с введением небольших возмущений в данные.	Adversarial attacks / Атаки с возмущениями	ПК-1
10.	Назовите метод реализации приоритетов в обучении с подкреплением на основе самооценки.	Итерации / Тестирование / Адаптация	УК-6
11.	Назовите метод представления контекстуальных бандитов.	Графики / Сценарии / Анализ	ПК-4
12.	Укажите основное отличие применения A* в непрерывном пространстве по сравнению с дискретным графом.	Непрерывность пространства / Непрерывное пространство	ПК-1
13.	Укажите рекомендацию по интеграции контекстуальных бандитов в системы рекомендаций для персонализации контента.	Использование контекста / Контекстуальный выбор	ПК-5
14.	Укажите приоритет в самооценке для контекстуальных бандитов.	Контекст / Стратегия / Эффективность	УК-6
15.	Укажите рекомендацию для публичного изложения экспертов в бандитах.	Структура / Визуалы / Дискуссия	ПК-4
16.	Укажите рекомендацию по применению A* в разработке видео-игр для оптимизации поведения персонажей.	Поиск оптимального пути / Оптимальный путь	ПК-5
17.	Укажите общий подход к решению задачи стохастического принятия решений с использованием бандитов.	Эксперимент и эксплуатация / Exploration and exploitation	ПК-1
18.	Назовите способ определения приоритетов в защите от атак на ML.	Анализ / Тестирование / Мониторинг	УК-6
19.	Назовите способ презентации научных результатов о бандитах.	Слайды / Модели / Примеры	ПК-4
20.	Укажите рекомендацию по защите моделей машинного обучения от adversarial атак.	Adversarial training / Противоречивый тренинг	ПК-5
21.	Укажите приоритет в самооценке для методов защиты ML-моделей.	Безопасность / Профилактика / Реагирование	УК-6
22.	Укажите рекомендацию по публичному представлению результатов A*-алгоритма.	Визуализация / Демонстрация / Объяснение	ПК-4
23.	Укажите рекомендацию по публичному представлению атак на методы машинного обучения.	Не раскрывать детали / Избегать раскрытия	ПК-5