

УТВЕРЖДЕНА

Приказом Ректора АНО ВО
«Центральный университет»
Ивашкевич Е.В.
от «19» января 2024 г. № 0119.37

**Рабочая программа дисциплины (модуля)
«ML System Design (Проектирование систем машинного обучения)»
дополнительной профессиональной программы – программы
профессиональной переподготовки «Академия data science»**

Траектория: Продуктовая аналитика

**Москва
2024**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Тематический план	4
3. Содержание дисциплины (модуля)	4
4. Учебно-методическое обеспечение	5
5. Материально-техническое обеспечение	5
6. Методические и оценочные материалы	7

1. Краткая характеристика дисциплины (модуля)

Изучение дисциплины (модуля) «ML System Design (Проектирование систем машинного обучения)» позволяет создавать интегрированные решения, обеспечивающие высокую производительность и устойчивость моделей в реальных условиях. Кроме того, знание проектирования систем машинного обучения способствует оптимизации ресурсов и ускоряет внедрение инноваций в различных сферах.

Цель изучения дисциплины (модуля): освоение принципов и методов проектирования эффективных, масштабируемых и надежных систем машинного обучения для решения практических задач.

Задачи изучения дисциплины (модуля):

- формирование знаний о задачах применения ML в компаниях и проведении предварительного анализа для оценки необходимости внедрения ML-решения;
- формирование знаний о ключевых компонентах архитектуры ML систем;
- формирование знаний об этапах жизненного цикла ML моделей;
- формирование знаний о способах оценки эффекта от ML решений;
- формирование знаний о принципах подготовки кода и моделей для продакшена;
- формирование знаний о методах мониторинга ML систем в продакшене;
- формирование знаний о принципах обеспечения масштабируемости и надежности ML систем;
- формирование умения формулировать техническую постановку ML задачи, выбирать метрики качества и бейзлайн;
- формирование умения разрабатывать концепцию ML системы с учетом бизнес-целей и ограничений;
- формирование умения разрабатывать дизайн ML-решения, включая функциональные требования и план реализации;
- формирование умения проектировать архитектуру ML системы, включая компоненты для обработки данных, обучения и инференса;
- формирование умения оценивать экономический эффект от внедрения ML-проекта;
- формирование умения формулировать требования к инфраструктуре с учетом нагрузки и масштабирования;
- формирование навыка проектирования ML систем, ориентированных на решение бизнес-задач;
- формирование навыка проверки жизнеспособности ML решения.

2. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		<i>Очная форма</i>				
		Аудиторная работа		Контроль	Самостоя тельная работа	
Лекции	Семинары (практичес кие занятия)					
1	Введение в ML System Design. Типы ML задач в бизнесе	1	4		24	Домашние задания
2	Оценка эффекта для бизнеса	1	5		24	Домашние задания
3	Дизайн, проработка и валидация ML решения	1	5		24	Домашние задания
4	Проектирование ML системы и разработка ML приложения	2	5		24	Домашние задания Кейсы
5	Обеспечение надежной работы ML приложений	2	5		26	Домашние задания Кейсы
6	Типичные ошибки в ML проектах. ML System Design интервью	2	5		26	Домашние задания Кейсы
	<i>Зачет с оценкой</i>			4		Проект
	Итого:	9	29	4	148	
	Объем дисциплины (модуля) (в ак. ч.)	148				

3. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Введение в ML System Design. Типы ML задач в бизнесе	Основные понятия машинного обучения. Классификация ML задач. Примеры применения ML в различных отраслях
2	Оценка эффекта для бизнеса	Метрики оценки эффективности ML решений. Анализ возврата инвестиций (ROI). Методы оценки экономического эффекта
3	Дизайн, проработка и валидация ML решения	Процесс разработки ML решения. Валидация моделей и тестирование. Принципы проектирования функциональных требований
4	Проектирование ML системы и разработка ML приложения	Архитектура ML системы. Интеграция с существующими системами. Разработка и развертывание ML приложений
5	Обеспечение надежной работы ML приложений	Методы мониторинга и поддержки ML систем. Обеспечение масштабируемости и производительности. Управление изменениями и обновлениями моделей
6	Типичные ошибки в ML проектах	Распространенные ошибки в проектировании ML систем. Подходы к предотвращению ошибок

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. Стружкин, Н. П. Базы данных: проектирование. Практикум : учебник для вузов / Н. П. Стружкин, В. В. Годин. — Москва : Издательство Юрайт, 2025. — 291 с. — (Высшее образование). — ISBN 978-5-534-00739-8. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/561215>.

Дополнительная литература:

1. Высоконагруженные приложения. Программирование, масштабирование, поддержка. — СПб.: Питер, 2018. — 640 с.: ил. — (Серия «Бестселлеры O'Reilly»). ISBN 978-5-4461-0512-0.

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) «ML System Design (Проектирование систем машинного обучения)» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, практические занятия, домашние задания, кейсы, проект, 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 удастся, 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.

Проект – исследовательская работа по дисциплине (модулю) и презентация

результатов.

Для успешной подготовки к проекту: четко определите цели и задачи проекта, распределите роли и обязанности между участниками, а также установите сроки выполнения каждой части работы. Регулярно проводите встречи для обсуждения прогресса и решения возникающих вопросов.

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
6	Хорошо	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			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.
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) «ML System Design (Проектирование систем машинного обучения)» оценивается следующим образом:

Активность	Вес	Описание
Домашние задания	50%	Набор задач по темам недели
Проекты	30%	Исследовательская работа по дисциплине (модулю) и презентация результатов
Кейсы	20%	Практическая работа слушателей над реальными или смоделированными задачами

Формула расчёта итоговой оценки по дисциплине (модулю) «ML System Design (Проектирование систем машинного обучения)»: « $0,5 \times$ среднее за домашние задания + $0,3 \times$ среднее за проекты + $0,2 \times$ среднее за кейсы».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание: Введение в ML System Design. Типы ML задач в бизнесе

1. Определите основные понятия машинного обучения и составьте краткий обзор (1-2 страницы).
2. Классифицируйте ML задачи по типам (например, классификация, регрессия, кластеризация) и приведите по 2 примера для каждой категории.
3. Исследуйте и опишите 3 примера применения ML в различных отраслях (например, здравоохранение, финансы, маркетинг).
4. Напишите краткий отчет о том, как ML может изменить бизнес-процессы в одной из выбранных вами отраслей.
5. Подготовьте презентацию (5-7 слайдов) о типах ML задач и их применении в бизнесе, включая примеры и визуализации.

Домашнее задание: Оценка эффекта для бизнеса

1. Изучите и опишите 3 основные метрики оценки эффективности ML решений (например, точность, полнота, F1-мера).
2. Рассчитайте ROI для гипотетического ML проекта, используя предоставленные данные (сделайте свои предположения о затратах и доходах).
3. Проведите анализ экономического эффекта внедрения ML в конкретной бизнес-области (выберите отрасль и обоснуйте выбор).
4. Подготовьте 1-2 страницы о том, как можно использовать метрики для оценки успеха ML проектов в реальном времени.
5. Создайте таблицу, в которой сопоставлены различные метрики и их применение в ML задачах, с примерами.

Домашнее задание: Дизайн, проработка и валидация ML решения

1. Опишите процесс разработки ML решения, включая все ключевые этапы (от сбора данных до развертывания модели).
2. Разработайте план валидации модели для конкретного ML проекта (определите методы валидации и тестирования).
3. Напишите о принципах проектирования функциональных требований для ML решения, включая примеры.
4. Создайте диаграмму, иллюстрирующую процесс разработки и валидации ML решения, включая ключевые этапы и взаимодействия.
5. Проведите анализ успешных и неудачных примеров валидации ML моделей, выделив ключевые уроки и рекомендации.

Примерные задания для кейсов

Кейс-задача: Проектирование ML системы и разработка ML приложения

Ситуация: Вы работаете в компании, которая хочет внедрить систему предсказания оттока клиентов для своего онлайн-сервиса. Ваша задача — спроектировать архитектуру ML системы.

Задания:

1. Опишите основные компоненты архитектуры ML системы, включая сбор данных, обработку, обучение модели и развертывание.
2. Предложите, как интегрировать новую ML систему с существующими системами компании (например, CRM, базы данных).
3. Опишите, какие технологии и инструменты вы бы использовали для разработки и развертывания ML приложения (например, языки программирования, библиотеки, облачные решения).
4. Подготовьте план по обеспечению безопасности данных в процессе разработки и развертывания системы.
5. Обоснуйте, как вы будете тестировать и валидировать систему перед ее запуском.

Кейс-задача: Обеспечение надежной работы ML приложений

Ситуация: Ваша компания развернула ML приложение для автоматизации процесса обработки заявок на кредит. Однако пользователи начали сообщать о сбоях и низкой производительности системы.

Задания:

1. Опишите методы мониторинга, которые вы бы использовали для отслеживания производительности ML приложения (например, метрики, логирование).
2. Разработайте стратегию поддержки ML системы, включая регулярные проверки и обновления.
3. Объясните, как вы будете обеспечивать масштабируемость приложения, чтобы оно могло обрабатывать увеличивающийся объем данных и запросов.
4. Опишите, как вы будете управлять изменениями и обновлениями модели, чтобы минимизировать риски и сбои.
5. Подготовьте отчет о том, какие шаги следует предпринять для устранения текущих проблем с производительностью и надежностью приложения.

Кейс-задача: Интеграция и развертывание ML решения

Ситуация: Ваша команда разработала модель для прогнозирования спроса на продукты в интернет-магазине. Теперь необходимо интегрировать эту модель в существующую систему управления запасами.

Задания:

1. Опишите шаги, необходимые для интеграции ML модели в систему управления запасами, включая взаимодействие с API и базами данных.
2. Разработайте план развертывания модели, включая выбор среды (локальная, облачная) и подходы к CI/CD (непрерывная интеграция и развертывание).
3. Объясните, как вы будете тестировать интеграцию модели с существующими системами на этапе развертывания.
4. Опишите, как вы будете обеспечивать поддержку и обновление модели после развертывания, включая стратегии мониторинга и обратной связи.
5. Подготовьте презентацию для заинтересованных сторон, в которой вы объясните, как внедрение ML модели повлияет на процессы управления запасами и повысит эффективность бизнеса.

Примерное описание к проекту

Задание для проекта: "Распространенные ошибки в проектировании ML систем и подходы к их предотвращению"

Описание проекта

Цель данного проекта — исследовать и проанализировать распространенные ошибки, возникающие при проектировании и разработке ML систем, а также предложить эффективные подходы и методы для их предотвращения. Участники проекта должны будут выбрать конкретные ошибки, проанализировать их причины и последствия, а также разработать рекомендации по предотвращению этих ошибок в будущем.

Этапы подготовки проекта

- 1. Выбор темы и формулирование проблемы**
 - Исследуйте различные аспекты проектирования ML систем.
 - Выберите 3-5 распространенных ошибок, которые вы хотите проанализировать.
 - Сформулируйте основные вопросы, на которые вы хотите ответить в проекте.
- 2. Сбор информации и анализ**
 - Изучите литературу и ресурсы по выбранным ошибкам (научные статьи, блоги, книги).
 - Проанализируйте примеры реальных проектов, где произошли ошибки, и их последствия.
 - Соберите информацию о подходах к предотвращению этих ошибок.
- 3. Разработка рекомендаций**
 - На основе собранной информации разработайте конкретные рекомендации по предотвращению выбранных ошибок.
 - Подготовьте презентацию, которая включает в себя основные выводы и рекомендации.
- 4. Подготовка к защите проекта**
 - Создайте слайды для презентации, включая визуализации и примеры.
 - Подготовьтесь к вопросам и обсуждениям, которые могут возникнуть во время защиты.
- 5. Защита проекта (1 день)**
 - Презентуйте свой проект перед аудиторией (можно в формате 10-15 минут).
 - Ответьте на вопросы и обсудите полученные результаты.

Критерии оценивания

- 1. Содержание проекта (40%)**
 - Глубина анализа выбранных ошибок.
 - Полнота и точность информации.
 - Качество предложенных рекомендаций.
- 2. Структура и оформление (20%)**
 - Логичность и последовательность изложения материала.
 - Качество слайдов и визуализаций (если используются).
 - Соответствие формальным требованиям (шрифт, размер, оформление).
- 3. Презентация (20%)**
 - Ясность и убедительность изложения.
 - Умение отвечать на вопросы и вести дискуссию.
 - Использование времени (соответствие отведенному времени).
- 4. Исследовательская работа (20%)**
 - Уровень самостоятельности и оригинальности в исследовании.
 - Умение использовать различные источники информации.
 - Критическое мышление и способность делать выводы.

Дополнительные рекомендации

- Работайте в команде, если это возможно, чтобы обмениваться идеями и опытом.
- Регулярно обсуждайте прогресс с преподавателем или наставником, чтобы получить обратную связь.
- Не забывайте о важности документирования вашего процесса работы и результатов для будущих проектов.