

УТВЕРЖДЕНА

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

**Рабочая программа дисциплины (модуля)
«Управление ML-продуктами»**

Направление подготовки: 02.04.01 Математика и компьютерные науки

Направленность (профиль) подготовки: Машинное обучение

Квалификация (степень) выпускника: магистр

Форма обучения: очная (с применением ДОТ)

Срок освоения программы: 2 года

Год набора: 2025

**Москва
2025**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения.....	5
3. Тематический план.....	7
4. Содержание дисциплины (модуля).....	7
5. Учебно-методическое обеспечение	9
6. Материально-техническое обеспечение	9
7. Методические и оценочные материалы	11

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Управление ML-продуктами» 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) «Управление ML-продуктами» позволяет развить навыки критического мышления и анализа данных, что является ключевым для принятия обоснованных решений в условиях неопределенности и быстроменяющихся технологий. Кроме того, знание принципов управления ML-продуктами помогает эффективно интегрировать машинное обучение в бизнес-процессы, обеспечивая создание ценности и конкурентных преимуществ.

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 Блок Факультативные дисциплины.

Дисциплина (модуль) изучается на 2 курсе в 3 семестре, доступна для прохождения при условии успешного завершения дисциплин (модулей) «Основы статистики», «SQL и базы данных», «Machine Learning (Машинное обучение)», «Продуктовая студия», «Python для анализа данных» и общеуниверситетского факультатива «Метрики бизнеса и продукта».

Цель изучения дисциплины (модуля): формирование навыков анализа данных и управления проектами в области машинного обучения для эффективного принятия решений и создания успешных ML-решений.

Задачи изучения дисциплины (модуля):

- изучить ключевые категории и направления AI/ML продуктов для понимания их применения в бизнес-средах;
- освоить принципы выбора и применения метрик для оценки эффективности ML-систем;
- разработать навыки создания стратегических планов развития ML-продуктов с учетом технических и бизнес-факторов;
- научиться анализировать данные и метрики для выявления соответствий между оффлайн- и онлайн-показателями;
- приобрести умения в координации процессов сбора и обработки данных для поддержки ML-исследований.

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- основные цели, типы и направления AI/ML продуктов;
- характеристику деятельности продакт-менеджера в AI/ML- продуктах;
- ключевые метрики качества для AI/ML продуктов: технические и продуктовые;
- принципы составления заданий для разметки данных, критерии оценки;
- методы сбора данных: логи, ручная разметка, полуавтоматическая разметка.

уметь:

- ориентироваться в целях, типах и направлениях AI/ML продуктов;
- подбирать метрики и делать разметку данных;
- разрабатывать ML roadmap для продуктов;
- на конкретных примерах анализировать сроки, качество данных, соответствие оффлайн и онлайн метрик.

владеть:

- навыками оценки качества AI/ML продуктов: подбирать метрики и делать разметку данных;
- навыком организации сбора данных и осуществления ML Research;
- навыком анализа применения ML-решений в продуктах
- навыком анализа проблем и решений в ML исследовании.

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
ОПК-2.	Способен создавать и исследовать новые математические модели в естественных науках, совершенствовать и разрабатывать концепции, теории и методы	ОПК-2.1.	Знает основные математические модели и методы, используемые в естественных науках, включая статистическое моделирование, дифференциальные уравнения и численные методы, а также современные подходы к исследованию и анализу данных
		ОПК-2.2.	Умеет разрабатывать и адаптировать математические модели для решения конкретных проблем в естественных науках, проводить их анализ и верификацию, а также интерпретировать полученные результаты в контексте научных исследований
		ОПК-2.3.	Имеет практический опыт создания и исследования математических моделей в

			рамках научных проектов или исследований, включая участие в публикациях, конференциях или коллаборациях, где были разработаны и апробированы новые концепции и методы
ПК-3.	Способен решать задачи профессиональной деятельности, формулировать результат, увидеть следствия полученного результата	ПК-3.1.	Знает основные принципы и методы решения задач профессиональной деятельности, а также способы формулирования и представления результатов, включая анализ последствий и их значимость в контексте проекта
		ПК-3.2.	Умеет применять математические и компьютерные методы для решения конкретных задач, формулировать четкие и обоснованные результаты, а также анализировать их последствия для дальнейших действий и решений
		ПК-3.3.	Имеет практический опыт в решении профессиональных задач, включая участие в проектах, где были получены результаты и проанализированы их следствия, что способствовало принятию обоснованных решений

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)					
1	Market & Strategy — Ideation	2				Домашние задания
2	ML Lifecycle & Stages	2				Домашние задания
3	Discovery: Проверка идеи и гипотез — Gate 1–2	2				Домашние задания
4	Discovery: Проверка концепции и быстрые прототипы — Gate 2–3	2				Домашние задания
5	Delivery: Подготовка данных и построение модели — Gate 4	2				Домашние задания
6	Delivery: Оценка качества и готовность к релизу — Gate 5	2				Домашние задания
7	Delivery: Масштабирование и оптимизация — Gate 5–6	2				Домашние задания
	<i>Зачет</i>			4		
	Итого:	14		4	96	
	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	Market & Strategy — Ideation	Цели и направления AI/ML продуктов: создание нового пользовательского опыта или оптимизация существующих процессов. Типы продуктов: рекомендации, чат-боты, ассистенты, системы предсказаний. Специфика работы продакт-менеджера с AI/ML-продуктами.
2	ML Lifecycle & Stages	Жизненный цикл ML-продукта: от исследования до масштабирования. Логика Stage-Gate: как принимаются решения от этапа к этапу. Подход PDCA в ML-разработке: как строится непрерывное улучшение. Роли в ML-команде: продакт, исследователь, инженер, аналитик, их зоны ответственности. Взаимосвязь UX-решений и ML-компонентов: где ML усиливает продукт, а где мешает.
3	Discovery: Проверка идеи и гипотез — Gate 1–2	Как отличить проблему в UX, данных или в алгоритме. Метрики результата в Discovery: что считать успехом идеи. Анализ данных как инструмент проверки гипотезы.

		<p>Формулирование ML-гипотезы: ценность, ограничения, критерии проверки.</p> <p>Определение базового уровня (baseline) и критериев Go/No-Go на Gate 1–2.</p>
4	Discovery: Проверка концепции и быстрые прототипы — Gate 2–3	<p>Проверка реализуемости (feasibility): качество данных, доступность сигналов, сложность решения.</p> <p>Экономика ML: стоимость вычислений, работы команды, риск ошибок.</p> <p>Как оценивать окупаемость (ROI) ML-инициатив.</p> <p>Быстрое прототипирование (“vibecoding”): как сделать проверку идеи без полной модели.</p> <p>Принятие решения: строить ML-модель или отказаться от идеи.</p>
5	Delivery: Подготовка данных и построение модели — Gate 4	<p>Способы получения данных: логирование, ручная разметка, полуавтоматические пайплайны.</p> <p>Разметка данных: как формировать задания, критерии качества, процессы проверки.</p> <p>Построение бейзлайна: минимально достаточная модель для оценки гипотезы.</p> <p>Генеративные модели и вопросы качества: алигнмент, контроль поведения модели.</p> <p>Связь качества данных и качества модели: где пределы и как их обнаружить.</p>
6	Delivery: Оценка качества и готовность к релизу — Gate 5	<p>Технические и продуктовые метрики качества ML-моделей.</p> <p>Соотношение оффлайн-метрик и реального онлайн-эффекта.</p> <p>Система контрольных метрик (guardrails): стабильность, надёжность, безопасность.</p> <p>Как определить готовность модели к продакшену.</p> <p>Наборы тестов качества: функциональные тесты, стресс-тесты, редкие и крайние случаи.</p>
7	Delivery: Масштабирование и оптимизация — Gate 5–6	<p>Дрейф моделей: типы дрейфа, причины, как обнаруживать и реагировать.</p> <p>Оптимизация ML-решений: производительность, стоимость, устойчивость.</p> <p>Мониторинг качества в реальном времени.</p> <p>Переобучение и обновление моделей: как строится итерационный цикл.</p> <p>Управление рисками и отклонениями поведения модели (alignment).</p>

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. Кацов, И. Машинное обучение для бизнеса и маркетинга : практическое руководство / И. Кацов. - Санкт-Петербург : Питер, 2019. - 512 с. - (Серия «IT для бизнеса»). - ISBN 978-5-4461-0926-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1783938>.

2. Мэрфи, К. П. Вероятностное машинное обучение. Дополнительные темы: предсказание, порождение, обнаружение, действие : практическое руководство / К. П. Мэрфи ; пер. с англ. А. А. Слинкина. – Москва : ДМК Пресс, 2024. - 768 с. – ISBN 978-5-93700-317-1. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2205085>.

3. Фомин, В. И. Информационный бизнес : учебник и практикум для вузов / В. И. Фомин. — 4-е изд., испр. и доп. — Москва : Издательство Юрайт, 2025. — 240 с. — (Высшее образование). — ISBN 978-5-534-14388-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/557608>.

Дополнительная литература:

1. Управление программными проектами : учебник для вузов / под редакцией Р. Ф. Маликова. — Москва : Издательство Юрайт, 2025. — 167 с. — (Высшее образование). — ISBN 978-5-534-14329-4. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/567576>.

2. Равичандиран, С. Глубокое обучение с подкреплением на Python. OpenAI Gym и TensorFlow для профи : практическое руководство / С. Равичандиран. - Санкт-Петербург : Питер, 2020. - 320 с. - (Серия «Библиотека программиста»). - ISBN 978-5-4461-1251-7. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1756109>.

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) «Управление ML-продуктами» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, домашние 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 удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

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.

Бонусные баллы — это оценки, которые студенты могут получить за выполнение дополнительных заданий.

Формат бонусных баллов позволяет студентам улучшить общую оценку по дисциплине (модулю) и стимулирует углубленное изучение материала.

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

Электронный документ

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) «Управление МЛ-продуктами»

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**, 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, 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.
9	Отлично	Зачтено	
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
6	Хорошо	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			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.
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 способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	Зачтено	
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) «Управление ML-продуктами» оценивается следующим образом:

Активность	Вес	Описание
Домашние задания	60%	Набор задач по темам недели
Зачет	40%	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

В рамках изучения дисциплины (модуля) возможно получение бонусных баллов.

Формула расчёта итоговой оценки по дисциплине (модулю) «Управление ML-продуктами»: $\langle 0,6 \times \text{среднее за домашние задания} + 0,4 \times \text{зачет} \rangle$.

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание

Кейс: AI-ассистент для студентов

Задача:

Разработайте концепцию AI-ассистента для студентов университета, который помогает:

- (а) быстро находить академическую информацию;

Электронный документ

- (б) объяснять сложные темы понятным языком;
- (в) решать учебные задачи (тесты, задачи, код).

Что нужно сделать:

1. Гипотезы и метрики (до 2х баллов)

Что нужно:

- **Продуктовая гипотеза (1 балл):**

- Опишите, какую *конкретную* пользовательскую или бизнес-метрику должен улучшить ассистент (например, "сокращение времени на поиск ответа", "рост удовлетворённости", "повышение вовлечённости в LMS")

- Продуктовые метрики

- **Техническая гипотеза (1 балл):**

- Почему выбранный подход (например, RAG, fine-tuning, rule-based) даст нужный эффект? Как он позволяет достигнуть целей продукта?

- Технические метрики

2. Концепция решения и риски (до 3х баллов)

Что нужно:

- **Архитектура решения (1 балл):**

- Опишите, из каких блоков будет состоять ассистент (miro, unidraw, etc)
- Какие данные используются?

- **Assumption Mapping (2 балла):**

- Сформулируйте ключевые предположения (по пользователям, данным, качеству ML).

- Распределите их по осям *важность / неопределённость* (или *impact / knownness*).

- Выделите 1–3 **критичных риска**, которые обязательно нужно проверять в PoC/MVP.

3. Proof of Concept (до 2 баллов)

Что нужно:

- **Экспериментальная проверка (1 балл):**

- Как протестировать критичные риски без полной реализации?

- **Метрики и критерии успеха (1 балл):**

- Какие численные или качественные показатели покажут, что PoC удался?
- Как вы будете их измерять?

4. MVP (до 3х баллов)

Что нужно:

- **Скоуп MVP (1 балл):**

- Какую минимальную функциональность можно запустить, чтобы ассистент начал приносить ценность?

- **Подготовка данных и контроль качества (1 балл):**

- Какие материалы использовать (методички, конспекты, видео)?
- Как обеспечить точность и полноту базы знаний?
- Кто будет разметчиком / валидатором?

- **Критерии успеха MVP (1 балл):**

- Сформулируйте, что будет считаться успешным MVP?
- Как это измерять?

Максимальный балл: **10**

Домашнее задание

Кейс "Данные и модели"

Контекст:

У вас есть машина времени и вы оказались в 2018 году. В «Ц-Банке» планируется запуск ML-модели для чат-бота поддержки. Чат-бот должен определять **интент** пользователя (цель и суть его сообщения), чтобы выдать точный ответ или передать

оператору. Готовых размеченных данных нет. Задача команды — **собрать и разметить датасет с нуля**, чтобы в дальнейшем обучить ML-модель.

Примеры интенгов:

- «Проверить баланс»,
- «Оформить карту»,
- «Пополнить счет»,
- «Не могу сделать перевод».

Вы работаете в команде, которая отвечает за сбор датасета.

Ваша цель — **спроектировать процесс разметки, контроль качества и способы улучшить результат**.

Как работать в команде:

Посмотрите на задание с разных сторон: не только как составитель задания на разметку, но и как разметчик (проверьте на ваших сокомандниках, насколько понятно задание, насколько удобно размечать данные).

Задания:

1. Выбор источников данных (1 балл)

Предложено не менее **трёх вариантов источников**, из которых можно взять примеры сообщений пользователей. Указаны **плюсы и минусы** каждого варианта и обосновано, какой из них лучше всего использовать на старте.

2. Инструкция и форма задания (2 балла)

Сформулирована понятная **инструкция для разметчиков**. Приложена **форма для задания**, в которой можно удобно размечать данные

3. Контрольные задания (2 балла)

Составлено **10 контрольных примеров** с правильной разметкой. К каждому примеру приложено обоснование, почему он может быть полезен как контрольный. Собрана таблица.

Пример	Почему важно его добавить как контрольный?
--------	--

4. Метрики оценки качества разметки (2 балла)

Сформулирован ответ на вопрос “Какие **метрики** нужно считать, чтобы понять, насколько качественно размечены данные?” Собрана таблица интерпретации каждой из метрик.

Метрика	Для чего на нее смотреть?
---------	---------------------------

5. Что делать при плохом качестве? (1 балл)

Предложен план действий на случай, если метрики качества данных будут ниже ожидаемых. Оформлена таблица.

Метрика	Порог	Что делать?
---------	-------	-------------

6. Как собрать данные без ручной разметки? (1 балл)

Предложен **альтернативный подход**, который позволит собрать датасет без ручной разметки.

7. Какие метрики важны при обучении модели? (1 балл)

Допустим, вы уже обучили первую версию модели. Сформулирован ответ на вопрос “Какие **метрики качества модели** вы будете отслеживать?” Указано:

- какие метрики нужно считать на валидации;
- какие — в проде;
- какие значения будут приемлемыми.

Общий балл: **10**

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Если Вы фиксируете личные цели развития как ML-продакт-менеджера, как одним словом называется такой ориентир?	цель	УК-6

2.	Если Вы регулярно пересматриваете продуктовую стратегию на основе данных, как одним словом называется такой подход?	итеративность	УК-6
3.	Если Вы оцениваете собственную эффективность в управлении ML-циклом, как одним словом называется такой анализ?	самооценка	УК-6
4.	Если Вы распределяете личное время между Discovery и Delivery этапами, как одним словом называется такое распределение?	планирование	УК-6
5.	Если Вы анализируете причины неудачи ML-гипотезы после Gate 2, как одним словом называется такой процесс?	рефлексия	УК-6
6.	Вы определяете приоритет ML-инициатив на основе ценности и ресурсов. Как одним словом называется такой процесс?	приоритизация	УК-6
7.	Если модель обучается на размеченных данных с известными ответами, как называется такой тип обучения одним словом?	supervised / обучение с учителем	ОПК-2
8.	Если изменение распределения входных данных приводит к ухудшению качества модели, как одним словом называется такое явление?	дрейф	ОПК-2
9.	Если модель оценивается без влияния на реальных пользователей, как одним словом называется такой тип оценки?	оффлайн / offline	ОПК-2
10.	Если проверяется реализуемость ML-идеи с точки зрения данных и сложности алгоритма, как одним словом называется такая проверка?	feasibility	ОПК-2
11.	Если модель используется как минимально достаточное решение для проверки гипотезы, как одним словом называется такой подход?	прототипирование	ОПК-2
12.	Если генеративная модель настраивается для соответствия заданным ограничениям поведения, как одним словом называется такой процесс?	alignment / алигнмент	ОПК-2
13.	Вы определяете, принесла ли ML-функция рост ключевой бизнес-метрики после релиза. Как одним словом называется такое изменение?	эффект	ПК-3
14.	Если решение о продолжении ML-инициативы принимается на основе критериев Go/No-Go, как одним словом называется такая точка принятия решения?	гейт / gate	ПК-3
15.	Если базовый уровень качества модели используется для сравнения с улучшенной версией, как называется такой уровень одним словом?	бейзлайн / baseline	ПК-3
16.	Если после запуска модели ключевая метрика ухудшилась, как одним словом называется такое явление?	деградация	ПК-3
17.	Если Вы оцениваете финансовую отдачу от ML-инициативы, как одним словом называется такой показатель?	ROI	ПК-3
18.	Если продуктовая метрика защищает систему от ухудшения пользовательского опыта, как одним словом называется такая метрика?	guardrail / гардрейл	ПК-3