



Федеральное государственное образовательное бюджетное учреждение высшего образования
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)

Кафедра искусственного интеллекта
Факультет информационных технологий и анализа больших данных

Документ подписан усиленной неквалифицированной электронной подписью.

Организация: «ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ
РОССИЙСКОЙ ФЕДЕРАЦИИ»

Сертификат: XV1yR5kZbNvM9E6V6xT1NmZJlOnKEoxE

Утверждено: Проректор по учебной и методической работе, Е.А. Каменева

Дата: 26.01.2026 г.

П.В. Никитин , Е.В. Романова , А.Н. Чупреева

Прикладные задачи машинного обучения

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки

01.03.02 - Прикладная математика и информатика,

Образовательная программа «Прикладное машинное обучение»

Профиль:

«Прикладное машинное обучение»

Рекомендовано

*Факультет информационных технологий и анализа больших данных
(протокол № 08 от 19.05.2026 г.)*

Одобрено

*Кафедра искусственного интеллекта
(протокол № 4 от 06.05.2026 г.)*



Содержание

1. Наименование дисциплины	3
2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине	3
3. Место дисциплины в структуре образовательной программы	
4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	7
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий	8
5.1. Содержание дисциплины	8
5.2. Учебно – тематический план	12
5.3. Содержание семинаров, практических занятий	13
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	15
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы	15
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю	20
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине	21
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	33
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	33
10. Методические указания для обучающихся по освоению дисциплины	39
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)	41
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	41



1. Наименование дисциплины

Прикладные задачи машинного обучения

2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
ПКП-1	Способность собирать наборы данных, в том числе больших данных, выполнять их подготовку для анализа в соответствии с решаемой прикладной задачей	Владеет навыками поиска внешних и внутренних источников данных для решения прикладной задачи	Знать: основные типы внешних (открытые государственные данные, API финансовых организаций, статистические сборники, данные ЦБ РФ) и внутренних (корпоративные базы данных, CRM-системы, журналы транзакций, результаты предыдущих аудиторских проверок) источников данных, а также методы оценки их пригодности для решения прикладной задачи Уметь: формулировать требования к необходимым данным исходя из бизнес-задачи, осуществлять поиск и доступ к идентифицированным источникам, оценивать полноту, достоверность и релевантность найденных данных, а также интегрировать данные из разнородных источников для формирования аналитической выборки
		Использует инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников	Знать: современные инструментальные средства для извлечения данных (библиотеки requests, BeautifulSoup, Scrapy для веб-скрапинга, SQLAlchemy и коннекторы к базам данных, API-клиенты), преобразования (pandas, dplyr, Spark SQL), хранения (реляционные и NoSQL базы



Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
			данных, объектное хранилище, Parquet, HDF5) и обработки (pandas, Dask, Apache Spark, modin) данных из разнородных источников Уметь: выбирать и применять соответствующие инструменты для каждого этапа работы с данными в зависимости от объема, структуры и скорости поступления данных, настраивать конвейеры ETL/ELT для интеграции разнородных источников и обеспечивать воспроизводимость процессов обработки
		Владеет современным инструментарием и практическими навыками визуализации данных	Знать: современные библиотеки и инструменты для визуализации данных (Matplotlib, Seaborn, Plotly, ggplot2, Tableau, Power BI), типы графиков и диаграмм, а также принципы выбора подходящего типа визуализации в зависимости от структуры данных и решаемой задачи (распределение, сравнение, связь, состав, географическое отображение) Уметь: создавать статические, интерактивные и анимированные визуализации для разведочного анализа данных, выявления закономерностей, обнаружения выбросов, проверки гипотез и презентации результатов, настраивая внешний вид графиков (цветовые схемы, легенды, аннотации) для улучшения восприятия
ПКП-6	Способность выполнять сборку модулей и компонент программной реализации	Демонстрирует знание архитектуры современных инструментальных средств реализации моделей ма-	Знать: архитектуру современных инструментальных средств реализации моделей машинного обучения (Scikit-learn, XGBoost/LightGBM, PyTorch, TensorFlow, Hugging



Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
	моделей машинного обучения и развертывания реализации моделей машинного обучения	шинного обучения и подходов к их интеграции	Face Transformers), включая их внутреннее устройство (вычисления с использованием векторизации, графы вычислений, автоматическое дифференцирование), а также основные подходы к их интеграции (конвейеризация с помощью Pipeline, обёртывание в REST API с использованием FastAPI/Flask, сериализация моделей (pickle, ONNX, TorchScript) и развёртывание в контейнерах Docker) Уметь: выбирать и комбинировать инструментальные средства для построения сквозного пайплайна машинного обучения (от предобработки данных до инференса), настраивать их совместную работу, а также интегрировать обученные модели в корпоративные информационные системы через API-интерфейсы или пакетную обработку
		Владеет практическим навыком интеграции компонент моделей машинного обучения	Знать: основные методы и подходы к интеграции компонентов моделей машинного обучения, включая сборку пайплайнов (FeatureUnion, ColumnTransformer), создание цепочек предобработки и обучения с использованием Pipeline, обёртывание моделей в сервисы с помощью FastAPI/Flask, а также способы интеграции нескольких моделей в единый ансамбль (стэкинг, блендинг, взвешенное голосование) Уметь: объединять отдельные компоненты (импутация пропусков, масштабирование, кодирование категориальных переменных, отбор признаков, обучение модели) в единый сквозной пайплайн, обеспечивая его корректную



Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
			работу на этапах обучения и применения, а также интегрировать несколько моделей в ансамбль для повышения точности прогнозов
		Владеет практическим навыком сборки и развертывания многокомпонентных моделей машинного обучения	Знать: современные инструменты и методы сборки многокомпонентных моделей машинного обучения (пайплайны в Scikit-learn, конвейеры данных в Apache Airflow/Prefect, системы управления версиями моделей DVC) и развертывания (контейнеризация с Docker, оркестрация с Kubernetes, асинхронные очереди задач Celery/RabbitMQ, сервера моделей TensorFlow Serving, BentoML, MLflow) Уметь: собирать многокомпонентные модели, включающие этапы предобработки данных, обучения и оценки, а также развертывать их в различных средах (локальной, на сервере, в облаке) с использованием контейнеризации и систем оркестрации для обеспечения масштабируемости и отказоустойчивости.



3. Место дисциплины в структуре образовательной программы

Дисциплина «Прикладные задачи машинного обучения» относится к «Профилю "Прикладное машинное обучение"».

4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

Очная форма обучения

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 7 (в часах)
Общая трудоемкость дисциплины	4/144	144
Контактная работа – Аудиторные занятия	50	50
<i>Лекции</i>	<i>16</i>	<i>16</i>
<i>Семинары, практические занятия</i>	<i>34</i>	<i>34</i>
<i>Самостоятельная работа</i>	<i>94</i>	<i>94</i>
<i>В том числе курсовая работа/курсовой проект</i>	<i>24</i>	<i>-</i>
Вид текущего контроля		
Вид промежуточной аттестации	Экзамен	Экзамен



5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

5.1. Содержание дисциплины

Тема 1. Методы оптимизации для больших данных и распределенных систем

Введение в распределенное обучение. Архитектуры распределенного обучения: Модели Master-Slave; Peer-to-Peer модели. Методы агрегации градиентов: синхронная vs асинхронная агрегация; преимущества и недостатки каждого подхода. Обработка больших данных: инструменты для работы с большими данными (Hadoop, Spark); влияние больших данных на выбор методов оптимизации. Обучение на графических процессорах (GPU): оптимизация алгоритмов для GPU; параллелизация вычислений. Методы распределенного SGD: алгоритмы для распределенного SGD; сравнение с традиционными методами. Масштабируемость методов оптимизации: проблемы масштабируемости и их решения; подходы к улучшению производительности на больших данных; применение распределенных методов в реальных задачах; Будущее распределенного обучения: новые тренды и направления исследований.

Тема 2. Методы распределенного и федеративного обучения

Введение в распределенное обучение: определение и основные концепции распределенного обучения; отличия между распределенным и централизованным обучением. Типы распределенного обучения: параллельное обучение: данные и вычисления; модельное параллельное обучение; данные параллельное обучение. Архитектуры распределенного обучения: Master-Slave архитектура; Peer-to-Peer архитектура; микросервисные архитектуры.

Федеративное обучение: определение и принципы федеративного обучения; особенности федеративного обучения по сравнению с традиционным обучением. Применения федеративного обучения: примеры использования в мобильных устройствах; применение в здравоохранении и финансовом секторе. Алгоритмы федеративного обучения: основные алгоритмы, такие как Federated Averaging (FedAvg) и его варианты; обсуждение улучшений и оптимизаций алгоритмов.

Безопасность и конфиденциальность: вопросы безопасности данных в распределенном и федеративном обучении; методы обеспечения конфиденциальности, такие как дифференциальная приватность.



Вызовы и ограничения: проблемы с коммуникацией и задержками; неоднородность данных и вычислительных мощностей; сложности с синхронизацией моделей.

Оценка производительности: метрики для оценки эффективности распределенного и федеративного обучения; сравнение с традиционными методами обучения.

Будущее распределенного и федеративного обучения: тенденции и перспективы развития технологий; возможные направления исследований и применения.

Тема 3. Большие языковые модели (LLM) и их архитектуры

Архитектуры LLM: обзор архитектур трансформеров; ключевые компоненты, такие как механизмы внимания и слои.

Обучение больших языковых моделей: методы обучения, включая обучение с учителем и без учителя; использование предобучения и дообучения.

Предобучение и дообучение: стратегии предобучения на больших объемах данных; адаптация моделей к специфическим задачам через дообучение. Тонкая настройка (fine-tuning): подходы к тонкой настройке моделей для конкретных приложений; влияние объема данных на эффективность тонкой настройки. Оценка производительности LLM: метрики для оценки качества генерации текста и понимания языка; сравнение с другими моделями.

Этика и ответственность в использовании LLM: проблемы предвзятости, конфиденциальности и безопасности; обсуждение последствий использования LLM в обществе.

Применения больших языковых моделей: примеры использования в чат-ботах, системах автоматического перевода и генерации контента.

Тенденции и будущее LLM: новые направления исследований и разработки; прогнозы по эволюции архитектур и методов обучения.

Интеграция LLM в бизнес-процессы: возможности применения в различных отраслях; влияние на автоматизацию и оптимизацию рабочих процессов.

Тема 4. Мульти-модальные модели и их интеграция

Введение в мульти-модальные модели: определение и основные характеристики мульти-модальных подходов; их роль в интеграции различных типов данных (текст, изображения, звук и т.д.).

Архитектуры мульти-модальных моделей: обзор архитектур, поддерживающих работу с несколькими модальностями; ключевые компоненты, такие как механизмы внимания и совместное представление данных.



Обучение мульти-модальных моделей: методы обучения, включая совместное обучение и предобучение на различных модальностях; использование больших объемов данных из разных источников.

Предобучение и дообучение мульти-модальных моделей: стратегии предобучения на многомодальных данных; адаптация моделей к специфическим задачам через дообучение на целевых модальностях.

Интеграция данных из разных модальностей: подходы к объединению информации из текста, изображений и звука; методы синхронизации и согласования данных.

Оценка производительности мульти-модальных моделей: метрики для оценки качества генерации и понимания данных из различных модальностей; сравнение с однородными моделями.

Применения мульти-модальных моделей: примеры использования в системах распознавания объектов, автоматическом переводе, генерации контента и других областях.

Этика и ответственность в использовании мульти-модальных моделей: проблемы предвзятости, конфиденциальности и безопасности; обсуждение последствий использования таких моделей в обществе.

Тенденции и будущее мульти-модальных моделей: новые направления исследований и разработки; прогнозы по эволюции архитектур и методов обучения.

Интеграция мульти-модальных моделей в бизнес-процессы: возможности применения в различных отраслях; влияние на автоматизацию, анализ данных и пользовательский опыт.

Тема 5. Современные исследования в машинном обучении

Обзор литературы по актуальным направлениям машинного обучения: анализ последних публикаций в ведущих научных журналах; обзор ключевых конференций и их материалов (NeurIPS, ICML, CVPR и др.); тенденции в исследованиях: от классических алгоритмов к глубокому обучению.

Предпринты и открытые исследования: изучение платформ, таких как arXiv, для поиска новейших исследований; обсуждение значимых предстоящих работ и их потенциального влияния на область; критический анализ методологий и результатов из предпринтов.

Практические примеры применения машинного обучения: кейсы успешного применения машинного обучения в различных отраслях (медицина, финансы, маркетинг, производство); разбор реальных проектов и стартапов, использующих машинное обучение для решения конкретных задач; влияние машинного обучения на бизнес-процессы и принятие решений.



Инновационные технологии и инструменты: обзор новых фреймворков и библиотек для разработки моделей машинного обучения; исследование новых подходов к интерпретации моделей; разработка и использование автоматизированных систем для машинного обучения (AutoML).

Этика и социальные аспекты применения машинного обучения: обсуждение проблем предвзятости в алгоритмах и их влияние на результаты; вопросы конфиденциальности данных и соблюдения этических норм; роль машинного обучения в решении социальных проблем.

Будущее машинного обучения: тренды и перспективы: прогнозы по развитию технологий и методов машинного обучения; обсуждение междисциплинарных исследований и их влияния на область; потенциал интеграции с другими технологиями (блокчейн, IoT, квантовые вычисления).



5.2. Учебно – тематический план

Очная форма обучения

п/ п	Наименование тем (разделов) дисциплины	Трудоемкость в часах				
		Всего	Контактная работа - Аудиторная работа			Самостоятельная работа
			Общая, в т.ч.:	Лекции	Семинары, практические занятия	
1	Методы оптимизации для больших данных и распределенных систем	30	10	4	6	20
2	Методы распределенного и федеративного обучения	44	8	2	6	36
3	Большие языковые модели (LLM) и их архитектуры	22	8	2	6	14
4	Мульти-модальные модели и их интеграция	30	18	6	12	12
5	Современные исследования в машинном обучении	18	6	2	4	12
В целом по дисциплине		144	50	16	34	94



5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарах, практических занятиях	Формы проведения занятий
Методы оптимизации для больших данных и распределенных систем	Методы оптимизации в машинном обучении, включая градиентный спуск и его стохастические варианты. Преимущества и недостатки различных подходов. Пример задания: оптимизация гиперпараметров моделей с использованием Bayesian Optimization и разработка адаптивных методов обучения для глубоких нейронных сетей.	Интерактивная форма, практикум в малых группах (2–4 студента), коллективное обсуждение решений. Семинары с ИТ-практиками, конкурсы, чемпионаты, мастер-классы, командные проектные задания.
Методы распределенного и федеративного обучения	Концепции распределенного и федеративного обучения, их преимущества и недостатки. Подходы к обеспечению безопасности и конфиденциальности данных. Пример задания: создание системы для обучения моделей на мобильных устройствах с использованием федеративного подхода и разработка распределенных систем для обработки больших объемов данных в облачных средах.	Интерактивная форма, практикум в малых группах (2–4 студента), коллективное обсуждение решений. Семинары с ИТ-практиками, конкурсы, чемпионаты, мастер-классы, командные проектные задания.
Большие языковые модели (LLM) и их архитектуры	Архитектуры больших языковых моделей, таких как Transformer, и их применение в обработке текста. Проблемы, связанные с обучением и использованием LLM, включая предвзятость и потребление ресурсов. Пример задания: разработка чат-ботов на основе LLM для поддержки клиентов и создание систем автоматического перевода текста с использованием больших языковых моделей.	Интерактивная форма, практикум в малых группах (2–4 студента), коллективное обсуждение решений. Семинары с ИТ-практиками, конкурсы, чемпионаты, мастер-классы, командные проектные задания.



Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарах, практических занятиях	Формы проведения занятий
Мульти-модальные модели и их интеграция	Интеграция различных типов данных, таких как текст, изображения и звук, для решения сложных задач. Вызовы, возникающие при обучении мульти-модальных моделей, а также способы оценки их производительности. Пример задания: разработка системы распознавания объектов на изображениях с текстовыми описаниями и мультимедийные поисковые движки, которые комбинируют текстовые и визуальные данные.	Интерактивная форма, практикум в малых группах (2–4 студента), коллективное обсуждение решений. Семинары с ИТ-практиками, конкурсы, чемпионаты, мастер-классы, командные проектные задания.
Современные исследования в машинном обучении	Текущие тренды в исследованиях машинного обучения, проблемы предвзятости и интерпретируемости моделей. Новые методы и подходы, предложенные в недавних публикациях. Пример задания: исследование по улучшению интерпретируемости глубоких нейронных сетей и разработка новых алгоритмов машинного обучения для применения в области экономики и финансов.	Интерактивная форма, практикум в малых группах (2–4 студента), коллективное обсуждение решений. Семинары с ИТ-практиками, конкурсы, чемпионаты, мастер-классы, командные проектные задания.



6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Методы оптимизации для больших данных и распределенных систем	1. В чём заключается компромисс между синхронной и асинхронной агрегацией градиентов при распределённом обучении? Какие факторы влияют на выбор подхода? 2. Как устроен метод эластичного усреднения (Elastic Averaging SGD) и за счёт чего он позволяет улучшить сходимость по сравнению с классическим распределённым SGD? 3. Какие проблемы масштабируемости возникают при обучении моделей машинного обучения на данных объёмом более 1 ТБ, и какие существуют подходы к их решению? 4. Как влияние архитектуры GPU-памяти (Shared Memory vs. Global Memory) на эффективность параллелизации вычислений градиента?	Работа с учебной литературой. Разбор вопросов по теме занятия.
Методы распределенного и федеративного обучения	1. В чём заключается различие между горизонтальным и вертикальным федеративным обучением? Приведите примеры задач, для которых подходит каждый из подходов. 2. Как работает алгоритм Federated Averaging (FedAvg) и почему усреднение весов модели на	Работа с учебной литературой. Разбор вопросов по теме занятия.



Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
	сервере эквивалентно усреднению градиентов при независимых идентично распределённых данных? 3. Какие дополнительные механизмы безопасности (помимо дифференциальной приватности) применяются в федеративном обучении для защиты от атак восстановления данных (например, атак по градиентам)? 4. Как проблема неоднородности данных (non-IID данные) влияет на сходимость федеративного обучения и какие существуют алгоритмы для её смягчения (например, FedProx, SCAFFOLD)?	
Большие языковые модели (LLM) и их архитектуры	1. В чём отличие механизмов Grouped-Query Attention (GQA) и Multi-Query Attention (MQA) от стандартного Multi-Head Attention, и как они влияют на скорость инференса LLM? 2. Почему в больших языковых моделях возникает проблема «галлюцинаций» (hallucinations), и какие подходы к её решению существуют (например, RLHF, chain-of-thought, retrieval-augmented generation)? 3. Как работает механизм KV-кэширования при авторегрессивной генерации текста в трансформерных декодерах, и почему он критически важен для производительности LLM?	Работа с учебной литературой. Разбор вопросов по теме занятия.



Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Мульти-модальные модели и их интеграция	1. Какие существуют стратегии объединения (fusion) информации из разных модальностей (раннее, позднее, гибридное), и каковы их преимущества и недостатки в задачах типа VQA (Visual Question Answering)? 2. Как работает механизм кросс-модального внимания в архитектурах типа Flamingo или LLaVA, и как он позволяет согласовывать текстовые и визуальные представления? 3. В чём заключается проблема выравнивания модальностей (modality alignment) при обучении мультимодальных моделей, и как она решается с использованием контрастивных методов (например, CLIP, ALIGN)? 4. Какие способы синхронизации потоков данных из разных модальностей (аудио, видео, текст) во времени существуют, и как обрабатываются ситуации асинхронного поступления данных в реальных системах?	Работа с учебной литературой. Разбор вопросов по теме занятия.
Современные исследования в машинном обучении	1. Каковы основные методологические критерии критической оценки качества статей на ведущих конференциях (NeurIPS, ICML, ICLR) при	Работа с учебной литературой. Разбор вопросов по теме занятия.



Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
	проведении систематического обзора литературы? 2. В чём отличие препечатных платформ (arXiv, bioRxiv) от рецензируемых журналов с точки зрения скорости распространения знаний и контроля качества, и как правильно интерпретировать результаты из препринтов? 3. Какие направления исследований в области машинного обучения сейчас являются наиболее активно развивающимися согласно анализу публикаций за последние 2–3 года (например, агентные системы, диффузионные модели, LLM-агенты)? 4. Как междисциплинарные исследования на стыке машинного обучения и квантовых вычислений могут повлиять на развитие оптимизационных методов в ближайшие 5–10 лет?	



6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю

Задание 1. Анализ и предобработка данных

Загрузить датасет с информацией о клиентах банка (возраст, доход, кредитная история, сумма кредита, цель кредита, просрочка). Выполнить первичный анализ: вывести статистику по числовым признакам, обработать пропуски (заполнить медианой для доходов и модой для цели кредита), закодировать категориальные переменные методом One-Hot Encoding и нормализовать числовые признаки с использованием StandardScaler.

Задание 2. Сравнение методов классификации

На датасете кредитного скоринга обучить три модели классификации: логистическую регрессию, дерево решений и случайный лес. Сравнить их качество по метрикам accuracy, precision, recall и ROC-AUC. Построить матрицы ошибок для каждой модели и сделать вывод о том, какая модель лучше подходит для задачи минимизации ложноположительных срабатываний.

Задание 3. Прогнозирование временного ряда

По данным о ежемесячных налоговых поступлениях за последние 5 лет построить прогноз на следующие 6 месяцев с использованием модели Prophet (Meta). Визуализировать исходный ряд, компоненты разложения (тренд, сезонность) и прогноз с доверительными интервалами. Оценить точность прогноза на отложенной выборке (например, последние 6 месяцев) с помощью метрик MAE и RMSE.

Задание 4. Построение рекомендательной системы

На датасете MovieLens (контекст финансовых продуктов можно заменить на банковские) реализовать рекомендательную систему на основе матричной факторизации (SVD). Разделить данные на обучающую и тестовую выборки, вычислить RMSE на тестовой выборке. Для заданного пользователя выдать топ-5 рекомендаций и объяснить, почему были предложены именно эти продукты.

Задание 5. Реализация пайплайна машинного обучения

Создать сквозной пайплайн с использованием Scikit-learn Pipeline, включающий этапы: импутация пропусков медианой, стандартизация числовых признаков, One-Hot кодирование категориальных признаков и обучение модели градиентного бустинга (LightGBM или XGBoost). Обучить пайплайн на данных о финансовых транзакциях для задачи обнаружения мошеннических операций и оценить качество с помощью Precision-Recall кривой.

Задание 6. Работа с API и интеграция модели

Разработать простое REST API приложение на FastAPI, которое загружает предобученную модель классификации (например, случайный лес для кредитного скоринга) и предоставляет два эндпоинта: /predict (одиночный прогноз) и



/batch_predict (загрузка CSV-файла с несколькими клиентами). Реализовать валидацию входных данных и возврат прогнозов в формате JSON. Продемонстрировать работоспособность с помощью Postman или curl.

Дополнительная информация

Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях Кафедры искусственного интеллекта Факультета информационных технологий и анализа больших данных.



7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

Перечень компетенций с указанием индикаторов их достижения в процессе освоения образовательной программы содержится в разделе 2. «Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине».

Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний

Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
ПКП-1. Способность собирать наборы данных, в том числе больших данных, выполнять их подготовку для анализа в соответствии с решаемой при-	Владеет навыками поиска внешних и внутренних источников данных для решения прикладной задачи	Знать: основные типы внешних (открытые государственные данные, API финансовых организаций, статистические сборники, данные ЦБ РФ) и внутренних (корпоративные базы данных, CRM-системы, журналы транзакций, результаты предыдущих	По задаче «Прогнозирование вероятности банкротства предприятия» составить спецификацию необходимых данных, определить не менее трёх внешних (например, данные ФНС, СПАРК, статистика ЦБ РФ) и двух внутренних (например, кредитная история, финансовые показатели компании за 3 года) источников, обосновать выбор каждого источника, сформулировать запросы на получение данных и предложить схему их объединения (вплоть до структуры итоговой таблицы признаков).



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
кладной задач		аудиторских проверок) источников данных, а также методы оценки их пригодности для решения прикладной задачи Уметь: формулировать требования к необходимым данным исходя из бизнес-задачи, осуществлять поиск и доступ к идентифицированным источникам, оценивать полноту, достоверность и релевантность найденных данных, а также интегрировать данные из разнородных источников для фор-	



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		мирования аналитической выборки	
	Использует инструментальные средства для извлечения, преобразования, хранения и обработки данных из разнородных источников	Знать: современные инструментальные средства для извлечения данных (библиотеки requests, BeautifulSoup, Scrapy для веб-скрапинга, SQLAlchemy и коннекторы к базам данных, API-клиенты), преобразования (pandas, dplyr, Spark SQL), хранения (реляционные и NoSQL базы данных, объектное хранилище, Parquet, HDF5) и обработки (pandas, Dask, Apache Spark,	Разработать и реализовать на Python конвейер с использованием библиотек pandas, SQLAlchemy и requests, который извлекает данные о курсах валют из открытого API ЦБ РФ в формате JSON, подключается к локальной базе данных SQLite для извлечения таблицы с ежедневными транзакциями клиентов, выполняет объединение, очистку и агрегацию данных, а затем сохраняет результат в формате Parquet для последующего анализа. Весь код должен быть задокументирован, а логи обработки должны позволять воспроизвести результат.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		modin) данных из разнородных источников Уметь: выбирать и применять соответствующие инструменты для каждого этапа работы с данными в зависимости от объёма, структуры и скорости поступления данных, настраивать конвейеры ETL/ELT для интеграции разнородных источников и обеспечивать воспроизводимость процессов обработки	
	Владеет современным инструмента-	Знать: современные библиотеки и инстру-	По датасету финансовых операций банка построить комплексный отчёт с визуализациями: гистограмму распределения сумм транзакций, ящик с усами для выявления вы-



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
	рием и практическими навыками визуализации данных	менты для визуализации данных (Matplotlib, Seaborn, Plotly, ggplot2, Tableau, Power BI), типы графиков и диаграмм, а также принципы выбора подходящего типа визуализации в зависимости от структуры данных и решаемой задачи (распределение, сравнение, связь, состав, географическое отображение) Уметь: создавать статические, интерактивные и анимированные визуализации для разведочного	бросов по каждому типу транзакции (с возможностью скрытия/показа категорий), интерактивный линейный график динамики среднего чека по месяцам с возможностью масштабирования, а также тепловую карту корреляции между числовыми признаками. Все графики должны быть подписаны, содержать легенду и сохранены в HTML-файл с возможностью самостоятельного изучения пользователем.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		анализа данных, выявления закономерностей, обнаружения выбросов, проверки гипотез и презентации результатов, настраивая внешний вид графиков (цветовые схемы, легенды, аннотации) для улучшения восприятия	
ПКП-6. Способность выполнять сборку модулей и компонент программной реализации моделей машинного обучения	Демонстрирует знание архитектуры современных инструментов реализации моделей машинного обучения и подходов к их интеграции	Знать: архитектуру современных инструментов реализации моделей машинного обучения (Scikit-learn, XGBoost/LightGBM, PyTorch, TensorFlow, Hugging Face Transformers), вклю-	Спроектировать архитектуру системы для кредитного скоринга, включающую этапы: предобработка данных (Pandas с обёрткой в Scikit-learn Pipeline), обучение модели (XGBoost с кросс-валидацией и подбором гиперпараметров через Optuna), сериализацию модели в формате ONNX, развёртывание модели в виде микросервиса на FastAPI с эндпоинтами /predict (для одного клиента) и /batch_predict (для загрузки CSV-файла). Описать все компоненты, их взаимодействие, форматы входных/выходных данных, а также предложить схему CI/CD для обновления модели в продакшене.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
и развертывания реализации моделей машинного обучения		чая их внутреннее устройство (вычисления с использованием векторизации, графы вычислений, автоматическое дифференцирование), а также основные подходы к их интеграции (конвейеризация с помощью Pipeline, обёртывание в REST API с использованием FastAPI/Flask, сериализация моделей (pickle, ONNX, TorchScript) и развёртывание в контейнерах Docker) Уметь: выбирать и комбинировать ин-	



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		струментальные средства для построения сквозного пайплайна машинного обучения (от предобработки данных до инференса), настраивать их совместную работу, а также интегрировать обученные модели в корпоративные информационные системы через API-интерфейсы или пакетную обработку	
	Владеет практическим навыком интеграции компонент моделей машинного обучения	Знать: основные методы и подходы к интеграции компонентов моделей машинного обучения, включая сборку пайплай-	На основе датасета кредитного скоринга построить сквозной пайплайн с использованием Scikit-learn Pipeline, включающий заполнение пропусков медианой, стандартизацию числовых признаков, one-hot кодирование категориальных признаков и обучение модели случайного леса. Затем построить ансамбль из трёх моделей (логистическая регрессия, случайный лес, градиентный бустинг) с использованием VotingClassifier, реализовать его обучение и предсказание, а также сравнить качество ансамбля с каждой



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		нов (FeatureUnion, ColumnTransformer), создание цепочек предобработки и обучения с использованием Pipeline, обёртывание моделей в сервисы с помощью FastAPI/Flask, а также способы интеграции нескольких моделей в единый ансамбль (стэкинг, блендинг, взвешенное голосование) Уметь: объединять отдельные компоненты (импутация пропусков, масштабирование, кодирование категориальных	отдельной моделью по метрике ROC-AUC. Весь код должен быть оформлен в виде воспроизводимого Jupyter Notebook.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		переменных, отбор признаков, обучение модели) в единый сквозной пайплайн, обеспечивая его корректную работу на этапах обучения и применения, а также интегрировать несколько моделей в ансамбль для повышения точности прогнозов	
	Владеет практическим навыком сборки и развертывания многокомпонентных моделей машинного обучения	Знать: современные инструменты и методы сборки многокомпонентных моделей машинного обучения (пайплайны в Scikit-learn, конвейеры данных в Apache	Разработать Docker-контейнер, содержащий многокомпонентный пайплайн машинного обучения (импутация пропусков, масштабирование, кодирование, XGBoost модель), и развернуть его как REST API сервис с помощью FastAPI внутри контейнера. Настроить передачу параметров через переменные окружения (путь к данным, параметры модели), подключить том для сохранения обученной модели и обеспечить возможность переобучения модели через отдельный эндпоинт. Запустить контейнер локально и выполнить тестовые запросы к обоим эндпоинтам, предоставив подтверждение работоспособности в виде логов и скриншотов.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		Airflow/Prefect, системы управления версиями моделей DVC) и развертывания (контейнеризация с Docker, оркестрация с Kubernetes, асинхронные очереди задач Celery/RabbitMQ, сервера моделей TensorFlow Serving, BentoML, MLflow) Уметь: собирать многокомпонентные модели, включающие этапы предобработки данных, обучения и оценки, а также развертывать их в различных средах (ло-	



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		кальной, на сервере, в облаке) с использованием контейнеризации и систем оркестрации для обеспечения масштабируемости и отказоустойчивости.	



Примеры практико-ориентированных заданий

1. На основе предложенного набора данных о клиентах разработать модель прогнозирования оттока, определить целевую переменную, выполнить подготовку данных, построить модель и оценить ее качество.
2. Провести разведочный анализ данных для предложенного датасета, выявить закономерности, аномалии и потенциальные проблемы качества данных, сформулировать выводы.
3. Выполнить предобработку данных: обработать пропущенные значения, выбросы и дубликаты, обосновать выбор применяемых методов.
4. Разработать модель прогнозирования спроса на основе исторических данных, провести оценку качества модели и интерпретацию результатов.
5. Сформировать новые признаки для предложенной задачи и оценить их влияние на качество модели.
6. Выполнить сегментацию клиентов с использованием метода кластеризации, определить характеристики полученных сегментов и предложить возможные направления их использования.
7. Реализовать модель выявления аномалий в данных, проанализировать результаты и оценить их практическую значимость.
8. Разработать рекомендательную систему на основе пользовательских данных, реализовать алгоритм рекомендаций и оценить качество полученных рекомендаций.
9. На основе предложенного прикладного кейса разработать архитектуру решения машинного обучения, определить основные компоненты системы и их взаимодействие.
10. Разработать прототип сервиса машинного обучения (например, API для получения предсказаний модели) и описать процесс его интеграции в прикладную систему.

Примеры вопросов для подготовки к промежуточной аттестации

Примерные вопросы для подготовки к экзамену

1. Охарактеризуйте процесс постановки прикладной задачи машинного обучения на основе бизнес-задачи. Какие этапы он включает?
2. Объясните роль целевой переменной и признаков в задачах машинного обучения. Какие требования к ним предъявляются?
3. Классифицируйте источники данных, используемые в прикладных задачах машинного обучения. Приведите примеры.



4. Раскройте содержание и этапы разведочного анализа данных. Какие задачи он позволяет решить?
5. Проанализируйте влияние качества данных на результаты машинного обучения. Какие виды проблем данных наиболее критичны?
6. Сравните задачи классификации и регрессии. В чем их принципиальные различия и области применения?
7. Объясните назначение базовой модели. Почему она является обязательным этапом разработки модели?
8. Раскройте принципы работы логистической регрессии и области её применения.
9. Охарактеризуйте основные метрики качества моделей классификации (precision, recall, F1, ROC-AUC). В каких случаях применяется каждая из них?
10. Объясните взаимосвязь метрик качества модели и бизнес-эффекта (например, стоимости ошибок).
11. Сравните деревья решений и ансамблевые методы. В чем их преимущества и недостатки?
12. Раскройте понятия переобучения и недообучения. Какие методы используются для их выявления и устранения?
13. Объясните роль feature engineering в задачах машинного обучения. Приведите примеры.
14. Раскройте сущность кластерного анализа и его применение в прикладных задачах.
15. Опишите метод К-средних: алгоритм, преимущества и ограничения.
16. Объясните, как интерпретируются результаты кластеризации и как они используются в бизнесе.
17. Раскройте сущность задачи обнаружения аномалий. Приведите примеры прикладных задач.
18. Сравните методы обнаружения аномалий и области их применения.
19. Раскройте принципы построения рекомендательных систем.
20. Сравните коллаборативную и контентную фильтрацию. В чем их преимущества и ограничения?
21. Объясните проблему «холодного старта» и предложите возможные способы её решения.
22. Охарактеризуйте методы оценки качества рекомендательных систем.
23. Опишите жизненный цикл модели машинного обучения от постановки задачи до внедрения.
24. Раскройте принципы интеграции моделей машинного обучения в прикладные системы.



25. Оцените влияние внедрения модели машинного обучения на бизнес-процессы. Какие метрики используются для оценки эффективности?

Пример экзаменационного билета

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ №

Вопрос 1 (15 баллов). Опишите проблему масштабируемости методов оптимизации при обучении больших языковых моделей (LLM) на данных объёмом более 1 ТБ. Какие существуют подходы к распределённому обучению (параллелизация данных, параллелизация моделей, смешанная параллелизация) и в чём их преимущества и недостатки?

Вопрос 2 (15 баллов). Что такое мультимодальная модель? Опишите архитектуру модели CLIP (Contrastive Language-Image Pre-training). Как происходит обучение совместного эмбединга-пространства для текста и изображения и как эта модель может быть применена в задаче поиска финансовых документов по текстовому запросу?

Вопрос 3 (30 баллов). Практическое задание (комплексное).

Контекст: Команда разработчиков создаёт чат-бота для финансовых консультаций на основе LLM. Исходная базовая модель (например, открытая модель Llama 3 8B) недостаточно хорошо отвечает на вопросы по налоговому законодательству Российской Федерации. У вас есть размеченный датасет из 500 пар «вопрос — правильный ответ» (собранный экспертами-аналитиками из ФНС). Задача — дообучить модель для повышения качества ответов.

Задание (выполняется в виде описания + псевдокода или реального кода в Colab):

1. Выбрать метод дообучения (full fine-tuning, LoRA или Q-LoRA). Обосновать выбор, указав преимущества выбранного метода с точки зрения потребления памяти GPU и скорости обучения.
2. Описать формат подготовки данных для дообучения (например, в стиле Alpaca или инструктивных данных). Привести пример одного обучающего примера.
3. Написать фрагмент кода на Python с использованием библиотеки transformers (Hugging Face) и PEFT для настройки LoRA: загрузка базовой модели, применение LoRA-адаптеров к слоям внимания (query, value), настройка гиперпараметров (rank, alpha, dropout).
4. Описать процесс обучения: функция потерь (например, кросс-энтропия для авторегрессивной генерации), оптимизатор, количество эпох, размер батча с учётом ограничений GPU.
5. Оценить качество дообученной модели: предложить метрики (например, BERTScore, ROUGE-L или экспертную оценку по шкале от 1 до 5) и описать



процедуру валидации на отложенной выборке.

6. Дополнительно (вариативно): предложить способ ускорения инференса после дообучения (например, квантизация модели до 4-bit с помощью bitsandbytes, использование vLLM или ONNX Runtime).

Примерные темы курсовой работы /курсового проекта

1. Разработка системы прогнозирования оттока клиентов с оценкой влияния модели на бизнес-метрики.
2. Разработка модели прогнозирования спроса на товары с использованием методов машинного обучения и анализа временных рядов.
3. Разработка рекомендательной системы для персонализации пользовательского опыта в цифровом сервисе.
4. Разработка модели кредитного скоринга с интерпретацией результатов и оценкой рисков.
5. Разработка системы сегментации клиентов с целью повышения эффективности маркетинговых стратегий.
6. Разработка системы выявления мошеннических операций на основе методов обнаружения аномалий.
7. Разработка модели прогнозирования вероятности совершения покупки пользователем.
8. Разработка системы динамического ценообразования на основе методов машинного обучения.
9. Разработка модели анализа текстовых отзывов пользователей для оценки удовлетворенности клиентов.
10. Разработка модели оценки лояльности клиентов на основе поведенческих данных.



8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Основная литература:

1. Андриянов, Н. А. Построение и оценка моделей машинного обучения. 01.04.02 «Прикладная математика и информатика», всех профилей (программы подготовки магистров) [Электронный ресурс] / Андриянов Н. А., Никитин П. В. Москва : Финансовый университет, 2023 140 с. Книга из коллекции Финансовый университет - Информатика <https://e.lanbook.com/book/345374>. [БИК ID: RU-LAN-BOOK-345374]
2. Коротеев, М.В. Практикум по машинному обучению на Python : учебное пособие по дисциплине "Машинное обучение" для студентов, обучающихся по направлению: "Прикладная информатика", всех профилей (программы подготовки бакалавров), институт онлайн-образования / М.В. Коротеев, В.А. Одинцова, Е.С. Плешакова ; Финуниверситет, Департамент анализа данных и машинного обучения факультета информационных технологий и анализа больших данных Москва : Финуниверситет, 2023 1 файл (1,64 Мб) : ил. Доступ из локальной сети Финуниверситета (чтение) <http://elib.fa.ru/rbook/books137315.pdf> (дата обращения 16.10.2023) + Текст : электронный. [БИК ID: RU\FA\books\137315]
3. Коротеев, М.В. Практикум по машинному обучению на Python. (2 часть) : учебное пособие по дисциплине "Машинное обучение" для студентов, обучающихся по направлению: "Прикладная информатика", всех профилей (программы подготовки бакалавров) / М.В. Коротеев, В.А. Одинцова ; Финуниверситет, Кафедра анализа данных и машинного обучения Факультета информационных технологий и анализа больших данных Москва : Финуниверситет, 2024 1 файл (1,87 Мб) : ил. Свободный доступ из сети Интернет (чтение) <http://elib.fa.ru/fbook/books139159.pdf> (дата обращения 22.05.2024) + Текст : электронный. [БИК ID: RU\FA\books\139159]
4. Коротеев, М.В. Основы машинного обучения на Python : Учебник / М.В. Коротеев Электрон. дан. Москва : КноРус, 2024 431 с. Режим доступа: book.ru Internet access <https://book.ru/book/952751> ISBN 978-5-406-12673-8. [БИК ID: RU\bookru\bibl\952751]

Дополнительная литература:

1. Никитин, П.В. Машинное обучение с подкреплением : Учебник / П.В. Никитин, С.А. Корчагин, Е.В. Романова Электрон. дан. Москва : КноРус,



2026 236 с. Режим доступа: book.ru Internet access <https://book.ru/book/959175> ISBN 978-5-406-15268-3. [БИК ID: RU\bookru\bibl\959175]

2. Золкин, А.Л. Языковые модели (LLM) для интерпретации генерации и адаптации знаний в процессе управленческой и экспертной деятельности : Учебник / А.Л. Золкин, И.А. Рамазанов, Н.Ю. Логунова, Р.А. Вербицкий Электрон. дан. Москва : Русайнс, 2026 200 с. Режим доступа: book.ru Internet access <https://book.ru/book/962323> ISBN 978-5-466-11261-0. [БИК ID: RU\bookru\bibl\962323]

Нормативные правовые акты:

-

9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Информационно-образовательный портал Финансового университета при Правительстве Российской Федерации <http://portal.ufrf.ru/>
2. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/> (<http://library.fa.ru/files/elibfa.pdf>)
3. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
4. Электронно-библиотечная система "Университетская библиотека ОНЛАЙН" <http://biblioclub.ru/>
5. • Электронно-библиотечная система Znanium <http://www.znanium.ru/>
6. Образовательная платформа "ЮРАЙТ" <https://urait.ru/>
7. • Электронно-библиотечная система издательства Проспект <http://ebs.prospekt.org/books>
8. Электронно-библиотечная система издательства "Лань" <https://e.lanbook.com/>
9. • Деловая онлайн-библиотека Alpina Digital <http://lib.alpinadigital.ru/>
10. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>
11. • Национальная электронная библиотека <http://нэб.рф/>



10. Методические указания для обучающихся по освоению дисциплины

Лекционные занятия направлены на формирование теоретических основ прикладного анализа данных и машинного обучения, а также на развитие понимания их применения при решении практических задач. В процессе освоения лекционного материала обучающимся рекомендуется систематически посещать занятия и вести конспект, отражающий ключевые понятия, методы и подходы, рассматриваемые в рамках дисциплины. Особое внимание следует уделять взаимосвязи теоретических моделей с прикладными задачами, в том числе задачами, возникающими в бизнес-практике.

Важным элементом освоения дисциплины является анализ примеров применения методов машинного обучения в различных предметных областях, что позволяет сформировать целостное представление о возможностях и ограничениях используемых подходов. Обучающимся также рекомендуется самостоятельно изучать дополнительные источники информации по рассматриваемым темам, а возникающие вопросы и затруднения выносить на обсуждение в рамках практических занятий.

Особое внимание в ходе изучения лекционного материала следует уделять вопросам постановки задач машинного обучения, интерпретации результатов моделирования, а также анализу взаимосвязи между качеством исходных данных, выбором методов и итоговой эффективностью решений. Существенную роль играет понимание места машинного обучения в процессе поддержки принятия решений.

Практические занятия по дисциплине организованы по проектному принципу и направлены на формирование устойчивых навыков решения прикладных задач машинного обучения. В ходе выполнения практических работ обучающиеся последовательно проходят основные этапы разработки решений, включая постановку задачи, подготовку и анализ данных, построение моделей, оценку их качества и интерпретацию полученных результатов.

Особое значение придается активному участию обучающихся в анализе прикладных кейсов, что способствует развитию способности применять теоретические знания на практике. В процессе выполнения заданий рекомендуется использовать современные инструментальные средства обработки данных и реализации моделей машинного обучения, уделяя внимание корректности выполнения всех этапов работы, а также качеству получаемых результатов.

Результаты выполнения практических заданий должны оформляться в виде отчетов, содержащих описание поставленной задачи, обоснование выбранных методов, анализ полученных результатов и выводы о возможности их практического применения. В рамках самостоятельной работы обучающимся рекомендуется



дорабатывать выполненные задания, проводить дополнительный анализ данных и моделей, а также изучать альтернативные подходы к решению поставленных задач.

Практическая направленность дисциплины предполагает необходимость регулярной работы с данными и инструментами анализа, а также формирование навыков комплексного подхода к решению задач машинного обучения. Важным результатом освоения дисциплины является способность разрабатывать прикладные решения, ориентированные на достижение практического эффекта и использование в реальных условиях.

При подготовке курсового проекта следует использовать нормативные документы Финансового университета, Методические рекомендации по планированию и организации внеаудиторной самостоятельной работы студентов по образовательным программам бакалавриата и магистратуры в Финансовом университете, утвержденные приказом Финуниверситета от 02.07.2021 г. № 1583/о (см. сайт Финансового Университета: на главной странице раздел "Наш университет" → "Единая правовая база Финуниверситета" → "Организация учебного процесса" → "Нормативные документы по самостоятельной работе").



11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем

Комплект лицензионного программного обеспечения:

1. Python 3.8
2. LMS Moodle
3. Docker
4. Реляционная СУБД
5. Anaconda
6. библиотеки Keras
7. PyTorch
8. Yandex Cloud
9. Mail Cloud
10. Виртуальные машины
11. Git (клиент)
12. Среда разработки (PyCharm, VS Code или др.)
13. Антивирус Kaspersky
14. Пакет офисных программ

Современные профессиональные базы данных и информационные справочные системы

1. Информационно-правовая система «Гарант»
2. Информационно-правовая система «Консультант Плюс»
3. Электронная энциклопедия: <http://ru.wikipedia.org/wiki/Wiki>
4. Система комплексного раскрытия информации «СКРИН» - <http://www.skrin.ru/>

Сертифицированные программные и аппаратные средства защиты информации

1. не предусмотрены

12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

1. **Учебная аудитория** для проведения учебных занятий, предусмотренных программой, в том числе групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенная оборудованием и техническими средствами обучения: мебель аудиторная



(столы, стулья, доска аудиторная), набор демонстрационного оборудования (проектор, экран)

2. **Помещение для самостоятельной работы** обучающихся оснащенное компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа к электронной информационно-образовательной среде Университета.
3. **Компьютерный класс** для проведения учебных занятий, предусмотренных программой, в том числе групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащенный оборудованием и техническими средствами обучения: мебель аудиторная (столы, стулья, доска аудиторная), персональные компьютеры, набор демонстрационного оборудования (проектор, экран)