



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: Z2g099ThSiJ8r/IokALaTyBFuLalP4Si

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

Дата: 02.06.2026 г.

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

Прикладные задачи машинного обучения

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки

09.03.03 - Прикладная информатика,

Образовательная программа «Прикладные информационные системы в экономике и
финансах»

Профиль:

«Инженерия данных»

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	6
4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	6
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий	7
5.1. Содержание дисциплины	7
5.2. Учебно – тематический план	9
5.3. Содержание семинаров, практических занятий	10
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	12
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы	12
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю	14
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине	15
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	6
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	9
10. Методические указания для обучающихся по освоению дисциплины	26
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)	28
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	28



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
ПKN-4	Способность проектировать и создавать интеллектуальные информационные системы, выбирать метод обучения в соответствии с анализом задачи	Демонстрирует знание основных понятий машинного обучения и интеллектуального анализа данных, понимание области и границ применимости, основные виды задач	Знать: основные понятия и типы задач машинного обучения (классификация, регрессия, кластеризация), критерии выбора метода обучения в зависимости от структуры данных и бизнес-требований, а также границы применимости различных алгоритмов Уметь: анализировать прикладную задачу и формулировать её в терминах машинного обучения, определять тип задачи и обосновывать предварительный выбор методов
		Демонстрирует знание популярных инструментальных средств машинного обучения, собирает датасет, строит модели, проводит их анализ и диагностику, делает содержательные выводы	Знать: популярные библиотеки и фреймворки для машинного обучения (Scikit-learn, Pandas, NumPy, XGBoost, LightGBM), основные этапы построения модели (предобработка данных, обучение, валидация, оценка качества) и методы диагностики (анализ переобучения, кросс-валидация, подбор гиперпараметров) Уметь: загружать и предобрабатывать данные, строить модели машинного обучения с использованием Scikit-learn, оценивать их качество с помощью метрик (accuracy,



Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
			F1, ROC-AUC, MSE) и интерпретировать полученные результаты для принятия решений
		Презентабельно демонстрирует результаты анализа данных и машинного обучения в форме, доступной непрофессионалу, структурирует отчет по проведенному анализу	Знать: способы визуализации результатов моделирования (матрица ошибок, ROC-кривая, важность признаков) и методы представления аналитических выводов для различных аудиторий (бизнес-заказчики, технические специалисты, руководство) Уметь: готовить отчёты о результатах анализа данных, включая визуализации и интерпретацию ключевых метрик, формулировать практические рекомендации на основе построенных моделей и адаптировать сложность изложения для неподготовленной аудитории
ПКП-1	Способность формировать требования, производить подготовку данных и проводить аналитические работы с использованием технологий больших данных	Демонстрирует знания в области компьютерных технологий, используемых для анализа больших данных	Знать: современные компьютерные технологии для работы с большими данными (распределенные вычисления, принципы работы Apache Spark и Hadoop, инструменты для хранения и обработки больших объёмов данных), их возможности и ограничения Уметь: выбирать и обосновывать применение конкретных технологий работы с большими данными в зависимости от объёма данных, скорости поступления и требований к обработке (пакетная или потоковая)
		Владеет навыками управления аналитическими работами на основе больших данных	Знать: методы и подходы к управлению аналитическими проектами в области больших данных (организация ETL/ELT-процессов, контроль качества данных, версионирование)



Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
			рование данных, управление метаданными), а также инструменты для организации пайплайнов обработки данных (Apache Airflow, Dagster) Уметь: планировать и организовывать процессы сбора, очистки, трансформации и агрегации больших данных из разнородных источников, обеспечивать воспроизводимость и масштабируемость аналитических процессов
		Выполняет аналитические работы на основе больших данных	Знать: методы и алгоритмы анализа больших данных (анализ временных рядов, обнаружение аномалий, масштабируемые версии классических алгоритмов ML), особенности их применения в распределённых средах Уметь: реализовывать аналитические алгоритмы с использованием инструментов для больших данных (PySpark, Dask), проводить исследовательский анализ больших массивов данных и интерпретировать полученные закономерности



3. Место дисциплины в структуре образовательной программы

Дисциплина «Прикладные задачи машинного обучения» относится к «Модулю "Технологии машинного обучения"».

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	3/108	108
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>58</i>	<i>58</i>
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	Зачет	Зачет



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. Введение в прикладное машинное обучение и глубокое обучение

От классических методов к глубокому обучению. Искусственные нейронные сети: перцептрон, многослойные сети, функции активации, обратное распространение ошибки. Понятие глубокого обучения. Обзор прикладных областей: компьютерное зрение, обработка естественного языка, рекомендательные системы, временные ряды. Платформы и фреймворки: PyTorch, TensorFlow, Hugging Face. Применение в финансовой сфере: анализ транзакций, прогнозирование временных рядов, обработка текстов документов.

Тема 2. Глубокое обучение в обработке естественного языка (NLP)

Представление текста: токенизация (BPE, SentencePiece), эмбединги (Word2Vec, GloVe, FastText). Рекуррентные нейронные сети (RNN, LSTM, GRU). Механизм внимания и архитектура трансформера. Современные модели: BERT, GPT, Llama, Mistral. Инструменты: Hugging Face Transformers. Прикладные задачи: классификация текстов (анализ тональности), извлечение именованных сущностей (NER), суммаризация, генерация текста. Кейсы: анализ финансовых новостей, обработка обращений клиентов, автоматическое заполнение документов.

Тема 3. Прикладные задачи компьютерного зрения в экономике и финансах

Свёрточные нейронные сети (CNN): архитектура, обучение фильтров. Современные архитектуры: ResNet, EfficientNet, Vision Transformer (ViT). Трансфер обучение: использование предобученных моделей. Детекция объектов: YOLO, DETR. Сегментация: U-Net, SAM. Генерация изображений: диффузионные модели, Stable Diffusion. Инструменты: OpenCV, ultralytics, diffusers. Прикладные задачи: распознавание документов (счета-фактуры, чеки, паспорта), анализ графиков и диаграмм, проверка подлинности подписей, визуальный мониторинг банкоматов и офисов.

Тема 4. Рекомендательные системы и прогнозирование в финансовой сфере



Классические подходы: коллаборативная фильтрация, матричная факторизация (SVD). Глубокое обучение в рекомендациях: нейронные коллаборативные фильтры (NCF), двухбашенные модели (two-tower), трансформеры для последовательностей (SASRec, BERT4Rec). Прогнозирование временных рядов: ARIMA, Prophet, LSTM, Temporal Fusion Transformer. Применение: рекомендация финансовых продуктов (кредиты, депозиты, страховые продукты), прогнозирование оттока клиентов (churn prediction), оценка кредитного риска по последовательностям транзакций, прогнозирование цен активов.



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	Введение в прикладное машинное обучение и глубокое обучение	26	12	4	8	14
2	Глубокое обучение в обработке естественного языка (NLP)	26	12	4	8	14
3	Прикладные задачи компьютерного зрения в экономике и финансах	28	14	4	10	14
4	Рекомендательные системы и прогнозирование в финансовой сфере	28	12	4	8	16
В целом по дисциплине		108	50	16	34	58



5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарах, практических занятиях	Формы проведения занятий
Введение в прикладное машинное обучение и глубокое обучение	1. Реализация многослойного перцептрона на PyTorch для классификации рукописных цифр MNIST. 2. Сравнение классических методов (логистическая регрессия) и нейронной сети. 3. Обсуждение областей применения глубокого обучения в финансах. 4. Настройка среды разработки (Google Colab, PyTorch).	Интерактивная форма, практикум в малых группах (2–4 студента), коллективное обсуждение решений, командные проектные задания.
Глубокое обучение в обработке естественного языка (NLP)	1. Токенизация текста с использованием библиотеки transformers. 2. Классификация тональности отзывов клиентов банка с использованием LSTM. 3. Дообучение (fine-tuning) модели BERT для задачи извлечения именованных сущностей (NER) из финансовых новостей. 4. Генерация текста с использованием GPT.	Интерактивная форма, практикум в малых группах (2–4 студента), коллективное обсуждение решений, командные проектные задания.
Прикладные задачи компьютерного зрения в экономике и финансах	1. Классификация изображений с использованием предобученной модели ResNet50 (трансфер обучение). 2. Распознавание документов: детекция и классификация полей на счёте-фактуре с использованием YOLO. 3. Сегментация изображений для анализа графиков и диаграмм.	Интерактивная форма, практикум в малых группах (2–4 студента), коллективное обсуждение решений, командные проектные задания.



Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарах, практических занятиях	Формы проведения занятий
	4. Генерация изображений с использованием Stable Diffusion.	
Рекомендательные системы и прогнозирование в финансовой сфере	1. Реализация матричной факторизации (SVD) для рекомендации кредитных продуктов на датасете транзакций. 2. Построение нейронной коллаборативной фильтрации (NCF) с использованием PyTorch. 3. Прогнозирование оттока клиентов банка с использованием LSTM на последовательностях транзакций. 4. Прогнозирование временных рядов (курс валют, цены акций) с использованием Prophet и LSTM.	Интерактивная форма, практикум в малых группах (2–4 студента), коллективное обсуждение решений, 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
Введение в прикладное машинное обучение и глубокое обучение	1. В чём заключается отличие глубокого обучения от классических методов машинного обучения? 2. Как работает алгоритм обратного распространения ошибки и какие проблемы могут возникать при его применении? 3. Какие функции активации используются в глубоких нейронных сетях и в чём их особенности? 4. Что такое переобучение и какие методы регуляризации применяются для его предотвращения? 5. Каковы основные этапы сквозного пайплайна машинного обучения в прикладном проекте?	Работа с учебной литературой. Разбор вопросов по теме занятия. Выполнение домашних заданий.
Глубокое обучение в обработке естественного языка (NLP)	1. Какие существуют методы векторизации текста и в чём преимущества эмбедингов перед one-hot encoding? 2. Как устроен механизм внимания и почему он стал ключевым в современных NLP-моделях? 3. В чём различие между архитектурами BERT и GPT с точки зрения задач и обучения? 4. Какие методы дообучения (fine-tuning) больших языковых моделей существуют и когда они применяются? 5. Как можно оценить качество модели генерации текста? Какие метрики для этого используются?	Работа с учебной литературой. Разбор вопросов по теме занятия. Выполнение домашних заданий.



Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Прикладные задачи компьютерного зрения в экономике и финансах	1. Как устроен свёрточный слой и чем он отличается от полносвязного? 2. Что такое трансфер обучение и как оно применяется в компьютерном зрении? 3. Какие существуют подходы к детекции объектов (двухэтапные и одноэтапные) и в чём их различия? 4. В чём разница между классификацией, детекцией и сегментацией изображений? 5. Как генеративные модели (например, диффузионные) могут быть использованы в финансовых приложениях?	Работа с учебной литературой. Разбор вопросов по теме занятия. Выполнение домашних заданий.
Рекомендательные системы и прогнозирование в финансовой сфере	1. В чём отличие коллаборативной фильтрации от контентной фильтрации в рекомендательных системах? 2. Как работает матричная факторизация и для решения каких задач она применяется? 3. Какие метрики используются для оценки качества рекомендаций и в чём их особенности? 4. Какие модели машинного обучения применяются для прогнозирования временных рядов и в чём их преимущества? 5. Как можно использовать последовательности транзакций для прогнозирования поведения клиента?	Работа с учебной литературой. Разбор вопросов по теме занятия. Выполнение домашних заданий.



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

Примерные задания контрольной работы

Задание 1. На датасете Fashion-MNIST сравнить качество классификации двух подходов:

1. логистическая регрессия (из библиотеки Scikit-learn);
2. многослойный перцептрон с двумя скрытыми слоями (на PyTorch).

Для каждого метода вычислить accuracy, precision, recall и F1-score. Сделать вывод о том, в каких случаях нейронная сеть даёт преимущество перед классическим методом.

Задание 2. На датасете классификации тональности (например, отзывы клиентов) сравнить четыре метода векторизации текста:

1. Bag-of-Words (CountVectorizer);
2. TF-IDF;
3. Word2Vec (предобученные векторы);
4. эмбединги из BERT (с замороженной моделью).

Для каждого метода обучить логистическую регрессию или простую нейронную сеть. Сравнить accuracy, время обучения и размер полученных векторов. Сделать вывод о применимости каждого метода в зависимости от объёма данных и вычислительных ресурсов.

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
ПKN-4. Способность проектировать и создавать интеллектуальные информационные системы, выбирать метод обучения в соответствии с анализом задачи	Демонстрирует знание основных понятий машинного обучения и интеллектуального анализа данных, понимание области и границ применимости, основные виды задач	Знать: основные понятия и типы задач машинного обучения (классификация, регрессия, кластеризация), критерии выбора метода обучения в зависимости от структуры данных и бизнес-требований, а также границы применимости различных алгоритмов	По описанию бизнес-задачи (например, прогнозирование оттока клиентов, сегментация налогоплательщиков, обнаружение мошеннических операций) определить тип задачи (классификация, регрессия, кластеризация), сформулировать требования к данным и предложить не менее двух подходящих методов машинного обучения с обоснованием выбора.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		Уметь: анализировать прикладную задачу и формулировать её в терминах машинного обучения, определять тип задачи и обосновывать предварительный выбор методов	
	Демонстрирует знание популярных инструментов машинного обучения, собирает датасет, строит модели, проводит их анализ и диагностику, делает содержательные выводы	Знать: популярные библиотеки и фреймворки для машинного обучения (Scikit-learn, Pandas, NumPy, XGBoost, LightGBM), основные этапы построения модели (предобработка данных, обучение, валидация, оценка качества) и методы диа-	На основе предоставленного датасета (например, данные о клиентах банка) выполнить полный цикл построения модели классификации: предобработать данные, обучить модели (логистическая регрессия, случайный лес, градиентный бустинг), сравнить их качество по метрикам F1 и ROC-AUC, выбрать лучшую модель и интерпретировать её результаты.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		гностики (анализ переобучения, кросс-валидация, подбор гиперпараметров) Уметь: загружать и преобразовывать данные, строить модели машинного обучения с использованием Scikit-learn, оценивать их качество с помощью метрик (accuracy, F1, ROC-AUC, MSE) и интерпретировать полученные результаты для принятия решений	
	Презентабельно демонстрирует результаты анализа данных	Знать: способы визуализации результатов моделирования (мат-	Подготовить презентацию (5-7 слайдов) по результатам проекта «Прогнозирование банкротства предприятий», включающую: описание задачи и данных, выбор модели, ключевые метрики качества, визуализацию важности признаков и практические реко-



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
	и машинного обучения в форме, доступной непрофессионалу, структурирует отчет по проведенному анализу	рица ошибок, ROC-кривая, важность признаков) и методы представления аналитических выводов для различных аудиторий (бизнес-заказчики, технические специалисты, руководство) Уметь: готовить отчёты о результатах анализа данных, включая визуализации и интерпретацию ключевых метрик, формулировать практические рекомендации на основе построенных моделей и адаптировать слож-	мендации для бизнеса. Защитить презентацию перед аудиторией с последующей дискуссией.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		ность изложения для неподготовленной аудитории	
ПКП-1. Способность формировать требования, производить подготовку данных и проводить аналитические работы с использованием технологий больших данных	Демонстрирует знания в области компьютерных технологий, используемых для анализа больших данных	Знать: современные компьютерные технологии для работы с большими данными (распределенные вычисления, принципы работы Apache Spark и Hadoop, инструменты для хранения и обработки больших объёмов данных), их возможности и ограничения Уметь: выбирать и обосновывать применение конкретных технологий работы с большими данными в	Для задачи обработки потоковых данных (логи посещения веб-сайта в реальном времени, объём 1 ТБ в сутки) выбрать и обосновать технологический стек для их хранения и анализа, сравнив не менее двух альтернативных решений (например, Apache Spark Streaming vs. Kafka Streams) по критериям масштабируемости, задержки и сложности внедрения.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		зависимости от объёма данных, скорости поступления и требований к обработке (пакетная или потоковая)	
	Владеет навыками управления аналитическими работами на основе больших данных	Знать: методы и подходы к управлению аналитическими проектами в области больших данных (организация ETL/ELT-процессов, контроль качества данных, версионирование данных, управление метаданными), а также инструменты для организации пайплайнов обработки данных (Apache Airflow,	Разработать архитектуру пайплайна обработки больших данных для агрегации логов мобильного приложения (10 млн событий в день) с использованием PySpark. Реализовать ETL-процесс: загрузка данных из Kafka, очистка, обогащение, агрегация и сохранение результатов в партиционированный parquet-файл.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		Dagster) Уметь: планировать и организовывать процессы сбора, очистки, трансформации и агрегации больших данных из разнородных источников, обеспечивать воспроизводимость и масштабируемость аналитических процессов	
	Выполняет аналитические работы на основе больших данных	Знать: методы и алгоритмы анализа больших данных (анализ временных рядов, обнаружение аномалий, масштабируемые версии классических алгоритмов	На датасете с 10 млн записей о транзакциях клиентов выполнить исследовательский анализ с использованием PySpark: рассчитать агрегированные метрики (сумма, среднее, медиана) в разрезе категорий товаров и регионов, выявить топ-5 аномалий по сумме транзакций с помощью метода Z-оценки, визуализировать распределения с помощью гистограмм и ящиков с усами.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		ML), особенности их применения в распределённых средах Уметь: реализовывать аналитические алгоритмы с использованием инструментов для больших данных (PySpark, Dask), проводить исследовательский анализ больших массивов данных и интерпретировать полученные закономерности	



Примеры практико-ориентированных заданий

Задание 1. Реализовать пайплайн классификации тональности отзывов клиентов банка с использованием LSTM и сравнить с линейной моделью на TF-IDF.

Задание 2. Выполнить дообучение модели BERT для извлечения именованных сущностей (дата, сумма, номер счёта) из текстов финансовых документов.

Задание 3. Применить предобученную модель YOLO для детекции полей на изображении счёта-фактуры и извлечения ключевой информации.

Задание 4. Построить рекомендательную систему кредитных продуктов на основе истории транзакций клиента с использованием коллаборативной фильтрации.

Задание 5. Разработать модель прогнозирования оттока клиентов на основе последовательностей транзакций с использованием LSTM.

Примеры вопросов для подготовки к промежуточной аттестации

Примерные вопросы для подготовки к зачету

1. В чём отличие глубокого обучения от классических методов машинного обучения?
2. Опишите архитектуру многослойного перцептрона. Как происходит обратное распространение ошибки?
3. Какие методы векторизации текста существуют? В чём преимущество эмбедингов?
4. Как устроен механизм внимания в трансформерах?
5. Сравните архитектуры BERT и GPT.
6. Что такое трансфер обучение и как оно применяется в компьютерном зрении?
7. Опишите архитектуру свёрточной нейронной сети. Какие основные типы слоёв в ней используются?
8. В чём разница между детекцией и сегментацией объектов?
9. Как работает коллаборативная фильтрация в рекомендательных системах?
10. Какие метрики используются для оценки качества рекомендаций?
11. Какие модели применяются для прогнозирования временных рядов? Назовите их преимущества и недостатки.
12. Как можно оценить качество модели классификации? Приведите примеры метрик.



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

Основная литература:

1. Татарникова, Т. М. Методы машинного обучения [Электронный ресурс] : учеб. пособие / Татарникова Т. М., Боженко В. В. Санкт-Петербург : ГУАП, 2023 100 с. Книга из коллекции ГУАП - Информатика <https://e.lanbook.com/book/461489> ISBN 978-5-8088-1885-9. [БИК ID: RU-LAN-BOOK-461489]
2. Андриянов, Н.А. Построение и оценка моделей машинного обучения : учебное пособие по дисциплине "Построение и оценка моделей машинного обучения" для студентов, обучающихся по направлению "Прикладная математика и информатика" всех профилей (программы подготовки магистров) / Н.А. Андриянов, П.В. Никитин ; Финуниверситет, Департамент анализа данных и машинного обучения Факультета информационных технологий и анализа больших данных Москва : Финансовый университет, 2023 1 файл (6,31 Мб) Свободный доступ из сети Интернет (чтение, печать) (дата обращения : 07.06.2023) ISBN 978-5-79421-445-1 + Текст : электронный. [БИК ID: RU\FA\books\136568]

Дополнительная литература:

1. Сахарова, Л. В. Методы оптимизации для машинного обучения : учебное пособие / Л. В. Сахарова, Г. В. Лукьянова ; Ростовский государственный экономический университет (РИНХ) Ростов-на-Дону : Издательско-полиграфический комплекс РГЭУ (РИНХ), 2023 87 с. : ил. Библиогр. в кн Режим доступа: электронная библиотечная система «Университетская библиотека ONLINE», требуется авторизация <https://biblioclub.ru/index.php?page=book&id=711220> ISBN 978-5-7972-3139-4. [БИК ID: BIBLIOCLUB\0000711220]
2. Фомичева, С. Г. Методы машинного обучения в задачах обеспечения информационной безопасности [Электронный ресурс] : учеб. пособие / Фомичева С. Г. Санкт-Петербург : ГУАП, 2023 136 с. Книга из коллекции ГУАП - Информатика <https://e.lanbook.com/book/341024> ISBN 978-5-8088-1822-4. [БИК ID: RU-LAN-BOOK-341024]

Нормативные правовые акты:

-



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

1. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
2. Электронно-библиотечная система "Университетская библиотека ОНЛАЙН" <http://biblioclub.ru/>
3. Электронно-библиотечная система Znanium <http://www.znanium.com>
4. Электронно-библиотечная система издательства "Лань" <https://e.lanbook.com/>
5. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>
6. Профессиональный ресурс по машинному обучению. – <https://machinelearning.ru>
7. Специализация «Машинное обучение и анализ данных» / МФТИ и Яндекс. – <https://www.coursera.org/specializations/machine-learning-data-analysis>
8. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/>
(<http://library.fa.ru/files/elibfa.pdf>)



10. Методические указания для обучающихся по освоению дисциплины

Рекомендации по подготовке к лекционным занятиям:

При подготовке к лекции целесообразно предварительно познакомиться с ее содержанием по рекомендованным источникам и выделить наиболее трудные вопросы. Во время лекций следует конспектировать содержание лекции. После занятий следует провести работу с конспектом: отредактировать записи, оформить конспект. При оформлении целесообразно выделять специальным образом названия тем и формулировки вопросов, основные определения, формулировки теорем и примеры. Сделанные записи нужно сопоставить с учебниками и учебными пособиями и в случае расхождений проконсультироваться с преподавателем.

Методические указания по проведению практических занятий:

Практические занятия структурно состоят из следующих компонент:

1. проверка наличия выполненного задания самостоятельной работы каждого студента;
2. выборочная проверка корректности выполнения домашнего задания;
3. разбор типичных ошибок, возникших в самостоятельной работе;
4. рассмотрение теоретических вопросов, связанных с текущим практическим занятием;
5. разбор методов выполнения практических заданий и решения задач;
6. корректировка заданий для самостоятельной работы студентов;
7. интерактивная форма – решение лабораторных работ по тематике занятия в малых группах (2-4 студента).

При подготовке к практическому занятию необходимо повторить или, если это требуется, изучить соответствующий теоретический материал. Во время занятия нужно точно следовать указаниям преподавателя к решению лабораторных работ. После занятий необходимо просмотреть полученные решения и восстановить имеющиеся пробелы или закончить решение. В случае затруднений отметить соответствующие задания и обратиться за консультацией к преподавателю.

Практические занятия проходят в интерактивной форме. Наряду с правильностью решения преподаватель учитывает также и активность студентов, направленную на решение предложенных задач.

При подготовке к текущему контролю 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 11.05.2021 г. № 1040/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. Jupyter Notebook (Windows 10)
3. LMS Moodle
4. Реляционная СУБД
5. библиотеки Keras
6. PyTorch
7. Антивирус Kaspersky

Современные профессиональные базы данных и информационные справочные системы

1. Язык программирования Python 3. <https://pythonworld.ru/>
2. Размещение ЭУК: <https://www.campus.fa.ru/>
3. Информационно-правовая система «Гарант»
4. Информационно-правовая система «Консультант Плюс»
5. Электронная энциклопедия: <http://ru.wikipedia.org/wiki/Wiki>
6. 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. **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)
2. 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)