

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: LSIgJZMTTrLuhUMkNA2+taYfAPIuTSeSR

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

Дата: 27.01.2026 г.

Р.В. Семенов , Е.А. Сальников , Э.Ф. Болтачев
Машинное обучение и искусственный интеллект в бизнесе

Рабочая программа дисциплины

для студентов, обучающихся по направлению подготовки

38.03.01 - Экономика,

Образовательная программа «Финансы и искусственный интеллект»

Профиль:

«Финансы и искусственный интеллект»

Рекомендовано

*Факультет информационных технологий и анализа больших данных
(протокол № 07 от 22.04.2026 г.)*

Одобрено

*Кафедра искусственного интеллекта
(протокол № 8 от 07.04.2026 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	10
5.3. Содержание семинаров, практических занятий	11
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	13
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине	14
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	13
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	13
10. Методические указания для обучающихся по освоению дисциплины	30
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем (при необходимости)	32
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	32



1. Наименование дисциплины

Машинное обучение и искусственный интеллект в бизнесе

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
УК-4	Способность использовать прикладное программное обеспечение при решении профессиональных задач	Использует основные методы и средства получения, представления, хранения и обработки данных	Знать: назначение и возможности прикладного ПО, используемого в задачах бизнес-аналитики и ML/AI (табличные процессоры, BI-инструменты, СУБД-клиенты, Python-экосистема для анализа данных и ML) Уметь: импортировать данные из типовых источников (файлы/БД), выполнять очистку, трансформации и подготовку признаков в выбранном инструментарии
		Демонстрирует владение профессиональными пакетами прикладных программ	Знать: типовые форматы данных и способы их получения, хранения и обработки, базовые принципы подготовки данных к анализу Уметь: применять прикладные программные средства для построения ML-решения: обучение модели, оценка качества, сравнение альтернатив, визуализация и интерпретация результатов для бизнес-пользователя
		Выбирает необходимое прикладное программное обеспечение в зависимости от решаемой задачи	Знать: базовые принципы воспроизводимости вычислений и документирования результатов как требование к профессиональному использованию ПО при решении прикладных задач



Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
			Уметь: обоснованно выбирать инструменты под задачу (например: BI для отчетности и витрин, Python для анализа и моделирования; локальный запуск — при необходимости воспроизводимости)
		Использует прикладное программное обеспечение для решения конкретных прикладных задач	Знать: конкретные возможности прикладного программного обеспечения (Python с библиотеками pandas/sklearn, табличные процессоры, BI-инструменты, СУБД-клиенты) для решения типовых прикладных задач в области экономики и финансов; типовые алгоритмы и последовательности действий в прикладном ПО для решения задач анализа данных, прогнозирования, оценки рисков и подготовки отчётности Уметь: выбирать и применять конкретные функции, методы и библиотеки прикладного ПО для решения поставленной прикладной задачи (очистка данных, построение модели, визуализация, формирование отчёта); настраивать параметры прикладного ПО и адаптировать его под особенности конкретной задачи (форматы данных, ограничения по времени, требования к воспроизводимости)
ПКП-3	Способность проводить анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры	Подготавливает данные для проведения работ по исследованию больших данных: извлечение, проверка, очистка, агрегация и оценка	Знать: модели и методы прогнозной финансовой аналитики, основанные на анализе больших данных и применении моделей машинного обучения; способы оценки качества прогнозных моделей для финансовых временных рядов (MAE, RMAPE, VaR-тесты) Уметь: применять методы анализа больших данных и мо-



Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
			дели машинного обучения для эффективной оценки финансовых инструментов и инвестиционного планирования; адаптировать модели машинного обучения под специфику финансовых данных (нестационарность, шумы, выбросы)
		Использует современные методы, стандарты, инструментальные средства и языки программирования для проведения анализа данных	Знать: аналитические инструменты и методики оценки и моделирования финансовых рисков инвестирования в различные проекты; методы стресс-тестирования и сценарного анализа инвестиционных проектов Уметь: использовать современный аналитический инструментарий для объективной оценки и минимизации рисков вложения капитала в финансовые активы и инвестиционные проекты
		Интерпретирует и визуализирует большие данные, полученные в результате анализа в соответствии с целями исследования	Знать: специфику интерпретации результатов анализа больших данных для разных аудиторий: перевод технических метрик (точность, Recall) на язык бизнеса (прибыль, снижение рисков) для менеджмента и регуляторов; методы и инструменты визуализации, эффективные для больших объёмов данных: интерактивные дашборды для исследования срезов, тепловые карты для выявления аномалий, визуализация сложных моделей (например, деревьев решений) для объяснения логики Уметь: интерпретировать результаты анализа больших данных для оценки их влияния на ключевые финансовые показатели и риски компании; создавать наглядные и ин-



Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
			терапевтивные визуализации, отвечающие на конкретные вопросы бизнеса и соответствующие целям исследования



3. Место дисциплины в структуре образовательной программы

Дисциплина «Машинное обучение и искусственный интеллект в бизнесе» относится к «Профилю "Финансы и искусственный интеллект"».

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

очно-заочная форма обучения

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 6 (в часах)
Общая трудоемкость дисциплины	4/144	144
Контактная работа – Аудиторные занятия	32	32
<i>Лекции</i>	<i>8</i>	<i>8</i>
<i>Семинары, практические занятия</i>	<i>24</i>	<i>24</i>
<i>Самостоятельная работа</i>	<i>112</i>	<i>112</i>
Вид текущего контроля	Проектная работа	Проектная работа
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. Введение в машинное обучение и искусственный интеллект

Понятие и виды искусственного интеллекта. Основные определения: интеллектуальный анализ данных, большие данные, машинное обучение, искусственный интеллект. Развитие искусственного интеллекта. История возникновения и этапы развития машинного обучения. Оценка доходности сектора высоких технологий.

Доходность применения искусственного интеллекта.

Большие данные и их основные характеристики (6V). Основные этапы цикла обработки данных. Интеллектуальный анализ данных и его роль в бизнесе. Применение машинного обучения в различных отраслях бизнеса. Ключевые концепции: выборка, признаки, целевые переменные, метрики оценки. Методы линейной алгебры и математической статистики применительно к обработке данных

Типы задач машинного обучения: классификация, регрессия, кластеризация, рекомендательные системы. Концепция переобучения и недообучения. Подбор гиперпараметров. Описание и постановка типовых задач машинного обучения (анализ кредитных рисков, маркетинговая сегментация клиентов).

Методы совершенствования моделей машинного обучения. Оптимизация гиперпараметров моделей. Обработка и предварительная подготовка данных. Создание синтетических признаков для обогащения набора данных. Работа с пропущенными данными. Предотвращение эффекта переобучения.

Снижение размерности факторного пространства. Метод главных компонент и его компьютерная реализация. Применение методов снижения размерности для улучшения качества моделей машинного обучения.

Нейронные сети как прогностический и оценочный инструмент. Их виды, архитектуры, методы обучения и оценки качества.

Тема 2. Алгоритмы и методы машинного обучения

Линейные модели (регрессии, логистическая регрессия): принцип работы, достоинства и недостатки. Задача прогнозирования продаж. Модель множественной линейной регрессии. Методы оценки качества моделей регрессии. Средняя



абсолютная ошибка предсказания, средняя относительная ошибка предсказания, коэффициент детерминации.

Задача кредитного скоринга. Модель логистической регрессии и ее реализация. Критерии оценки качества моделей классификации: полнота, F1, AUC, доля правильных ответов, точность. Ошибки первого и второго рода.

Деревья решений и случайные леса: структура, способы улучшения точности. Усиление (бустинг) деревьев решений. Бэггинг. Случайный лес. Метод опорных векторов (SVM): концепция, типы ядер, настройки параметров.

Байесовская классификация и вероятностные подходы. Метод ближайших соседей (kNN) и его особенности.

Тема 3. Практическое применение моделей машинного обучения в бизнесе

Алгоритмы обработки данных (EDA).

Маркетинг: персонализация предложений, сегментирование аудитории, повышение конверсии. Кластерный анализ и его компьютерная реализация. Задача сегментирования потребителей. Метод К-средних.

Управление рисками: кредитный скоринг, предотвращение мошенничества, прогноз дефолтов. Финансово-экономические задачи.

A/B тесты. Задача скоринга в машинном обучении. Прогнозирование сбыта и прогнозирование поставок на основе библиотеки NumPy.

Продажи и ценообразование: прогнозирование спроса, динамическое ценообразование. Баланс спрос/предложение - методы предиктивной оценки.



5.2. Учебно – тематический план

Очно-заочная 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 машинное обучение и искусственный интеллект	38	8	2	6	30
2	Алгоритмы и методы машинного обучения	53	12	3	9	41
3	Практическое применение моделей машинного обучения в бизнесе	53	12	3	9	41
В целом по дисциплине		144	32	8	24	112



5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарах, практических занятиях	Формы проведения занятий
Введение в машинное обучение и искусственный интеллект	Построение компьютерной системы визуализации данных о транзакциях покупателей в розничном магазине.	Лабораторная работа, интерактивная форма.
Алгоритмы и методы машинного обучения	Построение системы прогнозирования спроса и продаж розничного магазина. Построение системы классификации объектов недвижимости. Построение и оптимизация системы кредитного скоринга.	Лабораторная работа, интерактивная форма.
Практическое применение моделей машинного обучения в бизнесе	Разработка модели машинного обучения для прогнозирования цен на акции на фондовом рынке. Прогнозирование поставок. Сегментирование потребителей финансовых услуг.	Лабораторная работа, интерактивная форма.



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 машинное обучение и искусственный интеллект	Проблема больших данных и низкой эффективности оптимизаторов. Оптимизационные задачи. Метод наименьших квадратов. Методы объемного усреднения.	Работа с учебной литературой. Решение задач. Разбор вопросов по теме занятия. Выполнение домашних заданий.
Алгоритмы и методы машинного обучения	Методы совершенствования моделей машинного обучения. Оптимизация гиперпараметров моделей. Обработка и предварительная подготовка данных. Создание синтетических признаков для обогащения набора данных. Работа с пропущенными данными. Предотвращение эффекта переобучения. Работа с несбалансированными классами. Мультисетевые динамические модели.	Работа с учебной литературой. Разбор вопросов по теме занятия. Выполнение домашних заданий.
Практическое применение моделей машинного обучения в бизнесе	Применение методов снижения размерности для улучшения качества моделей машинного обучения. Применение методов снижения размерности для выявления латентных факторов. Оценка запасов на основе библиотеки NumPy. Баланс спрос/предложение - методы предиктивной оценки.	Работа с учебной литературой. Решение задач. Разбор вопросов по теме занятия. Выполнение домашних заданий.



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

Примерная тематика проектной работы

1. Разработка модели машинного обучения для прогнозирования цен акций на фондовом рынке.
2. Разработка модели машинного обучения для классификации кредитных заемщиков на основе их кредитной истории и других факторов
3. Разработка модели машинного обучения для оптимизации ассортимента розничного магазина по данным чеков покупателей.
4. Разработка модели машинного обучения для выявления дефектов производства на конвейере.
5. Разработка модели машинного обучения для определения риска дефолта корпоративного заемщика.
6. Разработка модели машинного обучения для прогнозирования цены на объекты недвижимости на основе макроэкономических факторов
7. Разработка модели машинного обучения для прогнозирования объема продаж мобильных устройств.
8. Разработка модели машинного обучения для прогнозирования оттока клиентов банка
9. Разработка модели машинного обучения для прогнозирования вероятности отклика клиента на маркетинговое предложение.
10. Разработка модели машинного обучения для прогнозирования страховых выплат.
11. Разработка модели машинного обучения для оптимизации цен на товары и услуги на основе анализа конкурентной среды и потребительского спроса
12. Разработка модели машинного обучения для автоматической классификации клиентов по уровню лояльности и вероятности повторных покупок
13. Разработка модели машинного обучения для автоматической оптимизации инвестиционного портфеля на основе анализа рисков и доходности активов

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
УК-4. Способность использовать прикладное программное обеспечение при решении профессиональных задач	Использует основные методы и средства получения, представления, хранения и обработки данных	Знать: назначение и возможности прикладного ПО, используемого в задачах бизнес-аналитики и ML/AI (табличные процессоры, BI-инструменты, СУБД-клиенты, Python-экосистема для анализа данных и ML) Уметь: импортировать данные из типо-	Загрузите бизнес-датасет (продажи, кредиты или маркетинг), выполните предобработку (пропуски, выбросы и кодирование категорий), сформируйте витрину признаков и сохраните результат в удобном формате.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		вых источников (файлы/БД), выполнять очистку, трансформации и подготовку признаков в выбранном инструментарии	
	Демонстрирует владение профессиональными пакетами прикладных программ	Знать: типовые форматы данных и способы их получения, хранения и обработки, базовые принципы подготовки данных к анализу Уметь: применять прикладные программные средства для построения ML-решения: обучение модели, оценка качества, сравнение аль-	Постройте baseline-модель (классификация или регрессия), проведите оценку качества, визуализируйте ключевые метрики и сделайте интерпретацию результата для бизнес-сценария.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		тернатив, визуализация и интерпретация результатов для бизнес-пользователя	
	Выбирает необходимое прикладное программное обеспечение в зависимости от решаемой задачи	Знать: базовые принципы воспроизводимости вычислений и документирования результатов как требование к профессиональному использованию ПО при решении прикладных задач Уметь: обоснованно выбирать инструменты под задачу (например: BI для отчетности и витрин, Python для анализа и моделирования; ло-	Подготовьте управленческий артефакт: выгрузите прогноз в таблицу и соберите простую отчетную форму с ключевыми показателями и пояснениями.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		кальный запуск — при необходимости воспроизводимости)	
	Использует прикладное программное обеспечение для решения конкретных прикладных задач	Знать: конкретные возможности прикладного программного обеспечения (Python с библиотеками pandas/sklearn, табличные процессоры, BI-инструменты, СУБД-клиенты) для решения типовых прикладных задач в области экономики и финансов; типовые алгоритмы и последовательности действий в прикладном ПО для решения задач анализа дан-	Разработайте в Jupyter Notebook сквозной воспроизводимый пайплайн для задачи бизнес-прогнозирования оттока клиентов, включающий предобработку данных, обучение модели классификации (градиентный бустинг), оценку качества по различным метрикам, в том числе ROC-AUC и оформление отчёта с фиксацией версий библиотек и комментариев.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		ных, прогнозирования, оценки рисков и подготовки отчётности Уметь: выбирать и применять конкретные функции, методы и библиотеки прикладного ПО для решения поставленной прикладной задачи (очистка данных, построение модели, визуализация, формирование отчёта); настраивать параметры прикладного ПО и адаптировать его под особенности конкретной задачи (форматы данных, ограничения	



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		по времени, требования к воспроизводимости)	
ПКП-3. Способность проводить анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры	Подготавливает данные для проведения работ по исследованию больших данных: извлечение, проверка, очистка, агрегация и оценка	Знать: модели и методы прогнозной финансовой аналитики, основанные на анализе больших данных и применении моделей машинного обучения; способы оценки качества прогнозных моделей для финансовых временных рядов (MAE, RMAPE, VaR-тесты) Уметь: применять методы анализа больших данных и модели машинного обучения для эффектив-	Задание 1. Разработайте модель машинного обучения на Python для определения наилучших поставщиков материалов и компонентов на основе анализа исторических данных о качестве, ценах и сроках поставок. Задание 2. Разработайте модель прогнозирования доходности портфеля акций на основе больших данных (цены закрытия, макроэкономические индикаторы, новостной фон).



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		ной оценки финансовых инструментов и инвестиционного планирования; адаптировать модели машинного обучения под специфику финансовых данных (нестационарность, шум, выбросы)	
	Использует современные методы, стандарты, инструментальные средства и языки программирования для проведения анализа данных	Знать: аналитические инструменты и методики оценки и моделирования финансовых рисков инвестирования в различные проекты; методы стресс-тестирования и сценарного анализа инвестиционных проектов	Задание 1. Разработайте модель машинного обучения на Python для анализа данных о качестве продукции и выявления причин возникновения дефектов с целью оптимизации производственных процессов и улучшения качества продукции. Задание 2. Разработайте модель машинного обучения на Python для оценки кредитного риска заёмщика на основе данных о транзакциях, социально-демографических и финансовых характеристиках заёмщика (возраст, доход, стаж, наличие других кредитов) с расчётом ожидаемых потерь.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		Уметь: использовать современный аналитический инструментарий для объективной оценки и минимизации рисков вложения капитала в финансовые активы и инвестиционные проекты	
	Интерпретирует и визуализирует большие данные, полученные в результате анализа в соответствии с целями исследования	Знать: специфику интерпретации результатов анализа больших данных для разных аудиторий: перевод технических метрик (точность, Recall) на язык бизнеса (прибыль, снижение рисков) для менеджмента и регу-	Задание 1. Подготовьте интерактивный дашборд с визуализацией прогнозных значений, доверительных интервалов и ключевых факторов влияния в динамике по разным регионам или категориям товаров на основе построенной модели прогнозирования спроса. Задание 2. Подготовьте аналитический отчет с визуализацией распределения вероятностей дефолта по сегментам клиентов и содержащий интерпретацию основных драйверов риска (например, какие именно транзакционные или макроэкономические параметры сильнее всего влияют на вероятность невозврата кредита) на основе результатов анализа кредитного риска заёмщиков.



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		ляторов; методы и инструменты визуализации, эффективные для больших объёмов данных: интерактивные дашборды для исследования срезов, тепловые карты для выявления аномалий, визуализация сложных моделей (например, деревьев решений) для объяснения логики Уметь: интерпретировать результаты анализа больших данных для оценки их влияния на ключевые финансовые показатели и риски ком-	



Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
		паний; создавать наглядные и интерактивные визуализации, отвечающие на конкретные вопросы бизнеса и соответствующие целям исследования	



Примеры практико-ориентированных заданий

Задание 1. Кредитный скоринг: «Одобрить или отказать? Анализ факторов дефолта заемщика»

Тип: Расчетно-аналитическая работа / Лабораторная работа.

Бизнес-контекст: Вы — финансовый аналитик в банке. Вам нужно оценить, насколько модель машинного обучения (логистическая регрессия) предсказывает дефолт заемщиков. Ваша задача — не обучить модель с нуля, а объяснить руководству, какие факторы влияют на дефолт, насколько точна модель и где она ошибается.

Исходные данные: датасет из 1000 клиентов банка (возраст, доход, кредитная история, сумма кредита, цель кредита, факт дефолта). Данные "зашумлены": есть пропуски и выбросы.

Задачи для студента:

Предобработка (аналитика): используя Python (Jupyter Notebook с готовыми ячейками), определить:

- Сколько пропусков в колонке "доход"? Предложить способ их заполнения (медиана или среднее) и обосновать выбор.

- Найти аномально высокий доход (>10 млн руб.) — выброс. Объяснить, почему его лучше удалить или заменить предельным значением.

Запуск модели (код может быть дан или его готовит студент): запускается скрипт, который строит логистическую регрессию и выводит:

- Коэффициенты модели (веса признаков).
- Матрицу ошибок (True/False Positive/Negative).
- ROC-кривую и значение AUC.

Интерпретация (ключевая часть): посмотреть на коэффициент при признаке "просрочка по прошлым кредитам". Он положительный. Что это означает? Ответ: Чем больше просрочек в прошлом, тем выше вероятность дефакта.

Модель выдала 20 ложных срабатываний (ошибочно отказала хорошим клиентам) и 5 ложных пропусков (пропустила плохих). Что опаснее для банка в текущей экономической ситуации? Потери процентов или прямые убытки от невозврата? Обосновать.

Результат (управленческое решение): студент пишет вывод на 1 страницу: "Рекомендую банку ужесточить требования к заемщикам с признаком X, так как он дает наибольший вес в модели. Модель имеет $AUC = 0.82$, что означает хорошее качество. Однако из-за высокой стоимости ошибки первого рода (отказ хорошему клиенту) предлагаю понизить порог принятия решения с 0.5 до 0.4".



Задание 2. Прогнозирование финансового временного ряда: «Спрогнозируй цену акции на завтра»

Тип: Командный проект / Кейс.

Бизнес-контекст: Вы — трейдер или риск-менеджер. Вам нужно оценить, можно ли предсказать цену закрытия акций Сбера (или любой другой компании) на основе исторических данных. Вам предоставлены готовые графики и метрики трех моделей: скользящее среднее (MA), линейная регрессия и нейронная сеть (LSTM — упрощенно).

Исходные данные: CSV-файл с историей цен акций за 3 года (Date, Open, High, Low, Close, Volume).

Задачи для студента:

- визуальный анализ: Построить график цены закрытия за весь период.
- визуально определить: есть ли тренд (растущий/падающий), есть ли сезонность (январский эффект и т.д.), есть ли резкие скачки (шоки).

Оценка моделей (без/с написанием кода): запускается скрипт, который обучает три модели и выводит метрики: MAE (средняя абсолютная ошибка) и RMSE (корень из среднеквадратичной ошибки).

Модель 1 (MA): MAE = 15 руб.

Модель 2 (LR): MAE = 12 руб.

Модель 3 (NN): MAE = 14 руб.

Вопрос: какая модель точнее предсказывает цену? (Линейная регрессия).

Вопрос: почему сложная нейронная сеть проиграла линейной регрессии? Ответ: данных мало, рынок не обладает глубокой нелинейной структурой, модель переобучилась.

Финансовый вывод: студент рассчитывает, сколько денег (в %) он бы потерял или заработал, если бы использовал предсказания лучшей модели для покупки акций. Если ошибка модели в 12 рублей составляет 2% от цены акции, то это приемлемо или нет? Сравнить с безрисковой ставкой (ключевая ставка ЦБ).

Результат: аналитический отчет "Гибридный подход: использовать линейную регрессию для тренда, но НЕ использовать для краткосрочных сделок из-за высокого шума".

Задание 3. Сегментация клиентов и управление рисками: «Кому предложить ипотеку со страховкой?»

Тип: Деловая игра / Лабораторная работа.

Бизнес-контекст: банк хочет предложить клиентам, берущим ипотеку, дополнительную финансовую услугу — страхование жизни и здоровья. Но предлагать



всем дорого и неэффективно. Нужно с помощью кластеризации (k-means, без учителя) выделить 3 группы риска и предложить разные условия.

Исходные данные: данные о 500 ипотечных заемщиках (возраст, первоначальный взнос, срок кредита, ежемесячный платеж, профессия (кодированная), семейное положение).

Задачи для студента: запуск кластеризации - готовится код, который делит клиентов на 3 кластера и визуализирует их на графике (например, "возраст" vs "первоначальный взнос").

Интерпретация кластеров (бизнес-логика):

- Кластер 1 (красный): Молодые (до 30 лет), низкий взнос, но высокий доход.
- Кластер 2 (синий): Средний возраст, средний взнос, работа по найму.
- Кластер 3 (зеленый): Возраст 45+, высокий взнос, но нестабильный доход (фриланс).

Выработка стратегии (финансовое решение): для каждого кластера студент предлагает процентную ставку по страховке (от 0.5% до 3% от суммы кредита).

Обосновать: *Кластеру 3 (нестабильный доход) — высокий процент или отказ. Кластеру 1 — низкий процент, но с требованием предоставить справку 2-НДФЛ. Кластеру 2 — базовый продукт.*

Расчет экономического эффекта: студент поверхностно анализирует: если обучить модель на исторических данных, сколько денег банк сэкономит на страховых выплатах, отсекая самых рискованных? (5-10% от портфеля).

Примеры вопросов для подготовки к промежуточной аттестации

Примерные вопросы для подготовки к экзамену

1. Типы машинного обучения: обучение с учителем, обучение без учителя
2. Интеллектуальный анализ данных, большие данные и машинное обучение: основные понятия и методы
3. Обучение с подкреплением
4. Методы и задачи машинного обучения
5. Большие данные и их основные характеристики (6V)
6. Основные этапы цикла обработки данных
7. Алгоритмы машинного обучения: линейная регрессия, деревья решений, нейронные сети
8. Понятие ансамбля в машинном обучении и принципы его построения
9. Метод опорных векторов
10. Случайные леса в машинном обучении
11. Модель множественной линейной регрессии
12. Модель логистической регрессии



13. Критерии оценки качества моделей классификации
14. Стекинг и блендинг в машинном обучении
15. Мультиномиальный наивный байесовский классификатор
16. Метод k-ближайших соседей
17. Структура сверточной нейронной сети, виды архитектур
18. Методы оптимизации моделей машинного обучения: градиентный спуск и стохастический градиентный спуск
19. Наивный байесовский классификатор
20. Кластерный анализ. Методы оценки качества моделей кластерного анализа
21. Методы снижения размерности и их использование
22. Характеристика задач классификации и регрессии
23. Обучение и тестирование моделей машинного обучения
24. Классификация моделей машинного обучения
25. Кросс-валидация в машинном обучении

** Промежуточная аттестация обучающихся проводится в формате тестирования*



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

Основная литература:

1. Андриянов, Н.А. Прикладные задачи машинного обучения и обработки больших данных : учебное пособие по дисциплине: "Прикладные задачи машинного обучения" для студентов, обучающихся по направлению: Бизнес-информатика, направленность программы "Управление информационными технологиями в цифровой экономике" (программы подготовки магистров) / Н.А. Андриянов ; Финуниверситет, Департамент анализа данных и машинного обучения факультета информационных технологий и анализа больших данных Москва : Финуниверситет, 2023 1 файл (4,55 Мб) : ил. Свободный доступ из сети Интернет (чтение, печать) <http://elib.fa.ru/fbook/books137580.pdf> (дата обращения 22.11.2023) + Текст : электронный. [БИК ID: RU\FA\books\137580]
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. Сидоркина, И.Г. Системы искусственного интеллекта : Учебное пособие / И.Г. Сидоркина Электрон. дан. Москва : КноРус, 2026 245 с. Режим доступа: book.ru Internet access <https://book.ru/book/959436> ISBN 978-5-406-14910-2. [БИК ID: RU\bookru\bibl\959436]

Дополнительная литература:

1. Каширин, В.В. Интеллектуальная экономика России третьего тысячелетия: модернизация инновации инвестиции цифровые технологии искусственный интеллект : Монография / В.В. Каширин Электрон. дан. Москва : Русайнс, 2020 149 с. Режим доступа: book.ru Internet access <https://book.ru/book/939545> ISBN 978-5-4365-5944-5. [БИК ID: RU\bookru\bibl\939545]



2. Сидоркина, И.Г. Системы искусственного интеллекта : Учебное пособие / И.Г. Сидоркина Электрон. дан. Москва : КноРус, 2024 245 с. Режим доступа: book.ru Internet access <https://book.ru/book/950668> ISBN 978-5-406-12171-9. [БИК ID: RU\bookru\bibl\950668]
3. Целых, А. Н. Применение временных рядов для анализа больших данных: учебное пособие по курсу «Математические методы анализа больших данных» : учебное пособие / А. Н. Целых, В. С. Васильев, Э. М. Котов ; Южный федеральный университет Ростов-на-Дону, Таганрог : Южный федеральный университет, 2021 86 с. : ил. Библиогр. в кн Режим доступа: электронная библиотечная система «Университетская библиотека ONLINE», требуется авторизация <https://biblioclub.ru/index.php?page=book&id=691448> ISBN 978-5-9275-3983-3. [БИК ID: BIBLIOCLUB\0000691448]

Нормативные правовые акты:

-

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

1. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/> (<http://library.fa.ru/files/elibfa.pdf>)
2. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
3. Электронно-библиотечная система "Университетская библиотека ОНЛАЙН" <http://biblioclub.ru/>
4. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>
5. Образовательная платформа "ЮРАЙТ" <https://urait.ru/>
6. Информационно-образовательный портал Финуниверситета: <https://org.fa.ru>



10. Методические указания для обучающихся по освоению дисциплины

Самостоятельная работа студентов проходит аудиторно и внеаудиторно. Организации самостоятельной работы служит учебно-тематический план изучения дисциплины. В этом плане указана тематика лекций, практических занятий, вопросы и задания для самостоятельного изучения.

При подготовке к лекции целесообразно предварительно познакомиться с ее содержанием по рекомендованным пособиям и выделить наиболее трудные вопросы. Во время лекций следует конспектировать содержание лекции. После занятий следует провести работу с конспектом: отредактировать записи, оформить конспект.

При подготовке к практическому занятию необходимо повторить или, если это требуется, изучить соответствующий теоретический материал. Практические занятия проходят, как правило, в интерактивной форме и преподаватель учитывает активность студентов, направленную на решение предложенных задач и в поиске ответов на вопросы.

Домашние задания следует выполнять регулярно при подготовке к практическим занятиям. Контроль за выполнением домашних заданий осуществляется в ходе практических занятий и выборочного собеседования.

Рекомендации для подготовки проектной работы

1. Выбор бизнес-задачи, а не алгоритма

Не начинайте с выбора «нейронной сети» или «случайного леса». Сначала сформулируйте финансовую или управленческую проблему: «Как снизить отток клиентов?», «Как точнее прогнозировать денежный поток?», «Какие факторы влияют на дефолт контрагента?». Алгоритм — это инструмент, который подбирается под задачу, а не наоборот.

2. Данные важнее модели

Качество результата на 80% определяется данными, и только на 20% — алгоритмом. Уделите основное время сбору, очистке и первичному анализу данных. Пропуски, выбросы, некорректные типы данных — самые частые причины провала проектов. Помните правило: «мусор на входе — мусор на выходе».

3. Простота — залог интерпретируемости

Для бизнеса и финансов предсказуемая, понятная модель лучше, чем сложный «черный ящик». Логистическая регрессия или дерево решений, которое можно объяснить за 2 минуты, часто ценнее, чем нейронная сеть с 10 слоями, которую никто не может проинтерпретировать. Начинайте с простых моделей и усложняйте только при явном улучшении качества.

4. Метрики должны говорить на языке бизнеса



Выбирайте и используйте ключевые метрики. В финансах например главные — это:

Для регрессии (прогнозы в рублях): MAE (средняя ошибка в рублях) — скажите «модель ошибается в среднем на 5000 руб.».

Для классификации (да/нет): Precision (точность — из всех предсказанных «дефолтов» сколько реальных) и Recall (полнота — из всех реальных «дефолтов» сколько нашли).

Выбор метрики — это выбор приоритета: ошибка первого рода (ложная тревога) или второго рода (пропуск цели). В финансах цена ошибок разная, и это определяет настройки модели.

5. Вывод — это управленческое решение

Результат работы — не график и не таблица с цифрами. Результат — это ответ на вопрос: «Что делать?». Ваш вывод должен содержать конкретные рекомендации: повысить лимит, отклонить заявку, сегментировать клиентов на 3 группы, изменить цену на товар. Если после прочтения вашего отчета непонятно, какое действие совершить — работа не завершена.

6. Структура отчета (типовая):

Бизнес-задача (зачем?)

Данные (откуда, какие, сколько)

Модель и ее настройки (что делали)

Качество модели (метрики на языке бизнеса)

Интерпретация (какие факторы важны)

Управленческое решение (что делать?)

7. Допустимые инструменты

Jupyter Notebooks (с подстановкой своих данных).

Библиотеки scikit-learn, pandas, matplotlib — в режиме вызова готовых функций.

Визуальные среды (при необходимости).

Результат должен быть осмыслен и защищен студентом устно.

8. Защита проекта (3 слайда и 5 минут)

На защите студент должен показать:

Слайд 1: Бизнес-задача и данные (коротко).

Слайд 2: Самый важный график или таблица с факторами (что влияет?).

Слайд 3: Конкретные рекомендации (что делаем с этим знанием?).

Устно: ответить на вопрос «Почему я выбрал именно эту метрику качества?» и «Почему модель можно или нельзя внедрять прямо сейчас?».



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. Microsoft Office (Windows)
4. Пакет офисных программ
5. библиотеки Keras
6. Антивирус Kaspersky

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. Язык программирования Python 3. <https://pythonworld.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)