

**Федеральное государственное образовательное бюджетное
учреждение высшего образования
«Финансовый университет при Правительстве Российской Федерации»
(Финуниверситет)**

Владикавказский филиал Финуниверситета

Кафедра «Корпоративные инфокоммуникационные системы»



С.Б. Волошин

Машинное обучение

Рабочая программа дисциплины
для студентов, обучающихся по направлению подготовки
09.03.04 Программная инженерия,
ОП «Технологии разработки программного обеспечения»

*Рекомендовано Ученым советом Владикавказского филиала
Финансового университета
(протокол от «15» сентября 2026 г. № 30)*

*Одобрено на заседании кафедры «Корпоративные инфокоммуникационные
системы»
(протокол от «10» сентября 2026 г. № 8)*

Владикавказ 2026

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
4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	4
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий.....	4
5.1. Содержание дисциплины	4
5.2. Учебно-тематический план	7
5.3. Содержание семинаров, практических занятий.....	8
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине.....	10
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы	10
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю	12
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине.....	25
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	34
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	34
10. Методические указания для обучающихся по освоению дисциплины.....	35
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем	45
11.1 Комплект лицензионного программного обеспечения	45
11.2 Современные профессиональные базы данных, и информационные справочные системы	46
11.3 Сертифицированные программные и аппаратные средства защиты информации	46
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.....	46

1. Наименование дисциплины

Дисциплина «Машинное обучение».

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
ОПК-1	Способен применять естественнонаучные и общетеоретические знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности	1. Демонстрирует знания основных методов математического анализа и моделирования, применяет их на практике для решения задач профессиональной деятельности	Знать: основы математики, линейной алгебры и статистики; основы программирования, особенно на Python; основные алгоритмы машинного обучения. Уметь: разрабатывать и обучать модели машинного обучения; обрабатывать и анализировать данные; оценивать и настраивать модели; визуализировать данные и результаты; эффективно коммуницировать результаты анализа данных.
		2. Проводит теоретические исследования по выбранной области профессиональной деятельности.	Знать: основы машинного обучения и математики, связанной с этой областью. Уметь: создавать и настраивать модели машинного обучения; анализировать данные и применять их в моделях; критически оценивать методы машинного обучения; эффективно коммуницировать результаты теоретических исследований.
		3. Проводит численные эксперименты на основе математических или информационных методов, делает выводы и обосновывает их.	Знать: основы машинного обучения и статистики, методы и модели машинного обучения. Уметь: планировать и проводить численные эксперименты; анализировать данные и извлекать выводы из численных экспериментов; обосновывать полученные результаты на основе математических и статистических методов; эффективно коммуницировать результаты численных экспериментов.

3. Место дисциплины в структуре образовательной программы

Дисциплина «Машинное обучение» является дисциплиной Общепрофессионального цикла обязательной части учебного плана образовательной программы «Технологии разработки программного обеспечения» по направлению подготовки 09.03.04 Программная инженерия, профиль «Технологии разработки программного обеспечения».

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

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 3 (в часах)	Семестр 4 (в часах)
Общая трудоемкость дисциплины	8/288	144	144
Контактная работа- Аудиторные занятия	100	50	50
<i>Лекции</i>	32	16	16
<i>Семинары, практические занятия</i>	68	34	34
Самостоятельная работа	188	94	94
Вид текущего контроля	Контрольная работа, проектная работа	Контрольная работа	Проектная работа
Вид промежуточной аттестации	Зачет, экзамен	Зачет	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

Основные понятия машинного обучения. Связь с другими дисциплинами. Контекст машинного обучения как дисциплины - анализ данных и искусственный интеллект. Сферы применения машинного обучения. Типы задач машинного обучения - обучение с учителем и без

учителя. Структура данных для машинного обучения. Инструментальные средства машинного обучения. Понятие модели машинного обучения.

Тема 2. Регрессия

Постановка задачи регрессии. Линейная регрессия с одной переменной - функция гипотезы, функция ошибки, метод градиентного спуска. Регрессия с несколькими переменными - множественная линейная регрессия, нормализация признаков, полиномиальная регрессия. Практическое построение регрессии - загрузка и представление данных, реализация метода градиентного спуска, оценка качества регрессии, подбор скорости обучения, знакомство с библиотекой `sklearn`.

Тема 3. Классификация

Постановка задачи классификации как задачи машинного обучения. Отличия от задачи регрессии. Структура данных для классификации. Логистическая регрессия - функция гипотезы, граница принятия решений, функция ошибки логистической регрессии, градиентный спуск для логистической регрессии, многоклассовая классификация, алгоритм “один против всех”.

Тема 4. Методы обучения с учителем

Универсальность методов обучения с учителем, общая постановка задачи. Линейные модели - линейная и логистическая регрессии - как единая модель. Полиномиальные модели. Метод опорных векторов, ядра - линейное, гауссово, другие. Перцептрон. Деревья решений. К ближайшим соседям. Наивная байесовская модель. Достоинства и недостатки разных типов моделей, их сравнительная характеристика, применимость. Применение этих моделей для решения задач классификации и регрессии.

Тема 5. Диагностика систем машинного обучения

Метрики эффективности машинного обучения - сравнение с функциями ошибки. Типичные метрики эффективности для моделей регрессии - MAE, MSE, RMSE, MSLE, MAPE и другие. Метрики эффективности для моделей классификации - accuracy, precision, recall, F1,

ROC, PR и другие. Недообучение и переобучение. Проблема bias-variance. Оценка сложности моделей. Обобщающая способность моделей, тестовый набор, кривые обучения. Методы борьбы с недо- и переобучением. Регуляризация. Задача выбора модели - кросс-валидация, гиперпараметры моделей, поиск по сетке, валидационный набор.

Тема 6. Предварительный анализ и обработка данных

Сбор данных для моделей обучения с учителем - реляционная форма данных, понятие чистых данных, оценка источников и объемов данных. Описательный (предварительный) анализ данных (EDA) - анализ репрезентативности, шкалы и типы, визуализация, проблема несбалансированности, обнаружение корреляций, аномалий в данных. Очистка и преобразование данных - удаление лишних признаков, удаление непоказательных объектов, заполнение отсутствующих значений, создание суррогатных признаков, преобразование шкал, воспроизводимость преобразования данных.

Тема 7. Задачи обучения без учителя

Задача кластеризации - постановка задачи, структура датасета, результат и интерпретация. Метод K средних - формализация, гиперпараметры, применимость. Другие методы кластеризации - DBSCAN, иерархическая, агломеративная кластеризация. Задача обнаружения аномалий. Задача понижения размерности - метод главных компонент, метод независимых компонент. Обучение с подкреплением.

Тема 8. Практическое использование моделей машинного обучения

Стохастический и пакетный градиентный спуск. Отбор признаков. Частичное обучение с учителем. Ансамблирование моделей - беггинг, бустинг, стакинг, случайный лес, XGBoost, CatBoost. Конвейеризация моделей машинного обучения. Основные этапы проекта по машинному обучению. Визуализация, интерпретация, представление и анализ результатов машинного обучения. Работа с разными типами данных - преобразование графической информации, методы векторизации текста.

5.2. Учебно-тематический план

№ п/ п	Наименование тем (разделов) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Всего	Контактная работа - Аудиторная работа			Самостоятельная работа	
			Общая, в т.ч.:	Лекции и	Семинары, практи		
1	Тема 1. Введение в машинное обучение	36	6	2	4	30	Опрос, решение задач, проверка выполнения домашних заданий самостоятельной работы
2	Тема 2. Регрессия	36	10	2	8	26	Опрос, решение задач, проверка выполнения домашних заданий самостоятельной работы
3	Тема 3. Классификация	36	6	2	4	30	Опрос, решение задач, проверка выполнения домашних заданий самостоятельной работы
4	Тема 4. Методы обучения с учителем	36	16	6	10	20	Опрос, решение задач, проверка выполнения домашних заданий самостоятельной работы
5	Тема 5. Диагностика систем машинного обучения	36	20	8	12	16	Опрос, решение задач, проверка выполнения домашних заданий самостоятельной работы, защита контрольной работы
6	Тема 6. Предварительный анализ и обработка данных	36	18	6	12	18	Опрос, решение задач, проверка выполнения домашних заданий самостоятельной работы
7	Тема 7. Задачи обучения без учителя	36	12	4	8	24	Опрос, решение задач, проверка выполнения домашних заданий самостоятельной работы
8	Тема 8. Практическое использование моделей машинного обучения	36	12	2	10	24	Опрос, решение задач, проверка выполнения домашних заданий самостоятельной работы, защита проектной работы

В целом по дисциплине	288	100	32	68	188	Согласно учебному плану: контрольная работа, проектная работа
Итого в %	100	34,7	32	68	65,3	

5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарах, практических занятиях	Формы проведения занятия
Тема 1. Введение в машинное обучение	Входной контроль. Изучение технологического стека анализа данных, построенного на базе языка программирования Python. Знакомство с библиотеками numpy, pandas, matplotlib. Выполнение задания по простому статистическому анализу данных инструментальными средствами, включающему продвинутое владение соответствующими инструментами.	Интерактивная форма: опрос, собеседование по домашним заданиям самостоятельной работы, решение задач с последующим коллективным обсуждением их результатов
Тема 2. Регрессия	Построение модели регрессии методом градиентного спуска своими руками. Знакомство с основами обучения параметров модели. Построение модели множественной регрессии своими руками. Построение модели библиотечными средствами. Сравнение результатов. Реализация модели регрессии на примере, близком к реальному (датасет boston). Интерпретация результатов моделирования.	Интерактивная форма: опрос, собеседование по домашним заданиям самостоятельной работы, решение задач с последующим коллективным обсуждением их результатов
Тема 3. Классификация	Реализация модели классификации (логистической регрессии) своими руками наподобие линейной регрессии. Использование библиотечных функций.	Интерактивная форма: опрос, собеседование по домашним заданиям самостоятельной работы, решение задач с последующим коллективным обсуждением их результатов
Тема 4. Методы обучения с учителем	Построение модели классификации на простых данных. Построение нескольких моделей, сравнение их эффективности. Построение модели на сложных данных с нелинейными и слабыми зависимостями. Оценка эффективности, выбор модели. Построение модели регрессии на сложных	Интерактивная форма: опрос, собеседование по домашним заданиям самостоятельной работы, решение задач с последующим

	данных. Построение разных типов регрессоров. Оценка эффективности, выбор модели.	коллективным обсуждением их результатов
Тема 5. Диагностика систем машинного обучения	Оценка эффективности модели регрессии с помощью разных метрик. Оценка эффективности классификации с помощью разных метрик. Построение диагностических кривых (PR, ROC) для бинарной классификации. Диагностика недо-и переобучения в модели классификации. Построение кривых обучения, их интерпретация. Тестовый набор данных Демонстрация необходимости кросс-валидации. Использование кросс-валидированных оценок качества модели. Демонстрация оптимизации гиперпараметров модели, использование валидационного набора для выбора модели. Поиск модели от простых к сложным.	Интерактивная форма: опрос, собеседование по домашним заданиям самостоятельной работы, решение задач с последующим коллективным обсуждением их результатов, выполнение контрольной работы
Тема 6. Предварительный анализ и обработка данных	Сбор и интеграция данных из разных источников. Основные численные характеристики датасета, анализ на чистоту. Нормализация признаков и заполнение отсутствующих значений на искусственном примере. Преобразование категориальных признаков в численные. Краткий алгоритм анализа данных на реальном примере. Интерпретация результатов. Отбор и инжиниринг признаков на примере, приближенном к реальному (датасет titanic).	Интерактивная форма: опрос, собеседование по домашним заданиям самостоятельной работы, решение задач с последующим коллективным обсуждением их результатов
Тема 7. Задачи обучения без учителя	Решение задачи кластеризации на искусственных данных. Метод К-средних. Метод локтя. Обнаружение аномалий на искусственных данных. Обнаружение аномалий как задача предварительного анализа данных. Понижение размерности на искусственных данных. Интерпретация результата. Понижение размерности для визуализации данных.	Интерактивная форма: опрос, собеседование по домашним заданиям самостоятельной работы, решение задач с последующим коллективным обсуждением их результатов
Тема 8. Практическое использование моделей машинного обучения	Использование простых ансамблей моделей. Оптимизация гиперпараметров моделей. Визуализация, интерпретация и представление результатов машинного обучения исходя из бизнес-задачи моделирования. Построение конвейера машинного обучения. Построение модели классификации текстов. Знакомство с	Интерактивная форма: опрос, собеседование по домашним заданиям самостоятельной работы, решение задач с последующим коллективным

	разными методами векторизации текстов. Построение модели классификации изображений. Представление изображений в виде вектора.	обсуждением их результатов, выполнение проектной работы
--	---	---

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
Тема 1. Введение в машинное обучение	Изучение истории машинного обучения как дисциплины, связи с другими областями знания. Изучение базовых библиотек анализа данных на языке программирования Python - numpy, pandas, matplotlib по официальной документации (при необходимости).	Изучение материалов лекций и литературы, предложенной преподавателем, поиск и анализ информации, содержащейся в Интернет-ресурсах. Разбор вопросов, отводимых на самостоятельное освоение, выполнение домашних заданий самостоятельной работы
Тема 2. Регрессия	Знакомство с нормированием данных, адаптивной скоростью обучения в ручной реализации метода градиентного спуска. Векторизация вычислений в модели множественной регрессии. Использование множественной регрессии для моделирования полиномиальных признаков и нелинейных зависимостей (корень, экспонента, синус, логарифм, обратная пропорциональность). Визуализация исходных данных и результатов моделирования в реальном случае. Анализ и представление результатов моделирования и совершение предсказания. Статистический анализ достоверности предсказания.	Изучение материалов лекций и литературы, предложенной преподавателем, поиск и анализ информации, содержащейся в Интернет-ресурсах. Разбор вопросов, отводимых на самостоятельное освоение, выполнение домашних заданий самостоятельной работы
Тема 3. Классификация	Реализация алгоритма “один против всех”. Оценка эффективности классификации. Сравнение двух реализаций.	Изучение материалов лекций и литературы, предложенной преподавателем, поиск и анализ информации, содержащейся в Интернет-ресурсах.

		Разбор вопросов, отводимых на самостоятельное освоение, выполнение домашних заданий самостоятельной работы
Тема 4. Методы обучения с учителем	Визуализация результатов классификации. Построение и анализ отчета от классификации. Оценка скорости обучения и работы разных моделей машинного обучения на большом объеме данных. Визуализация результатов классификации, интерпретация полученных результатов. Совершение предсказания и оценка его надежности.	Изучение материалов лекций и литературы, предложенной преподавателем, поиск и анализ информации, содержащейся в Интернет-ресурсах. Разбор вопросов, отводимых на самостоятельное освоение, выполнение домашних заданий самостоятельной работы
Тема 5. Диагностика систем машинного обучения	Автоматизация построения модели и подсчета выбранных метрик. Построение сводного отчета. Обоснование выбора метрик эффективности исходя из поставленной задачи. Визуализация диагностических кривых для множественной классификации. Использование регуляризованных моделей - ridge, lasso, elastic net. Демонстрация позитивного влияния регуляризации на точность модели. Построение и робастная оценка эффективности модели на данных, близких к реальным. Использование кросс-валидации. Построение модели и ее последующая оптимизация. Построение ряда простых моделей, диагностика и выбор класса модели, оптимизация и оценка качества.	Изучение материалов лекций и литературы, предложенной преподавателем, поиск и анализ информации, содержащейся в Интернет-ресурсах. Разбор вопросов, отводимых на самостоятельное освоение, выполнение домашних заданий самостоятельной работы
Тема 6. Предварительный анализ и обработка данных	Анализ распределения каждого атрибута и связь с целевой переменной. Визуализация. Оценка влияния нормализации признаков на качество обучения. Заполнение отсутствующих значений на реальном примере. Группировка данных и продвинутое алгоритмы решкалирования. Полный алгоритм анализа данных на реальном примере. Анализ влияния на	Изучение материалов лекций и литературы, предложенной преподавателем, поиск и анализ информации, содержащейся в Интернет-ресурсах. Разбор вопросов, отводимых на самостоятельное освоение, выполнение

	моделирование. Алгоритмический отбор признаков, отбор по итогам оценки важности признаков по итогам построения простых моделей.	домашних заданий самостоятельной работы
Тема 7. Задачи обучения без учителя	Решение задачи кластеризации на реальных данных. Интерпретация результата. Обнаружение аномалий в реальных данных. Интерпретация результатов. Понижение размерности для задачи снижения объема данных. Понижение размерности как инжиниринг признаков.	Изучение материалов лекций и литературы, предложенной преподавателем, поиск и анализ информации, содержащейся в Интернет-ресурсах. Разбор вопросов, отводимых на самостоятельное освоение, выполнение домашних заданий самостоятельной работы, выполнение
Тема 8. Практическое использование моделей машинного обучения	Сложные ансамбли моделей - XGBoost, CatBoost. Оптимизация гиперпараметров, выбор модели. Использование различных типов графики, инфографики для оптимального представления результатов моделирования. Знакомство с разными задачами обработки текста - рубрикация, перевод, анализ тональности, генерация. Знакомство с разными задачами анализа изображений - идентификация объекта, обнаружение объекта, распознавание объекта.	Изучение материалов лекций и литературы, предложенной преподавателем, поиск и анализ информации, содержащейся в Интернет-ресурсах. Разбор вопросов, отводимых на самостоятельное освоение, выполнение домашних заданий самостоятельной работы

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

Примерный перечень тем для подготовки к опросу

1. Входной контроль. Изучение технологического стека анализа данных, построенного на базе языка программирования Python.
2. Выполнение задания по простому статистическому анализу данных инструментальными средствами, включающему продвинутое владение соответствующими инструментами.
3. Построение модели регрессии методом градиентного спуска своими руками. Знакомство с основами обучения параметров модели.

4. Построение модели множественной регрессии своими руками. Построение модели библиотечными средствами. Сравнение результатов.
5. Реализация модели регрессии на примере, близком к реальному (датасет boston). Интерпретация результатов моделирования.
6. Реализация модели классификации (логистической регрессии) своими руками наподобие линейной регрессии.
7. Использование библиотечных функций.
8. Построение модели классификации на простых данных. Построение нескольких моделей, сравнение их эффективности.
9. Построение модели на сложных данных с нелинейными и слабыми зависимостями. Оценка эффективности, выбор модели.
10. Построение модели регрессии на сложных данных. Построение разных типов регрессоров. Оценка эффективности, выбор модели.
11. Оценка эффективности модели регрессии с помощью разных метрик.
12. Оценка эффективности классификации с помощью разных метрик.
13. Построение диагностических кривых (PR, ROC) для бинарной классификации.
14. Диагностика недо-и переобучения в модели классификации. Построение кривых обучения, их интерпретация. Тестовый набор данных
15. Демонстрация необходимости кросс-валидации. Использование кросс-валидированных оценок качества модели.
16. Демонстрация оптимизации гиперпараметров модели, использование валидационного набора для выбора модели. Поиск модели от простых к сложным.
17. Сбор и интеграция данных из разных источников.
18. Основные численные характеристики датасета, анализ на чистоту.
19. Нормализация признаков и заполнение отсутствующих значений

на искусственном примере.

20. Преобразование категориальных признаков в численные.

21. Краткий алгоритм анализа данных на реальном примере.

Интерпретация результатов.

22. Отбор и инжиниринг признаков на примере, приближенном к реальному (датасет titanic).

23. Решение задачи кластеризации на искусственных данных.

24. Метод К-средних. Метод локтя.

25. Обнаружение аномалий на искусственных данных.

26. Обнаружение аномалий как задача предварительного анализа данных.

27. Понижение размерности на искусственных данных.

Интерпретация результата. Понижение размерности для визуализации данных.

28. Использование простых ансамблей моделей.

29. Оптимизация гиперпараметров моделей.

30. Визуализация, интерпретация и представление результатов машинного обучения исходя из бизнес-задачи моделирования.

31. Построение конвейера машинного обучения.

32. Построение модели классификации текстов.

33. Знакомство с разными методами векторизации текстов.

34. Построение модели классификации изображений.

35. Представление изображений в виде вектора.

Примеры задач

Задача 1. Проанализировать набор данных о стоимости недвижимости и подготовить его к моделированию.

1. Загрузите датасет о недвижимости (например, Boston Housing или собственный с признаками: площадь, количество комнат, район, этаж, расстояние до центра и т.д.).

2. Изучите структуру данных: определите типы признаков (числовые, категориальные), проверьте наличие пропущенных значений.
3. Определите тип задачи машинного обучения (регрессия/классификация) и обоснуйте выбор.
4. Визуализируйте распределение целевой переменной (цены) и ключевых признаков с помощью гистограмм и boxplot.
5. Подготовьте краткий отчёт с описанием датасета, типа задачи, потенциальных проблем данных и способов их решения.

Задача 2. Построить модель линейной регрессии для прогнозирования цен на недвижимость.

1. Проведите предварительную обработку данных:
 - заполните пропущенные значения (средним, медианой или с помощью интерполяции);
 - закодируйте категориальные признаки;
2. Разделите данные на обучающую и тестовую выборки (80 % / 20 %).
3. Реализуйте простую линейную регрессию для одного признака (например, площади) с использованием градиентного спуска:
 - задайте функцию гипотезы $h_{\theta}(x)=\theta_0+\theta_1x$;
 - реализуйте функцию стоимости $J(\theta)$;
 - напишите код для обновления параметров θ_0 , θ_1 на каждой итерации.
4. Обучите модель, выберите оптимальную скорость обучения (α) и количество итераций.
5. Оцените модель на тестовой выборке, рассчитав MAE, MSE, RMSE и R^2 .
6. Визуализируйте результаты: постройте график с реальными и предсказанными ценами, а также график сходимости функции стоимости.

Задача 3. Преобразовать задачу регрессии в задачу классификации (например, определить, является ли объект дорогим/дешёвым) и построить классификационную модель.

1. Преобразуйте целевую переменную (цену) в бинарную метку:
 - объекты с ценой выше медианы — класс 1 («дорогой»);
 - остальные — класс 0 («дешёвый»).
2. Повторите шаги предварительной обработки из предыдущих задач.
3. Постройте модель логистической регрессии:
 - реализуйте сигмоидальную функцию $g(z)=1/(1+e^{-z})$;
 - задайте функцию гипотезы $h_{\theta}(x)=g(\theta^T x)$;
 - обучите модель с помощью градиентного спуска.
4. Оцените качество модели, рассчитав accuracy, precision, recall, F1-score и построив матрицу ошибок.
5. Визуализируйте границу принятия решений для двух наиболее важных признаков.
6. Сравните результаты с моделью регрессии и сделайте выводы о применимости каждого подхода.

Примеры домашних заданий самостоятельной работы

Задача 1 Сравнить различные модели машинного обучения для прогнозирования цен на недвижимость и выбрать лучшую.

1. Подготовьте данные (используйте обработанный датасет из предыдущих задач).
2. Обучите следующие модели на обучающей выборке:
 - линейную регрессию;
 - полиномиальную регрессию (степень 2);
 - метод опорных векторов для регрессии;
 - случайный лес ;
 - k-ближайших соседей .

3. Оцените каждую модель на тестовой выборке с помощью MAE, RMSE и R^2 .
4. Сравните время обучения и предсказания для каждой модели.
5. Постройте график сравнения метрик качества для всех моделей.
6. Выберите лучшую модель и обоснуйте свой выбор, учитывая точность, интерпретируемость и вычислительную сложность.

Задача 2. На датасете Adult Income обучите:

случайный лес;

XGBoost;

ансамбль из нескольких моделей (стейкинг).

Сравните результаты по accuracy и F1-score.

Задача 3. Создайте pipeline в sklearn, включающий:

масштабирование признаков;

отбор признаков (SelectKBest);

классификатор (например, SVM).

Обучите модель и оцените качество на тестовой выборке.

Задача 4. Для обученной модели случайного леса: визуализируйте важность признаков; постройте partial dependence plots для 2–3 важных признаков.

Примерные задания контрольной работы (семестр 3)

Блок 1. Работа с датасетом «Ирисы» (классификация)

1. Загрузите встроенный в библиотеку sklearn датасет «Ирисы». Охарактеризуйте датасет: укажите количество объектов, признаков, классов, типы данных. Постройте гистограммы распределения для каждого признака. Сделайте вывод о характере распределения признаков (нормальное/ненормальное).

2. Загрузите датасет «Ирисы». Постройте pairplot (матрицу диаграмм рассеяния) для всех пар признаков с цветовой маркировкой по классам. На основе визуализации сделайте вывод о возможности линейной разделимости классов.

3. Загрузите датасет «Ирисы». Разделите данные на обучающую и тестовую выборки (70% / 30%). Постройте модели классификации с использованием:

- логистической регрессии;
- метода опорных векторов (SVM);
- дерева решений;
- случайного леса.

Оцените точность каждой модели на тестовой выборке. Представьте результаты в виде таблицы. Сделайте вывод о лучшей модели.

4. Загрузите датасет «Ирисы». Примените метод главных компонент (PCA) для понижения размерности до 2 признаков. Визуализируйте получившиеся данные с учётом целевой переменной (используйте точечную диаграмму с цветовой маркировкой классов). Сравните визуализацию с исходными данными. Сделайте вывод об информативности двух главных компонент.

5. Загрузите датасет «Ирисы». Постройте модель классификации методом k -ближайших соседей (k -NN). Подберите оптимальное значение k с помощью кросс-валидации (5 фолдов). Постройте график зависимости точности от k . Сравните результаты с моделью без подбора гиперпараметра.

6. Загрузите датасет «Ирисы». Проведите масштабирование признаков (стандартизацию). Постройте модель SVM на исходных и масштабированных данных. Сравните точность моделей. Объясните влияние масштабирования на работу SVM.

7. Загрузите датасет «Ирисы». Постройте матрицу ошибок для модели случайного леса. Проанализируйте типы ошибок классификации. Рассчитайте точность, полноту и F1-метрику для каждого класса.

Представьте результаты в виде отчёта.

Блок 2. Работа с датасетом «Бостон» (регрессия)

8. Загрузите встроенный в библиотеку `sklearn` датасет «Бостон». Охарактеризуйте датасет: количество объектов, признаков, тип целевой переменной. Постройте `boxplot` для каждого признака, чтобы выявить выбросы. Сделайте вывод о качестве данных.

9. Загрузите датасет «Бостон». Постройте тепловую карту корреляции между признаками и целевой переменной. Выделите 3 наиболее значимых признака для предсказания цены дома. Обоснуйте выбор.

10. Загрузите датасет «Бостон». Разделите данные на обучающую и тестовую выборки (80% / 20%). Постройте модель линейной регрессии. Оцените качество модели с помощью метрик:

- среднеквадратичной ошибки (MSE);
- средней абсолютной ошибки (MAE);
- коэффициента детерминации (R^2).

Сделайте вывод об эффективности модели.

11. Загрузите датасет «Бостон». Примените метод градиентного бустинга для построения регрессионной модели. Сравните результаты с линейной регрессией по метрикам MSE, MAE и R^2 . Представьте результаты в виде таблицы.

12. Загрузите датасет «Бостон». Проведите кросс-валидацию (10 фолдов) для модели линейной регрессии. Рассчитайте среднее значение и стандартное отклонение R^2 по фолдам. Сделайте вывод о стабильности модели.

13. Загрузите датасет «Бостон». Примените метод Lasso-регрессии для отбора признаков. Определите, какие признаки были исключены из модели. Объясните, как регуляризация влияет на интерпретируемость модели.

Блок 3. Работа с датасетом «Цифры» (классификация изображений)

14. Загрузите встроенный в библиотеку sklearn датасет «Цифры». Охарактеризуйте датасет: размер изображений, количество классов, общее число объектов. Визуализируйте первые 16 изображений с подписями истинных классов.

15. Загрузите датасет «Цифры». Постройте модель многослойного перцептрона (MLP) с одним скрытым слоем (100 нейронов). Обучите модель на 80% данных, протестируйте на оставшихся 20%. Рассчитайте точность, полноту, F1-метрику и постройте матрицу ошибок. Сделайте вывод об эффективности классификации.

16. Загрузите датасет «Цифры». Примените метод опорных векторов с RBF-ядром. Оптимизируйте гиперпараметры C и γ с помощью GridSearchCV (5 фолдов). Сравните результаты с базовой моделью SVM без оптимизации.

17. Загрузите датасет «Цифры». Примените метод PCA для понижения размерности до 10 главных компонент. Постройте модель логистической регрессии на исходных данных и на данных после PCA. Сравните точность и время обучения. Сделайте вывод о влиянии понижения размерности на производительность модели.

18. Загрузите датасет «Цифры». Постройте t-SNE-визуализацию данных в 2D-пространстве с цветовой маркировкой по классам. Проанализируйте, насколько хорошо классы разделены в пониженном пространстве. Сравните с результатами PCA.

19. Загрузите датасет «Цифры». Проведите аугментацию данных (поворот, сдвиг, масштабирование изображений). Обучите модель CNN на исходных и аугментированных данных. Сравните точность на тестовой выборке. Сделайте вывод о влиянии аугментации на обобщающую способность модели.

20. Загрузите датасет «Цифры». Постройте ансамблевую модель

(например, голосование нескольких классификаторов: SVM, случайного леса, MLP). Сравните точность ансамбля с точностью отдельных моделей. Объясните преимущества ансамблирования.

21.

Примерные задания проектной работы (семестр 4)

1. Загрузите датасет **Boston Housing**. Проведите предобработку данных: заполните пропуски (если есть), закодируйте категориальные признаки (если есть) и масштабируйте числовые признаки. Обучите модель линейной регрессии на исходных данных и на масштабированных данных. Сравните значения метрик R^2 и MSE на тестовой выборке. Сделайте вывод о влиянии масштабирования признаков на производительность модели.

2. Загрузите датасет **Mall Customer Segmentation Data**. Примените метод k -means для кластеризации клиентов с разным количеством кластеров ($k=2,3,4,5$). Постройте график зависимости инерции (inertia) от количества кластеров (метод «локтя»). Визуализируйте результаты кластеризации для оптимального k с помощью диаграммы рассеяния по двум наиболее значимым признакам. Сделайте вывод о наиболее подходящем количестве сегментов клиентов.

3. Загрузите датасет **Hourly Energy Consumption**. Выберите один временной ряд (один регион). Разделите данные на обучающую и тестовую выборки (последние 10 % данных — тестовая выборка). Обучите модель ARIMA на обучающей выборке. Постройте прогноз на тестовый период и сравните его с реальными значениями. Оцените точность прогноза с помощью метрик MAE и RMSE. Сделайте вывод о применимости модели ARIMA для прогнозирования электропотребления.

4. Загрузите датасет **NSL-KDD**. Проведите предобработку категориальных признаков (one-hot encoding или label encoding). Разделите данные на признаки и целевую переменную (нормальный трафик / аномалия). Обучите модель случайного леса на исходных данных. Затем

примените метод SMOTE для балансировки классов и обучите модель на сбалансированных данных. Сравните точность, полноту и F1-метрику на тестовой выборке для обеих моделей. Сделайте вывод о влиянии балансировки классов на обнаружение аномалий.

5. Загрузите датасет **COVID-19 World Vaccination Progress**. Выберите одну страну. Сформируйте признаки: «доля вакцинированных», «число новых случаев за неделю» и т.д. Разделите данные на обучающую и тестовую выборки. Обучите модель градиентного бустинга для прогнозирования числа новых случаев на следующую неделю. Оцените модель с помощью MAE и R^2 . Постройте график реальных и предсказанных значений. Сделайте вывод о возможности прогнозирования заболеваемости на основе данных о вакцинации.

6. Загрузите датасет **MovieLens 20M**. Постройте простую рекомендательную систему на основе коллаборативной фильтрации (user-based или item-based). Рассчитайте точность рекомендаций с помощью метрики RMSE на отложенной тестовой выборке. Затем примените матричную факторизацию (например, SVD) и оцените точность этой модели. Сравните результаты. Сделайте вывод о преимуществах матричной факторизации для рекомендательных систем.

7. Загрузите датасет **CIFAR-10**. Обучите простую свёрточную нейронную сеть (CNN) с 2–3 свёрточными слоями на исходных данных. Оцените точность модели на тестовой выборке. Затем проведите аугментацию данных (поворот, сдвиг, изменение яркости). Обучите ту же архитектуру CNN на аугментированных данных. Сравните точность двух моделей. Сделайте вывод о влиянии аугментации на обобщающую способность CNN для классификации изображений.

8. Загрузите датасет **Telco Customer Churn**. Проведите предобработку данных (обработка пропусков, кодирование категориальных переменных). Разделите данные на обучающую и тестовую выборки (80% / 20%). Обучите модель логистической регрессии для предсказания оттока.

Постройте матрицу ошибок и рассчитайте точность, полноту, F1-метрику. Затем обучите модель случайного леса на тех же данных. Сравните метрики двух моделей. Сделайте вывод о лучшей модели для задачи предсказания оттока клиентов.

9. Загрузите датасет **Flowers Recognition**. Примените предобработку: измените размер изображений до фиксированного (например, 150×150), нормализуйте пиксели. Обучите простую CNN с 2 свёрточными слоями. Оцените точность на тестовой выборке. Далее используйте трансферное обучение: дообучите модель **VGG16** (заморозив базовые слои) на датасете цветов. Сравните точность простой CNN и дообученной VGG16. Сделайте вывод об эффективности трансферного обучения для классификации изображений с ограниченным числом данных.

10. Загрузите датасет **IMDb Reviews**. Проведите предобработку текста: приведите к нижнему регистру, удалите пунктуацию и стоп-слова, выполните стемминг. Векторизуйте тексты с помощью TF-IDF. Обучите модель SVM для классификации тональности. Оцените точность и F1-метрику на тестовой выборке. Затем используйте встраивание слов (Word2Vec или GloVe) и обучите простую RNN или LSTM-сеть на тех же данных. Сравните результаты двух подходов. Сделайте вывод о лучших методах обработки текста для анализа тональности.

11. Загрузите датасет **20 Newsgroups**. Проведите предобработку текста (очистка, токенизация). Векторизуйте тексты с помощью Bag-of-Words. Обучите наивный байесовский классификатор для определения темы статьи. Оцените точность и полноту модели. Затем примените TF-IDF-векторизацию и обучите тот же классификатор. Сравните метрики. Сделайте вывод о влиянии способа векторизации текста на качество классификации тем.

12. Загрузите датасет **European Soccer Database**. Выберите одну лигу и сезон. Сформируйте признаки (средние показатели команд, результаты последних матчей и т.д.). Разделите данные на обучающую и

тестовую выборки. Обучите модель случайного леса для предсказания исхода матча (победа, ничья, поражение). Оцените точность и матрицу ошибок. Постройте график важности признаков. Сделайте вывод о ключевых факторах, влияющих на результат матча.

13. Загрузите датасет **Brazilian E-commerce**. Проведите предобработку (обработка пропусков, создание признаков: расстояние, день недели и т.д.). Разделите данные на обучающую и тестовую выборки. Обучите модель градиентного бустинга (XGBoost) для предсказания времени доставки. Оцените модель по метрикам MAE, RMSE, R^2 . Постройте график важности признаков. Сделайте вывод о главных факторах, влияющих на сроки доставки.

14. Загрузите датасет **Global Power Plant Database**. Проведите предобработку (агрегация по странам/регионам, создание признаков мощности, типа топлива и т.д.). Разделите данные на обучающую и тестовую выборки. Обучите модель множественной линейной регрессии для прогнозирования энергопотребления страны. Оцените модель по R^2 и MAE. Визуализируйте предсказанные и реальные значения. Сделайте вывод о влиянии типа энергии и географии на энергопотребление.

15. Загрузите датасет **Give Me Some Credit**. Проведите предобработку (заполнение пропусков, масштабирование). Разделите данные на обучающую и тестовую выборки. Обучите модель логистической регрессии для оценки кредитоспособности. Постройте ROC-кривую и рассчитайте AUC. Затем обучите модель XGBoost на тех же данных и сравните AUC. Сделайте вывод о лучшей модели для оценки кредитоспособности и ключевых факторах риска.

16. Загрузите датасет **Mobile App Store**. Проведите предобработку (создание признаков: рейтинг, количество отзывов, категория и т.д.). Разделите данные на обучающую и тестовую выборки. Обучите модель регрессии (линейная регрессия или случайный лес) для предсказания рейтинга приложения. Оцените модель по MAE и R^2 . Постройте график

важности признаков. Сделайте вывод о характеристиках, наиболее влияющих на популярность приложения.

17. Загрузите датасет **NYC Taxi Trip Duration**. Проведите предобработку (извлечение признаков из даты и времени, расчёт расстояния по координатам). Разделите данные на обучающую и тестовую выборки. Обучите модель градиентного бустинга для предсказания продолжительности поездки. Оцените модель по RMSE и MAE. Постройте график важности признаков. Сделайте вывод о факторах, наиболее сильно влияющих на время поездки, и предложите способы оптимизации маршрутов.

Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях кафедры.

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

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

Примерные вопросы для подготовки к зачету (семестр 3)

1. Понятие машинного обучения. Отличие машинного обучения от других областей программирования.

2. Классификация задач машинного обучения. Примеры задач из различных классов.
3. Основные понятия машинного обучения: набора данных, объекты, признаки, атрибуты, модели, параметры.
4. Структура и представление данных для машинного обучения.
5. Инструментальные средства машинного обучения.
6. Задача регрессии: постановка, математическая формализация.
7. Метод градиентного спуска для парной линейной регрессии.
8. Понятие функции ошибки: требования, использование, примеры.
9. Множественная и нелинейная регрессии.
10. Нормализация признаков в задачах регрессии.
11. Задача классификации: постановка, математическая формализация.
12. Метод градиентного спуска для задач классификации.
13. Логистическая регрессия в задачах классификации.
14. Множественная и многоклассовая классификация. Алгоритм “один против всех”.
15. Метод опорных векторов в задачах классификации.
16. Понятие ядра и виды ядер в методе опорных векторов.
17. Метод решающих деревьев в задачах классификации.
18. Метод k ближайших соседей в задачах классификации.
19. Однослойный перцептрон в задачах классификации.

Примерные вопросы для подготовки к экзамену (семестр 4)

1. Метрики эффективности и функции ошибки: назначение, примеры, различия.
2. Понятие набора данных (датасета) в машинном обучении. Требования, представление. Признаки и объекты.
3. Шкалы измерения признаков. Виды шкал, их характеристика.
4. Понятие чистых данных. Определение, очистка данных.

5. Основные этапы проекта по машинному обучению.
6. Предварительный анализ данных: задачи, методы, цели.
7. Проблема отсутствующих данных: причины, исследование, пути решения.
8. Проблема несбалансированных классов: исследование, пути решения.
9. Понятие параметров и гиперпараметров модели. Обучение параметров и гиперпараметров. Поиск по сетке.
10. Понятие недо- и переобучения. Определение, пути решения.
11. Диагностика модели машинного обучения. Методы, цели.
12. Проблема выбора модели машинного обучения. Сравнение моделей.
13. Измерение эффективности работы моделей машинного обучения. Метрики эффективности.
14. Метрики эффективности моделей классификации. Виды, характеристика, выбор.
15. Метрики эффективности моделей регрессии. Виды, характеристика, выбор.
16. Перекрестная проверка (кросс-валидация). Назначение, схема работы.
17. Методы векторизации текстов для задач машинного обучения.
18. Представление графической информации в моделях машинного обучения.
19. Задачи без учителя. Кластеризация. Метод k средних.
20. Задачи без учителя. Обнаружение аномалий.
21. Задачи без учителя. Понижение размерности. Метод PCA.
22. Конвейеры в библиотеке sklearn. Назначение, использование.
23. Использование методов визуализации данных для предварительного анализа.

24. Исследование коррелированности признаков: методы, цели, выводы.
25. Решкалирование данных. Виды, назначение, применение. Нормализация и стандартизация данных.
26. Преобразование категориальных признаков в числовые.
27. Ансамблевые модели машинного обучения. Виды ансамблирования.
28. Конвейеризация моделей машинного обучения.
29. Методы визуализации данных для машинного обучения.
30. Задача выбора модели. Оценка эффективности, валидационный набор.

Примеры практико-ориентированных заданий для подготовки к экзамену (семестр 4)

Практико-ориентированное задание 1.

Предоставлен набор данных о стоимости недвижимости.

Надо разработать и обучить модель машинного обучения для прогнозирования цен на недвижимость на основе различных атрибутов, таких как площадь, количество комнат, район и др.

Для этого необходимо выполнить следующие шаги:

1. Подготовка данных: Импортировать и очистить данные, выполнить предварительную обработку, включая обработку пропущенных значений и кодирование категориальных признаков.
2. Разработка модели: Выбор и настройка модели машинного обучения (например, линейной регрессии или случайного леса), разбиение данных на обучающую и тестовую выборки и обучение модели на обучающих данных.
3. Оценка модели: Оценить эффективность модели на тестовой выборке с помощью таких показателей, как средняя абсолютная ошибка (MAE) или коэффициент детерминации (R^2).

Практико-ориентированное задание 2.

Предоставлен набор данных о стоимости недвижимости. - разработать и обучить модель машинного обучения для прогнозирования цен на недвижимость на основе различных атрибутов, таких как площадь, количество комнат, район и др.

Для этого необходимо выполнить следующие шаги:

1. Подготовка данных
2. Разработка модели
3. Оценка модели
4. Настройка модели
5. Визуализация данных и результатов
6. Отчет и коммуникаци.

Пример экзаменационного билета

Федеральное государственное образовательное бюджетное
учреждение высшего образования
«Финансовый университет при Правительстве Российской Федерации»
(Финансовый университет)

Кафедра: **Корпоративные инфокоммуникационные системы**
Дисциплина: **Машинное обучение**
Филиал: **Владикавказский**; Форма обучения: **Очная**
Семестр: **4** Направление: **09.03.04 Программная инженерия**
Профиль: **Технологии разработки программного обеспечения**

ЭКЗАМЕНАЦИОННЫЙ БИЛЕТ № 1

1. **Теория.** Представление графической информации в моделях машинного обучения.
(20 баллов)
2. **Практико-ориентированное задание.** Разработать модель машинного обучения для прогнозирования цен на недвижимость на основе атрибутов, таких как площадь, количество комнат, район, этаж, расстояние до центр. Изучите структуру данных. Реализуйте простую линейную регрессию для одного признака: площади) (20 баллов)
3. **Практико-ориентированное задание.** На основании разработанной модели машинного обучения из задания 2. Обучите модель, подберите оптимальную скорость обучения (α) и количество итераций. Оцените модель на тестовой выборке, рассчитав MAE, MSE, RMSE и R2. Визуализируйте результаты: постройте график с реальными и предсказанными ценами, а также график сходимости функции стоимости.
(20 баллов)

Подготовил: _____ М.А Ковалева

На основе перечня теоретических вопросов и практико-ориентированных заданий, утвержденного на заседании кафедры «Корпоративные инфокоммуникационные системы» протокол № ____ от _____.2026 г.

Утверждаю:
Заведующий кафедрой _____ М.А Ковалева
Дата _____.2026г.

**Примеры оценочных средств для проверки индикаторов
достижения компетенций, формируемых дисциплиной**

Код и наименование компетенции	Наименование индикатора достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
ОПК-1 Способен применять естественнонаучные и инженерные знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности	1. Демонстрирует знания основных методов математического анализа и моделирования, применяет их на практике для решения задач профессиональной деятельности	Знать: основы математики, линейной алгебры и статистики; основы программирования, особенно на Python; основные алгоритмы машинного обучения. Уметь: разрабатывать и обучать модели машинного обучения; обрабатывать и анализировать данные; оценивать и настраивать модели; визуализировать данные и результаты; эффективно коммуницировать результаты анализа данных.	Вопросы: 1. Понятие машинного обучения. Отличие машинного обучения от других областей программирования. 2. Классификация задач машинного обучения. Примеры задач из различных классов. Практико-ориентированное задание Предоставлен набор данных о стоимости недвижимости. Надо разработать и обучить модель машинного обучения для прогнозирования цен на недвижимость на основе различных атрибутов, таких как площадь, количество комнат, район и др. Для этого необходимо выполнить следующие шаги: 1. Подготовка данных: Импортировать и очистить данные, выполнить предварительную обработку, включая обработку пропущенных значений и кодирование категориальных признаков. 2. Разработка модели: Выбор и настройка модели машинного обучения (например, линейной регрессии или случайного леса), разбиение данных на обучающую и тестовую выборки и обучение модели на обучающих данных. 3. Оценка модели: Оценить

			<p>эффективность модели на тестовой выборке с помощью таких показателей, как средняя абсолютная ошибка (MAE) или коэффициент детерминации (R^2).</p> <p>4. Настройка модели: Попробовать различные гиперпараметры модели и методы регуляризации для улучшения ее работы.</p> <p>5. Визуализация данных и результатов: Создать визуализации для изучения данных и представления результатов моделирования. Например, графики зависимости цен от различных атрибутов.</p> <p>6. Отчет и коммуникация: Подготовить отчет, включающий описание проблемы, методологии, результатов и выводов. Представить результаты визуализации и оценки модели в понятной форме.</p>
2. Проводит теоретические исследования по выбранной области профессиональной деятельности.	<p>Знать: основы машинного обучения и математики, связанной с этой областью.</p> <p>Уметь: создавать и настраивать модели машинного обучения; анализировать данные и применять их в моделях; критически оценивать методы машинного обучения; эффективно коммуницировать результаты теоретических исследований.</p>	<p>Вопросы:</p> <p>1. Основные понятия машинного обучения: набора данных, объекты, признаки, атрибуты, модели, параметры.</p> <p>2. Структура и представление данных для машинного обучения.</p> <p>3. Инструментальные средства машинного обучения.</p> <p>4. Задача регрессии: постановка, математическая формализация.</p> <p>5. Метод градиентного спуска для парной линейной регрессии.</p> <p>Практико-ориентированное задание</p> <p>Предоставлен набор данных о стоимости недвижимости. - разработать и обучить модель машинного обучения для прогнозирования цен на</p>	

		<p>недвижимость на основе различных атрибутов, таких как площадь, количество комнат, район и др.</p> <p>Для этого необходимо выполнить следующие шаги:</p> <ol style="list-style-type: none"> 1. Подготовка данных 2. Разработка модели 3. Оценка модели 4. Настройка модели 5. Визуализация данных и результатов 6. Отчет и коммуникации
3. Проводит численные эксперименты на основе математических или информационных методов, делает выводы и обосновывает их.	<p>Знать: основы машинного обучения и статистики, методы и модели машинного обучения.</p> <p>Уметь: планировать и проводить численные эксперименты; анализировать данные и извлекать выводы из численных экспериментов; обосновывать полученные результаты на основе математических и статистических методов; эффективно коммуницировать результаты численных экспериментов.</p>	<p>Вопросы:</p> <ol style="list-style-type: none"> 1. Понятие набора данных (датасета) в машинном обучении. Требования, представление. Признаки и объекты. 2. Шкалы измерения признаков. Виды шкал, их характеристика. 3. Понятие чистых данных. Определение, очистка данных. 4. Основные этапы проекта по машинному обучению. <p>Предварительный анализ данных: задачи, методы, цели</p> <ol style="list-style-type: none"> 5. Метрики эффективности моделей классификации.. 6. Метрики эффективности моделей регрессии 7. Методы визуализации данных для машинного обучения.. <p>.</p> <p>Практико-ориентированное задание</p> <p>Предоставлен набор данных о стоимости недвижимости - разработать и обучить модель машинного обучения для прогнозирования цен на недвижимость на основе различных атрибутов, таких как площадь, количество комнат, район и др.</p> <p>Для этого необходимо выполнить следующие шаги:</p> <ol style="list-style-type: none"> 1. Подготовка данных

			2. Разработка модели 3. Оценка модели 4. Настройка модели 5. Визуализация данных и результатов 6. Отчет и коммуникация
--	--	--	--

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

Основная литература

1. Колдаев, В. Д. Структуры и алгоритмы обработки данных : учебное пособие / В. Д. Колдаев. - Москва : РИОР : ИНФРА-М, 2021. - 296 с. - ЭБС ZNANIUM.com. - URL: <https://znanium.com/catalog/product/1230215>. – Текст : электронный.

2. Коротеев, М. В. Основы машинного обучения на Python : учебник / М. В. Коротеев. — Москва : КноРус, 2024. — 431 с. — ЭБС BOOK.ru. — URL: <https://book.ru/book/952751>. — Текст : электронный.

Дополнительная литература

3. Нагаева, И. А. Основы алгоритмизации и программирования: практикум : учебное пособие / И. А. Нагаева, И. А. Кузнецов. – Москва : Берлин : Директ-Медиа, 2021. – 169 с. – ЭБС Университетская библиотека ONLINE. – URL: <https://biblioclub.ru/index.php?page=book&id=598404>. – Текст : электронный.

4. Златопольский, Д. М. Программирование: типовые задачи, алгоритмы, методы: учебное пособие / Д. М. Златопольский. — 4-е изд. (эл.). — Москва : Лаборатория знаний, 2020. — 226 с.: ил. — ЭБС Университетская библиотека ONLINE. — URL: <https://biblioclub.ru/index.php?page=book&id=222873>. – Текст : электронный.

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 Znanium
<http://www.znanium.com>
3. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» <http://biblioclub.ru/>
4. Электронно-библиотечная система издательства «ЮРАЙТ»
<https://www.biblio-online.ru>
5. Электронная библиотека издательского дома «Гребенников»
<https://grebennikon.ru>
6. Электронно-библиотечная система издательства «Лань»
<https://e.lanbook.com>
7. Pylru 1.0.9 [Электронный ресурс]: сайт. – Режим доступа:
<https://pypi.python.org/pypi/pylru>
8. Python Data Analysis Library [Электронный ресурс]: сайт. – Режим доступа: <http://pandas.pydata.org/>
9. Python Documentation [Электронный ресурс]: сайт. – Режим доступа: <http://python.org/doc/>
10. Scikit-learn Machine Learning in Python [Электронный ресурс]: сайт. – Режим доступа: <http://scikit-learn.org>

10. Методические указания для обучающихся по освоению дисциплины

Методика освоения дисциплины предусматривает подготовку обучающихся к лекциям, семинарам и практическим занятиям, выполнение студентами самостоятельной внеаудиторной работы.

Рекомендации по подготовке к лекционным занятиям.

Для наиболее полного освоения дисциплины студентам необходимо:

- перед каждой лекцией просматривать рабочую программу дисциплины, ее основные вопросы и рекомендуемую литературу. Это позволит сэкономить время на записывание основных вопросов темы;
- перед очередной лекцией просматривать материалы предыдущих,

чтобы освоение материала не оставляло пробелов.

Рекомендации по подготовке к семинарам, практическим занятиям.

Студентам следует:

- проработать теоретический материал к занятию по рекомендованным литературным источникам и лекциям;
- использовать при подготовке к занятию нормативно-правовые документы, научные публикации, информационный материал, рекомендуемый преподавателем;
- перед занятиями задать вопросы по невыясненным в ходе самостоятельной подготовки темам или отдельным положениям темы;
- в ходе занятия давать четкие и исчерпывающие ответы на вопросы;
- на занятии демонстрировать понимание обсуждаемых тем и вопросов.

Студентам, пропустившим занятия по различным причинам, необходимо перед очередным занятием отработать пропущенный материал, подготовив его самостоятельно.

Методические рекомендации по выполнению различных форм самостоятельной работы

Студентам при организации самостоятельной работы следует руководствоваться Приказом Финансового университета № 1040/о от 11.05.2021г. «Об утверждении методических рекомендаций по планированию и организации внеаудиторной самостоятельной работы студентов по образовательным программам бакалавриата и магистратуры в Финансовом университете».

Самостоятельная работа содержит в себе различные виды и формы работ. Самостоятельная работа студентов включает в себя выполнение различного рода заданий, которые ориентированы на более глубокое усвоение материала изучаемой дисциплины. По теме учебной дисциплины студентам предлагается перечень заданий для самостоятельной работы.

В ходе изучения дисциплины предусмотрены следующие формы

самостоятельной работы:

- подготовка к опросу;
- разбор вопросов, отводимых на самостоятельное освоение,
- решение задач;
- выполнение контрольной работы, проектной работы;
- подготовка к зачету и экзамену.

К выполнению заданий для самостоятельной работы предъявляются следующие требования: задания должны выполняться самостоятельно и представляться в установленный срок, а также должны соответствовать установленным требованиям по оформлению.

Студентам следует:

- руководствоваться графиком самостоятельной работы, определенным РПД;
- выполнять все плановые задания, выдаваемые преподавателем для самостоятельного выполнения, разбирать на занятиях и консультациях неясные вопросы;
- прорабатывать соответствующие теоретические и практические разделы дисциплины, фиксируя неясные фрагменты для их обсуждения на консультации.

Методические рекомендации для обучающихся по выполнению контрольной работы

Контрольная работа является обязательной формой внеаудиторной самостоятельной работы студентов по дисциплине и может реализовываться как в письменном виде, так и с использованием информационных технологий и специализированных программных продуктов.

Цель выполнения контрольной работы, содержащей комплект заданий – овладение студентами навыками решения типовых расчетных задач, формирование учебно-исследовательских навыков, закрепление умений самостоятельно работать с различными источниками информации; проверка сформированности компетенций.

Целью выполнения контрольной работы является углубление и закрепление теоретических знаний и практических навыков студентов по дисциплине.

Контрольная работа по дисциплине выполняется по вариантам.

Содержание заданий контрольных работ охватывают основной материал соответствующих разделов (тем) дисциплин. Контрольные задания разрабатываются по многовариантной системе. Варианты контрольных работ равноценны по объему и сложности.

Контрольная работа выполняется студентом под руководством преподавателя кафедры «Корпоративные инфокоммуникационные системы», ведущим семинарские (практические) занятия.

Контрольная работа состоит из нескольких частей. Состав контрольной работы и очередность размещения отдельных частей:

- титульный лист;
- основная часть;
- список использованных источников;
- приложения (при наличии).

Титульный лист является первой страницей и заполняется по определенным правилам.

Основная часть выполняется согласно заданиям (вопросам) контрольных работ.

В список использованных источников включаются названия законодательных актов, нормативных документов, книг, статей, учебных пособий и т. п., которые, так или иначе, использовались студентом при выполнении работы.

В Приложения выносятся вспомогательные материалы, которые не содержат основную информацию, либо материалы, которые сложно разместить по тексту работы (большие схемы, таблицы, графические материалы, расчетные справочные данные, образцы первичных документов и т.п.). Непременным условием включения данных материалов в

приложение является ссылка на них в тексте работы.

Требования к выполнению контрольной работы:

- четкость и последовательность изложения материала (решения) в соответствии с составленным планом;
- наличие обобщений и выводов, сделанных на основе изучения информационных источников по данной теме;
- предоставление в полном объеме решений имеющихся в задании практических задач;
- использование современных способов поиска, обработки и анализа информации;
- самостоятельность выполнения.

Требования к оформлению контрольной работы.

Контрольная работа выполняется на компьютере (гарнитура Times New Roman, шрифт 13 или 14) через 1-1,5 интервала с полями: верхнее, нижнее - 2; правое - 3; левое - 1,5. Отступ первой строки абзаца - 1,25. Нумерация страниц – внизу в центре.

Иллюстративный материал (схемы, диаграммы, рисунки, таблицы и др.) встраивается в текст работы или выносится в Приложения.

При написании допускаются только общепринятые сокращения (например, тыс. руб.).

В тексте обязательны ссылки на литературные источники, лучше всего постраничные.

Объем контрольной работы составляет не более 6 страниц, не включая таблиц, графиков и т.п. (при наличии).

Законченная контрольная работа, содержащая все требуемые элементы оформления, вставленная в папку (или файл) и скрепленная по левому краю, сдается на кафедру или непосредственно руководителю контрольной работы – преподавателю; ведущему семинарские (практические) занятия по дисциплине. Он осуществляет проверку контрольной работы, а также оказывает помощь при подготовке к ее защите.

Контрольная работа защищается в назначенные сроки. Защита работы проводится до начала сессии (в крайнем случае, до начала экзамена по соответствующему предмету). При защите студент кратко излагает основные положения работы, последовательность ее выполнения, свои предложения.

При защите работы студент должен свободно ориентироваться в изложенном материале работы; ответить на все замечания преподавателя; уметь отвечать на вопросы преподавателя по выполненной работе.

Оценка контрольных работ студентов проводится в процессе текущего контроля успеваемости студентов.

Критерии оценки контрольной работы

Оценка «отлично» (5 баллов) выставляется студенту, показавшему всесторонние, систематизированные, глубокие знания вопросов контрольной работы /и/или умение уверенно применять их на практике при решении конкретных задач.

Оценка «хорошо» (3-4 балла) выставляется студенту, если он твердо знает материал контрольной работы, грамотно и, по существу, излагает его /и/или умеет применять полученные знания на практике при решении конкретных задач, но допускает некоторые неточности.

Оценка «удовлетворительно» (1-2 балла) выставляется студенту, показавшему фрагментарный, разрозненный характер знаний, обнаружившему нарушения логической последовательности в изложении материала, но при этом владеющему основными вопросами, выносимыми на контрольную работу и необходимыми для дальнейшего обучения /и/или умеющему применять полученные знания по образцу в стандартной ситуации.

Оценки «неудовлетворительно» (0 баллов) заслуживает студент, который не знает большей части основного содержания выносимых на контрольную работу вопросов, тем дисциплины, допускает грубые ошибки в формулировках основных понятий /и/или не умеет использовать полученные знания при решении типовых практических задач.

При оценивании контрольной работы на «неудовлетворительно» она должна быть переделана (исправлена) в соответствии с полученными замечаниями, сдана на проверку заново и защищена не позднее срока окончания ее приёма и защиты.

Методические рекомендации для обучающихся по выполнению проектной работы

Проектная работа является обязательной формой внеаудиторной самостоятельной работы студентов по дисциплине.

Целью проектной работы является развитие у студентов способности прогнозировать, проектировать, моделировать, формирование учебно-исследовательских навыков, закрепление умений самостоятельно работать с различными источниками информации; проверка сформированности компетенций.

Проектная работа может выполняться как индивидуально, так и в составе группы. Количество групп и их численный состав определяет преподаватель, ведущий семинарские занятия.

Заказчиками выполнения проекта могут являться представители работодателей. В этом случае проектная работа выполняется исходя из потребностей заказчика.

Выполнение проекта предполагает:

- диагностику ситуации (проблематизация, целеполагание, конкретизация цели, форматирование проекта);
- проектирование (уточнение цели, функций, задач и плана работы; теоретическое моделирование методов и средств решения задач; детальная проработка этапов решения конкретных задач; пошаговое выполнение запланированных проектных действий; систематизация и обобщение полученных результатов, конструирование предполагаемого результата, пошаговое выполнение проектных действий);
- рефлексия (выяснение соответствия полученного результата замыслу; определение качества полученного продукта; перспективы его развития и

использования);

- фиксация результатов в виде выполненного проекта.

Проектная работа состоит из нескольких частей. Состав проектной работы и очередность размещения отдельных частей:

- титульный лист;
- основная часть;
- список использованных источников;
- приложения (при наличии).

Титульный лист является первой страницей проектной работы и заполняется по определенным правилам.

Основная часть выполняется согласно заданию преподавателя или исходя из потребностей заказчика.

В список использованных источников включаются названия законодательных актов, нормативных документов, книг, статей, учебных пособий и т. п., которые, так или иначе, использовались студентом при выполнении работы.

Иллюстративный материал (схемы, диаграммы, рисунки, таблицы и др.) встраивается в текст работы или выносится в Приложения. В Приложения выносятся вспомогательные материалы, которые не содержат основную информацию, либо материалы, которые сложно разместить по тексту работы (большие схемы, таблицы, графические материалы, расчетные справочные данные, образцы первичных документов и т.п.). Непременным условием включения данных материалов в приложение является ссылка на них в тексте работы.

Требования к оформлению проектной работы.

Проектная работа выполняется на компьютере на одной стороне белой бумаги формата А4 (210х297 мм). Размер шрифта -13 или 14, междустрочный интервал – одинарный или полуторный.

Размеры полей: левое - 30 мм, правое - 15 мм, верхнее - 20 мм, нижнее - 20 мм. Отступ первой строки абзаца - 1,25. Нумерация страниц – внизу в центре.

При написании допускаются только общепринятые сокращения (например, тыс. руб.).

Общий объем проектной работы составляет не более 10 страниц, не включая таблицы, графики и т.п. (при наличии), а также приложения (при наличии).

В тексте обязательны ссылки на литературные источники, лучше всего постраничные.

Законченная проектная работа, содержащая все требуемые элементы оформления, вставленная в папку (или файл) и скрепленная по левому краю, сдается на кафедру или непосредственно руководителю проектной работы – преподавателю; ведущему семинарские (практические) занятия по дисциплине. Он осуществляет проверку проектной работы, а также оказывает помощь при подготовке к ее защите.

Проектная работа защищается в назначенные сроки. Защита проектной работы проводится до начала сессии (в крайнем случае, до начала экзамена по соответствующему предмету). При защите студент кратко излагает основные положения работы, последовательность ее выполнения, свои предложения.

При защите проектной работы студент должен свободно ориентироваться в изложенном материале работы; ответить на все замечания преподавателя; уметь отвечать на вопросы преподавателя по проектной работе.

Оценка проектных работ студентов проводится в процессе текущего контроля успеваемости.

Критерии оценки проектной работы

Оценка «отлично» (5 баллов) выставляется студенту, если проектная работа отличается творческим (креативным) подходом, собственным

оригинальным отношением автора к идее проекта; содержит полную диагностику ситуации, а также теоретическое моделирование методов и детальную проработку этапов решения конкретных задач; в работе сделаны необходимые выводы, намечены перспективы использования проекта, спланированы действия по его продвижению; работа отличается грамотным оформлением в точном соответствии с установленными правилами, с соблюдением логической последовательности изложения материала; студент в работе выдвигает новые идеи и трактовки, демонстрирует способность анализировать материал; на дополнительные вопросы при защите проектной работы даны полные ответы.

Оценка «хорошо» (3-4 балла) выставляется студенту, если проектная работа содержит достаточно полную диагностику ситуации, а также теоретическое моделирование методов и этапов решения конкретных задач; в работе сделаны выводы, намечены перспективы использования проекта; работа оформлена правильно с учетом 1-2 мелких погрешностей или 2-3 недочетов, исправленных самостоятельно или по требованию преподавателя; в работе соблюдена логическая последовательность изложения материала; студент в работе демонстрирует творческие способности и хорошую способность анализировать материал. На дополнительные вопросы при защите проектной работы даны не совсем полные ответы.

Оценка «удовлетворительно» (1-2 балла) выставляется студенту, если проектная работа содержит отдельные элементы моделирования методов и этапов решения конкретных задач; в работе сделаны выводы, намечены перспективы использования проекта; работа выполнена и оформлена правильно, но в ней допущены 1-2 погрешности или одна грубая ошибка; в работе соблюдена логическая последовательность изложения материала; студент в работе демонстрирует удовлетворительную способность анализировать материал; допущены ошибки при ответе на дополнительные вопросы при защите проектной работы.

Оценки «неудовлетворительно» (0 баллов) заслуживает студент, если в работе отсутствуют элементы моделирования; студент в работе не проявил способность анализировать, прогнозировать и проектировать; в работе отсутствует логическая последовательность изложения материала, допущены грубые ошибки, которые обучающийся не может исправить даже по требованию преподавателя или работа не выполнена полностью.

При оценивании проектной работы на «неудовлетворительно» работа должна быть переделана (исправлена) в соответствии с полученными замечаниями, сдана на проверку заново и защищена не позднее срока окончания ее приёма и защиты.

Оценка результатов текущего контроля успеваемости и промежуточной аттестации обучающихся осуществляется в соответствии с Балльно-рейтинговой системой Финансового университета (Приказ Финансового университета № 2187/о от 01.10.2024 г. «Об утверждении Положения о проведении текущего контроля успеваемости и промежуточной аттестации студентов, обучающихся по образовательным программам высшего образования в Финансовом университете»).

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

11.1 Комплект лицензионного программного обеспечения

- 1) Антивирусная защита Kaspersky Security для виртуальных и облачных сред;
- 2) Windows, Microsoft Office или Astra Linux, Libre Office.
- 3) Среда разработки Jupyter Notebook.
- 4) Браузер.

11.2 Современные профессиональные базы данных, и информационные справочные системы

1. Информационно-правовая система «Гарант»: <https://www.garant.ru>
2. Большая Российская энциклопедия: <https://bigenc.ru/>
3. Система комплексного раскрытия информации «СКРИН» - <http://www.skrin.ru/>.

11.3 Сертифицированные программные и аппаратные средства защиты информации

Не используются

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

Учебная аудитория для проведения учебных занятий, предусмотренных программой бакалавриата, оснащенная оборудованием и техническими средствами обучения:

Аудитория № 45

Специализированная мебель:

Стол (учительский) – 1 шт.

Стол компьютерный – 1 шт.

Стол (студенческий) двухместный – 13 шт.

Стулья – 27 шт.

Доска меловая – 1 шт.

Технические средства обучения:

Компьютер в сборе – 1 шт.

Доска интерактивная – 1 шт.

Подключение к сети «Интернет» и обеспечение доступа в электронную информационно-образовательную среду Финансового университета

Учебная аудитория для проведения учебных занятий, предусмотренных программой бакалавриата, оснащенная оборудованием и техническими средствами обучения

Аудитория № 42

Специализированная мебель:

Стол компьютерный – 20 шт.

Стол (двухместный) – 7 шт.

Стул – 34 шт.

Шкаф – 1 шт.

Технические средства обучения:

Компьютер в сборе – 20 шт.

Мультимедиа-проектор – 1 шт.

Экран настенный – 1 шт.

Подключение к сети «Интернет» и обеспечение доступа в электронную информационно-образовательную среду Финансового университета

Помещение для самостоятельной работы обучающихся:

Кабинет № 55. Читальный зал:

Специализированная мебель:

Стол – 20 шт.

Стул – 40 шт.

Шкаф для книг – 4 шт.

Стеллаж книжный – 13 шт.

Стеллаж выставочный – 4 шт.

Технические средства обучения:

Компьютер в сборе – 6 шт.

Телевизор – 1 шт.

Помещение для самостоятельной работы обучающихся оснащено компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Финансового университета