

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Силин Яков Петрович

Должность: Ректор

Дата подписания: 04.06.2026 15:21:57

Уникальный программный идентификатор:

24f866be2aca16484036a8cbb3c509a9571b60ff

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФГБОУ ВО «Уральский государственный экономический университет»

Одобрена

на заседании кафедры

09.12.2025 г.

протокол № 12

И.о. зав. кафедрой Кольева Н.С.

Утверждена

Советом по учебно-методическим  
вопросам и качеству образования

16 декабря 2025 г.

протокол № 4

Председатель

Карх Д.А.

(подпись)



### РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

|                         |   |
|-------------------------|---|
| Наименование дисциплины | Управление бизнес-аналитической информацией                         |
| Направление подготовки  | 38.04.01 Экономика  |
| Профиль                 | Экономика организаций: бизнес-анализ и управление результативностью |
| Форма обучения          | очно-заочная  |
| Год набора              | 2026  |
| Разработана:            |   |
| Доцент, к.э.н.          |   |
| Кольева Н.С.            |   |

Екатеринбург  
2025 г.

## СОДЕРЖАНИЕ

|  |           |
|--|-----------|
| <b>ВВЕДЕНИЕ</b>  | <b>3</b>  |
| <b>1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ</b>   | <b>3</b>  |
| <b>2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП</b>  | <b>3</b>  |
| <b>3. ОБЪЕМ ДИСЦИПЛИНЫ</b>   | <b>3</b>  |
| <b>4. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ОПОП</b>   | <b>3</b>  |
| <b>5. ТЕМАТИЧЕСКИЙ ПЛАН</b>  | <b>6</b>  |
| <b>6. ФОРМЫ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ<br/>ШКАЛЫ ОЦЕНИВАНИЯ</b>  | <b>7</b>  |
| <b>7. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ</b>  | <b>8</b>  |
| <b>8. ОСОБЕННОСТИ ОРГАНИЗАЦИИ ОБРАЗОВАТЕЛЬНОГО<br/>ПРОЦЕССА ПО ДИСЦИПЛИНЕ ДЛЯ ЛИЦ С ОГРАНИЧЕННЫМИ<br/>ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ</b>   | <b>11</b> |
| <b>9. ПЕРЕЧЕНЬ ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ УЧЕБНОЙ ЛИТЕРАТУРЫ,<br/>НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ</b>   | <b>11</b> |
| <b>10. ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ВКЛЮЧАЯ ПЕРЕЧЕНЬ<br/>ЛИЦЕНЗИОННОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И<br/>ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ, ОНЛАЙН КУРСОВ,<br/>ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО<br/>ПРОЦЕССА ПО ДИСЦИПЛИНЕ</b> | <b>12</b> |
| <b>11. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ<br/>ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО<br/>ДИСЦИПЛИНЕ</b>  | <b>13</b> |

## ВВЕДЕНИЕ

Рабочая программа дисциплины является частью основной профессиональной образовательной программы высшего образования - программы магистратуры, разработанной в соответствии с ФГОС ВО

|         |   |
|---------|---|
| ФГОС ВО | Федеральный государственный образовательный стандарт высшего образования - магистратура по направлению подготовки 38.04.01 Экономика (приказ Минобрнауки России от 11.08.2020 г. № 939) |
|---------|---|

### 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Формирование знаний, умений и навыков в области обработки больших данных в корпорациях, применения статистических, математических и эмпирических методов анализа экономической информации и извлечения знаний из больших массивов данных.

### 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина относится к части, формируемой участниками образовательных отношений.

### 3. ОБЪЕМ ДИСЦИПЛИНЫ

| Промежуточная аттестация | Часов            |                                 |        |              |  | 3.е. |
|--------------------------|------------------|---------------------------------|--------|--------------|--|------|
|                          | Всего за семестр | Контактная работа .(по уч.зан.) |        |              | Самостоятельная работа в том числе подготовка контрольных и курсовых |      |
|                          |                  | Всего                           | Лекции | Лабораторные |  |      |
| Семестр 3                |                  |                                 |        |              |  |      |
| Зачет с оценкой          | 144              | 24                              | 8      | 16           | 120  | 4    |

### 4. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОСВОЕНИЯ ОПОП

В результате освоения ОПОП у выпускника должны быть сформированы компетенции, установленные в соответствии ФГОС ВО.

| Шифр и наименование компетенции | Индикаторы достижения компетенций |
|---------------------------------|-----------------------------------|
| аналитический                   |                                   |

|  |  |
|--|--|
| <p>ПК-1 Определение направлений развития организации</p> | <p>ИД-1.ПК-1 Знать: Методики оценки деятельности в соответствии с разработанными показателями; Методы сборки, анализа, систематизации, хранения и поддержания в актуальном состоянии информации для бизнес-анализа; Предметная область и специфика деятельности организации в объеме, достаточном для решения задач бизнес-анализа; Инструменты, техники моделирования, анализа бизнес-ситуации и предметной области; Возможности использования свободно распространяемого программного обеспечения в организации в объеме, необходимом для целей бизнес-анализа; Перспективные и существующие цифровые технологии и цифровые возможности для бизнеса в контексте предметной области и специфики деятельности организации; Управление рисками; Теория межличностной и групповой коммуникации в деловом взаимодействии; Методика организации и управления временем (тайм-менеджмент); Способы оценки и основы обеспечения информационной безопасности в объеме, необходимом для целей бизнес-анализа; Стратегии управления изменениями; Языки и инструменты визуального моделирования</p>   |
|  | <p>ИД-2.ПК-1 Уметь: Планировать, организовывать и проводить встречи и обсуждения с заинтересованными сторонами, в том числе с использованием инструментов (платформ) онлайн-коммуникаций; Использовать техники эффективных коммуникаций; Выявлять, регистрировать, анализировать и классифицировать риски, разрабатывать комплекс мероприятий по управлению ими ;Оформлять результаты бизнес-анализа в соответствии с выбранными подходами; Определять связи и зависимости между элементами информации для бизнес-анализа; Применять ИТ-инструменты (приложения и платформы) для обеспечения работ по бизнес-анализу; Представлять информацию для бизнес-анализа различными способами и в различных форматах для обсуждения с заинтересованными сторонами; Анализировать внутренние/внешние факторы и условия, влияющие на деятельность организации; Анализировать требования заинтересованных сторон с точки зрения критериев качества, определяемых выбранными подходами; Производить анализ деятельности организации; Производить анализ деятельности организации; Отбирать, применять и адаптировать соответствующие методы, инструменты и техники анализа бизнес-ситуации и предметной области, включая методы анализа данных; Пользоваться системами анализа и визуализации данных</p> |
|  | <p>ИД-3.ПК-1 Иметь практический опыт: Оценка текущего состояния организации; Определение параметров будущего состояния организации; Выявление, анализ и оценка несоответствия между параметрами текущего и будущего состояний организации; Оценка бизнес-возможностей организации, необходимых для проведения стратегических изменений в организации.</p>  |

|   |  |
|---|--|
| <p>ПК-2 Разработка стратегии управления изменениями в организации</p> | <p>ИД-1.ПК-2 Знать: Теория заинтересованных сторон; Методики оценки деятельности организации в соответствии с разработанными показателями; Сбор, анализ, систематизация, хранение и поддержание в актуальном состоянии информации для бизнес-анализа; Методика организации и управления временем (тайм-менеджмент); Стратегии управления изменениями; Способы оценки и основы обеспечения информационной безопасности в объеме, необходимом для целей бизнес-анализа; Теория межличностной и групповой коммуникации в деловом взаимодействии; Перспективные и существующие цифровые технологии и цифровые возможности для бизнеса в контексте предметной области и специфики деятельности организации; Возможности использования свободно распространяемого программного обеспечения в организации в объеме, необходимом для целей бизнес-анализа; Основанные на планах и на изменениях (гибкие) подходы к организации проектной работы; Управление рисками.</p>   |
|   | <p>ИД-2.ПК-2 Уметь: Планировать, организовывать и проводить встречи и обсуждения с заинтересованными сторонами, в том числе с использованием инструментов (платформ) онлайн-коммуникаций; Использовать техники эффективных коммуникаций; Выявлять, регистрировать, анализировать и классифицировать риски, разрабатывать комплекс мероприятий по управлению ими; Оформлять результаты бизнес-анализа в соответствии с выбранными подходами; Определять связи и зависимости между элементами информации для бизнес-анализа; Представлять информацию для бизнес-анализа различными способами и в различных форматах для обсуждения с заинтересованными сторонами; Анализировать внутренние/внешние факторы и условия, влияющие на деятельность организации; Анализировать требования заинтересованных сторон с точки зрения критериев качества, определяемых выбранными подходами; Применять ИТ-инструменты (приложения и платформы) для обеспечения работ по бизнес-анализу; Производить анализ деятельности организации; Моделировать объем и границы работ; Моделировать текущее и будущее (желаемое) состояние организации; Отбирать, применять и адаптировать соответствующие методы, инструменты и техники анализа бизнес-ситуации и предметной области, включая методы анализа данных; Выбирать и обосновывать модели управления изменениями.</p> |
|   | <p>ИД-3.ПК-2 Иметь практический опыт: Определение цели и задач стратегических изменений в организации; Определение основных параметров и ключевых показателей эффективности разрабатываемых стратегических изменений в организации; Определение критериев оценки успеха стратегических изменений в организации; Оценка соответствия изменений стратегическим целям организации; Определение заинтересованных сторон, которые должны быть вовлечены в инициативу по реализации стратегических изменений в организации; Определение основных аспектов организации, которые могут быть затронуты стратегическими изменениями; Выбор стратегии в составе группы экспертов; Определение промежуточных состояний при реализации выбранной стратегии изменений в организации; Разработка планов реализации стратегических изменений в организации; Мониторинг процесса проведения стратегических изменений в организации</p>  |

|  |   |
|--|---|
| организационно-управленческий                            |   |
| ПК-3 Обоснование подходов, используемых в бизнес-анализе | ИД-1.ПК-3 Знать: Теория межличностной и групповой коммуникации в деловом взаимодействии; Теория конфликтов; Методы, техники, процессы и инструменты управления требованиями; Языки и инструменты визуального моделирования; Управление рисками; Теория менеджмента; Управление ресурсами; Теория организационного развития; Теория систем; Методика организации временем и управления им (тайм-менеджмент); Способы оценки и основы обеспечения информационной безопасности в объеме, необходимом для целей бизнес-анализа; Основанные на планах и на изменениях (гибкие) подходы к организации проектной работы.                                 |
|  | ИД-2.ПК-3 Уметь: Планировать, организовывать и проводить встречи и обсуждения с заинтересованными сторонами в том числе с использованием инструментов (платформ) онлайн-коммуникаций; Выявлять, регистрировать, анализировать и классифицировать риски, разрабатывать комплекс мероприятий по управлению ими; Оформлять результаты бизнес-анализа в соответствии с выбранными подходами; Применять IT-инструменты (приложения и платформы) для обеспечения работ по бизнес-анализу; Отбирать, применять и адаптировать соответствующие методы, инструменты и техники анализа бизнес-ситуации и предметной области, включая методы анализа данных. |
|  | ИД-3.ПК-3 Иметь практический опыт: Определение подхода к проведению бизнес-анализа; Определение подхода к работе с информацией для бизнес-анализа; Определение подхода к работе с заинтересованными сторонами; Определение подхода к разработке различных типов требований; Определение подхода к работе с изменениями различных типов требований; Определение подхода к оценке эффективности работы по бизнес-анализу  |

## 5. ТЕМАТИЧЕСКИЙ ПЛАН

| Тема      | Часов  |             |                                |              |                      |                |                                 |
|-----------|--|-------------|--------------------------------|--------------|----------------------|----------------|---------------------------------|
|           | Наименование темы  | Всего часов | Контактная работа (по уч.зан.) |              |                      | Самост. работа | Контроль самостоятельной работы |
|           |  |             | Лекции                         | Лабораторные | Практические занятия |                |                                 |
| Семестр 3 |  | 144         |                                |              |                      |                |                                 |
| Тема 1.   | Статистические модели: критерии и методы их оценивания. Инструменты построения статистических моделей (ПК-1, ПК-2, ПК-3) | 26          | 2                              | 4            |                      | 20             |                                 |
| Тема 2.   | Регрессионные модели (ПК-1, ПК-2, ПК-3)  | 26          | 2                              | 4            |                      | 20             |                                 |
| Тема 3.   | Модели классификации (ПК-1, ПК-2, ПК-3)  | 24          | 4                              | 2            |                      | 18             |                                 |
| Тема 4.   | Ансамблевые модели (ПК-1, ПК-2, ПК-3)  | 20          |                                | 2            |                      | 18             |                                 |
| Тема 5.   | Методы кластеризации и понижения размерности (ПК-1, ПК-2, ПК-3)  | 22          |                                | 2            |                      | 20             |                                 |
| Тема 6.   | Нейронные сети (ПК-1, ПК-2, ПК-3)  | 26          |                                | 2            |                      | 24             |                                 |

## 6. ФОРМЫ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ШКАЛЫ ОЦЕНИВАНИЯ

| Раздел/Тема                             | Вид оценочного средства    | Описание оценочного средства  | Критерии оценивания  |
|---|----------------------------|---|--|
| Текущий контроль (Приложение 4)         |                            |   |  |
| Темы 1-2                                | Тест<br>(приложение 4)     | Тест состоит из 50-ти вопросов                                      | 10 баллов  |
| Темы 3-4                                | Тест<br>(приложение 4)     | Тест состоит из 16-ти вопросов                                      | 10 баллов  |
| Темы 5-6                                | Реферат<br>(приложение 4)  | Объем реферата от 15 до 25 страниц                                  | 10 баллов  |
| Промежуточная аттестация (Приложение 5) |                            |   |  |
| 3 семестр<br>(ЗаО)                      | Билет для зачета с оценкой | Билет состоит из 2 теоретических вопросов и 1 практического задания | оцениваются в баллах: 25 баллов за теоретическое задание и 50 баллов за практическое задание, итого 100 баллов |

### ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ

Показатель оценки освоения ОПОП формируется на основе объединения текущего контроля и промежуточной аттестации обучающегося.

Показатель рейтинга по каждой дисциплине выражается в процентах, который показывает уровень подготовки студента.

Текущий контроль. Используется 100-балльная система оценивания. Оценка работы студента в течении семестра осуществляется преподавателем в соответствии с разработанной им системой оценки учебных достижений в процессе обучения по данной дисциплине.

В рабочих программах дисциплин и практик закреплены виды текущего контроля, планируемые результаты контрольных мероприятий и критерии оценки учебных достижений.

В течение семестра преподавателем проводится не менее 3-х контрольных мероприятий, по оценке деятельности студента. Если посещения занятий по дисциплине включены в рейтинг, то данный показатель составляет не более 20% от максимального количества баллов по дисциплине.

Промежуточная аттестация. Используется 5-балльная система оценивания. Оценка работы студента по окончанию дисциплины (части дисциплины) осуществляется преподавателем в соответствии с разработанной им системой оценки достижений студента в процессе обучения по данной дисциплине. Промежуточная аттестация также проводится по окончанию формирования компетенций.

Порядок перевода рейтинга, предусмотренных системой оценивания, по дисциплине, в пятибалльную систему.

Высокий уровень – 100% - 70% - отлично, хорошо.

Средний уровень – 69% - 50% - удовлетворительно.

| Показатель оценки | По 5-балльной системе | Характеристика показателя  |
|-------------------|-----------------------|--|
| 100% - 85%        | отлично               | обладают теоретическими знаниями в полном объеме, понимают, самостоятельно умеют применять, исследовать, идентифицировать, анализировать, систематизировать, распределять по категориям, рассчитать показатели, классифицировать, разрабатывать модели, алгоритмизировать, управлять, организовать, планировать процессы исследования, осуществлять оценку результатов на высоком уровне   |
| 84% - 70%         | хорошо                | обладают теоретическими знаниями в полном объеме, понимают, самостоятельно умеют применять, исследовать, идентифицировать, анализировать, систематизировать, распределять по категориям, рассчитать показатели, классифицировать, разрабатывать модели, алгоритмизировать, управлять, организовать, планировать процессы исследования, осуществлять оценку результатов.<br><br>Могут быть допущены недочеты, исправленные студентом самостоятельно в процессе работы (ответа и т.д.) |
| 69% - 50%         | удовлетворительно     | обладают общими теоретическими знаниями, умеют применять, исследовать, идентифицировать, анализировать, систематизировать, распределять по категориям, рассчитать показатели, классифицировать, разрабатывать модели, алгоритмизировать, управлять, организовать, планировать процессы исследования, осуществлять оценку результатов на среднем уровне.<br>Допускаются ошибки, которые студент затрудняется исправить самостоятельно.  |
| 49 % и менее      | неудовлетворительно   | обладают не полным объемом общих теоретическими знаниями, не умеют самостоятельно применять, исследовать, идентифицировать, анализировать, систематизировать, распределять по категориям, рассчитать показатели, классифицировать, разрабатывать модели, алгоритмизировать, управлять, организовать, планировать процессы исследования, осуществлять оценку результатов.<br>Не сформированы умения и навыки для решения профессиональных задач                                       |
| 100% - 50%        | зачтено               | характеристика показателя соответствует «отлично», «хорошо», «удовлетворительно»   |
| 49 % и менее      | не зачтено            | характеристика показателя соответствует «неудовлетворительно»  |

## 7. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

### 7.1. Содержание лекций

|   |
|---|
| <p>Тема 1. Статистические модели: критерии и методы их оценивания. Инструменты построения статистических моделей (ПК-1, ПК-2, ПК-3)<br/>         Задача машинного обучения. Интеллектуальный анализ данных. Обучение с учителем и без учителя. Вероятностная постановка задачи. Регрессионная функция. Байесов классификатор. Метод главных компонент. Проверка статистических гипотез.</p> |
| <p>Тема 2. Регрессионные модели (ПК-1, ПК-2, ПК-3)<br/>         Регрессия. Проверка значимости и доверительные интервалы для коэффициентов. Подготовка данных. Переобучение.</p>  |
| <p>Тема 3. Модели классификации (ПК-1, ПК-2, ПК-3)<br/>         Наивный байесовский классификатор. Дискриминантный анализ. МНК для задачи классификации. Логистическая регрессия.</p>   |

## 7.2 Содержание практических занятий и лабораторных работ

|   |
|---|
| <p>Тема 2. Регрессионные модели (ПК-1, ПК-2, ПК-3)<br/><br/>         Линейная регрессия. Функционал качества и градиентный спуск.</p>   |
| <p>Тема 3. Модели классификации (ПК-1, ПК-2, ПК-3)<br/><br/>         Классификация в бинарных пространствах с использованием классических моделей. Бинарные деревья решений. Поиск логических закономерностей в данных. Алгоритмы выделения ассоциативных правил. Анализ последовательностей знаков или событий.<br/>         Дискриминантный анализ. Метод опорных векторов. Ядерные функции машины опорных векторов. Деревья классификации, случайный лес и логистическая регрессия. Процедуры сравнения эффективности моделей классификации.<br/>         Ирисы Фишера и метод k-ближайших соседей. Наивный классификатор Байеса. Классификация в линейном дискриминантном пространстве. Нелинейные классификаторы в R. Модель мультиномиального логита.</p> |
| <p>Тема 4. Ансамблевые модели (ПК-1, ПК-2, ПК-3)<br/><br/>         Решающие деревья. Случайный лес. Градиентный бустинг. Применение ансамблевых моделей.</p>  |
| <p>Тема 5. Методы кластеризации и понижения размерности (ПК-1, ПК-2, ПК-3)<br/><br/>         Задача кластеризации. Группы методов. Метод k-средних. Иерархическая кластеризация. Агломеративный алгоритм. DBSCAN. Оценки качества кластеризации.<br/>         Метод главных компонент. Сингулярное разложение матрицы и связь с PCA. Применение PCA на данных. Многомерное шкалирование. t-SNE.<br/>         Рекомендательные системы. Методы коллаборативной фильтрации. Методы с матричными разложениями.</p>   |
| <p>Тема 6. Нейронные сети (ПК-1, ПК-2, ПК-3)<br/><br/>         Обучение нейросети. Сверточные сети. Рекуррентные сети. Современные архитектуры. Введение в TensorFlow. Классификация изображений на Tensorflow.</p>   |

### 7.3. Содержание самостоятельной работы

|   |
|---|
| Тема 2. Регрессионные модели (ПК-1, ПК-2, ПК-3)<br>Изучение основной и дополнительной литературы, интернет-источников по теме. Разбор практических примеров. Изучение функций языка R. Выполнение практических работ.                         |
| Тема 3. Модели классификации (ПК-1, ПК-2, ПК-3)<br>Изучение основной и дополнительной литературы, интернет-источников по теме. Разбор практических примеров. Изучение функций языка R. Выполнение практических работ.                         |
| Тема 4. Ансамблевые модели (ПК-1, ПК-2, ПК-3)<br>Изучение основной и дополнительной литературы, интернет-источников по теме. Разбор практических примеров. Изучение функций языка R. Выполнение практических работ.                           |
| Тема 5. Методы кластеризации и понижения размерности (ПК-1, ПК-2, ПК-3)<br>Изучение основной и дополнительной литературы, интернет-источников по теме. Разбор практических примеров. Изучение функций языка R. Выполнение практических работ. |
| Тема 6. Нейронные сети (ПК-1, ПК-2, ПК-3)<br>Изучение основной и дополнительной литературы, интернет-источников по теме. Разбор практических примеров. Изучение функций языка R. Выполнение практических работ.                               |

7.3.1. Примерные вопросы для самостоятельной подготовки к зачету/экзамену  
Приложение 1.

7.3.2. Практические задания по дисциплине для самостоятельной подготовки к зачету/экзамену  
Приложение 2.

7.3.3. Перечень курсовых работ  
Не предусмотрено.

7.4. Электронное портфолио обучающегося  
Материалы не размещаются.

7.5. Методические рекомендации по выполнению контрольной работы  
Не предусмотрено.

7.6 Методические рекомендации по выполнению курсовой работы  
Не предусмотрено.

## 8. ОСОБЕННОСТИ ОРГАНИЗАЦИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ ДЛЯ ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

### *По заявлению студента*

В целях доступности освоения программы для лиц с ограниченными возможностями здоровья при необходимости кафедра обеспечивает следующие условия:

- особый порядок освоения дисциплины, с учетом состояния их здоровья;
- электронные образовательные ресурсы по дисциплине в формах, адаптированных к ограничениям их здоровья;
- изучение дисциплины по индивидуальному учебному плану (вне зависимости от формы обучения);
- электронное обучение и дистанционные образовательные технологии, которые предусматривают возможности приема-передачи информации в доступных для них формах.
- доступ (удаленный доступ), к современным профессиональным базам данных и информационным справочным системам, состав которых определен РПД.

## 9. ПЕРЕЧЕНЬ ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ УЧЕБНОЙ ЛИТЕРАТУРЫ, НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Сайт библиотеки УрГЭУ

<http://lib.usue.ru/>

### **Основная литература:**

2. Сарычев А.П. Регрессионный анализ динамических систем [Электронный ресурс]: Монография. - Москва: ООО "Научно-издательский центр ИНФРА-М", 2022. - 229 – Режим доступа: <https://znanium.com/catalog/product/1865377>

3. Акопов А. С. Имитационное моделирование [Электронный ресурс]: учебник и практикум для вузов. - Москва: Юрайт, 2023. - 389 – Режим доступа: <https://urait.ru/bcode/511425>

4. Зараменских Е. П. Основы бизнес-информатики [Электронный ресурс]: учебник и практикум для вузов. - Москва: Юрайт, 2024. - 470 – Режим доступа: <https://urait.ru/bcode/536967>

5. Колошкина И. Е. Автоматизация проектирования технологической документации [Электронный ресурс]: учебник и практикум для вузов. - Москва: Юрайт, 2024. - 371 – Режим доступа: <https://urait.ru/bcode/543895>

6. Кузнецова Е. В. Управление портфелем проектов как инструмент реализации корпоративной стратегии [Электронный ресурс]: учебник для вузов. - Москва: Юрайт, 2024. - 177 – Режим доступа: <https://urait.ru/bcode/537265>

7. Маслевич Т.П. Управление бизнес-процессами: от теории к практике [Электронный ресурс]: Учебное пособие. - Москва: ООО "Научно-издательский центр ИНФРА-М", 2025. - 206 – Режим доступа: <https://znanium.com/catalog/product/2187631>

### **Дополнительная литература:**

2. Ниворожкина Л.И., Арженовский С.В. Статистические методы анализа данных [Электронный ресурс]: Учебник. - Москва: Издательский Центр РИО, 2016. - 333 – Режим доступа: <https://znanium.com/catalog/product/556760>

3. Гоулман Д., Друкер П., Дэвенпорт Т. Управление бизнесом [Электронный ресурс]:. - Москва: ООО "Альпина Паблишер", 2016. - 290 с. – Режим доступа: <https://znanium.com/catalog/product/701800>

4. Кислицын Е. В., Першин В. К. Компьютерное имитационное моделирование: системная динамика и агенты [Электронный ресурс]:учебное пособие. - Екатеринбург: [Издательство УрГЭУ], 2016. - 122 – Режим доступа: <http://lib.usue.ru/resource/limit/ump/16/p487700.pdf>

5. Ли В. А. Моделирование делового совершенства организации [Электронный ресурс]:учебное пособие. - Екатеринбург: Издательство УрГЭУ, 2020. - 124 – Режим доступа: <http://lib.usue.ru/resource/limit/ump/20/p493241.pdf>

6. Паштова Л.Г., Шохин Е.И., Булава И.В., Гермогентова М.Н., Мингалиев К.Н., Сетченкова Л.А., Слепнева Т.А., Талимова Л.А., Серегина Е.Ю., Орлова О.В., Малофеев С.Н., Корнилова Е.В., Лихачева О.Н., Хотинская Г.И., Паштова Л.Г., Шохин Е.И. Корпоративные финансы и управление бизнесом [Электронный ресурс]:Монография. - Москва: КноРус, 2020. - 400 с. – Режим доступа: <https://book.ru/book/932075>

7. Яковлев В.Б. Регрессионный анализ. Расчеты в Excel и Statistica [Электронный ресурс]:Учебное пособие. - Москва: Русайнс, 2021. - 176 с. – Режим доступа: <https://book.ru/book/938371>

8. Паштова Л.Г., Шохин Е.И., Булава И.В., Гермогентова М.Н., Мингалиев К.Н., Сетченкова Л.А., Слепнева Т.А., Талимова Л.А., Серегина Е.Ю., Орлова О.В., Малофеев С.Н., Корнилова Е.В., Лихачева О.Н., Хотинская Г.И., Паштова Л.Г., Шохин Е.И. Корпоративные финансы и управление бизнесом [Электронный ресурс]:Монография. - Москва: КноРус, 2021. - 392 с. – Режим доступа: <https://book.ru/book/940048>

9. Андреева Л.В., Бодрова Т.В., Зубарева Е.В., Ингман Н.И., Крятова Л.А. Учетно-аналитическая система: теория и практика [Электронный ресурс]:Монография. - Москва: Издательско-торговая корпорация "Дашков и К", 2020. - 292 с. – Режим доступа: <https://znanium.com/catalog/product/1449821>

10. Соколов Г.А., Сагитов Р. В. Введение в регрессионный анализ и планирование регрессионных экспериментов в экономике [Электронный ресурс]:Учебное пособие. - Москва: ООО "Научно-издательский центр ИНФРА-М", 2022. - 202 с. – Режим доступа: <https://znanium.com/catalog/product/1836638>

11. Тельнов, Ю. Ф., Фёдоров, И. Г. Инжиниринг предприятия и управление бизнес-процессами. Методология и технология [Электронный ресурс]:учебное пособие для студентов магистратуры, обучающихся по направлению «прикладная информатика». - Москва: ЮНИТИ-ДАНА, 2017. - 207 с. – Режим доступа: <https://www.iprbookshop.ru/81628.html>

## **10. ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ВКЛЮЧАЯ ПЕРЕЧЕНЬ ЛИЦЕНЗИОННОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ, ОНЛАЙН КУРСОВ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ**

### **Перечень лицензионного программного обеспечения:**

Astra Linux Common Edition. Договор №0417-ПО/2019 от 08.05.2019, Акт №Sk000343 от 24.05.2019 и Контракт № 35-У/2018 от 13.06.2018, Акт № УТ213 от 17.12.2018. Срок действия лицензии - без ограничения срока.

Язык программирования R.Лицензия GNU GPL 2.Срок действия лицензии - без ограничения срока.

R Studio (среда для языка программирования R).Лицензия GNU Affero General Public License v3.Срок действия лицензии - без ограничения срока.

Microsoft Windows 10 .Договор № 52/223-ПО/2020 от 13.04.2020, Акт № Tr000523459 от 14.10.2020. Срок действия лицензии -Без ограничения срока.

Microsoft Office 2016. Договор № 52/223-ПО/2020 от 13.04.2020, Акт № Tr000523459 от 14.10.2020 Срок действия лицензии -Без ограничения срока.

МойОфис стандартный. Соглашение № СК-281 от 7 июня 2017. Дата заключения - 07.06.2017. Срок действия лицензии - без ограничения срока.

**Перечень информационных справочных систем, ресурсов информационно-телекоммуникационной сети «Интернет»:**

Справочно-правовая система Консультант +. Договор № 143/223-У/2025 от 02.12.2025 Срок действия лицензии до 31.12.2026

Справочно-правовая система Гарант. Договор № 58419 от 22 декабря 2015. Срок действия лицензии -без ограничения срока

**Статистические методы в управлении инновациями**

<https://openedu.ru/course/ITMOUniversity/INMAN/>

**Р для лингвистов: программирование и анализ данных**

<https://openedu.ru/course/hse/RLING/>

**11. ОПИСАНИЕ МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ**

Реализация учебной дисциплины осуществляется с использованием материально-технической базы УрГЭУ, обеспечивающей проведение всех видов учебных занятий и научно-исследовательской и самостоятельной работы обучающихся:

Специальные помещения представляют собой учебные аудитории для проведения всех видов занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду УрГЭУ.

Все помещения укомплектованы специализированной мебелью и оснащены мультимедийным оборудованием спецоборудованием (информационно-телекоммуникационным, иным компьютерным), доступом к информационно-поисковым, справочно-правовым системам, электронным библиотечным системам, базам данных действующего законодательства, иным информационным ресурсам служащими для представления учебной информации большой аудитории.

Для проведения занятий лекционного типа презентации и другие учебно-наглядные пособия, обеспечивающие тематические иллюстрации.

### 7.3.1. Примерные вопросы для самостоятельной подготовки к экзамену

1. Вероятностная постановка задачи обучения с учителем. Функция потерь. Средний (ожидаемый) риск. Эмпирический риск. Регрессионная функция и байесов классификатор. Неустраняемая (байесовская) ошибка.
2. Ошибки 1-го и 2-го рода. Чувствительность, специфичность, точность, полнота. ROC-кривая. Площадь под ROC-кривой.
3. Принцип минимизации эмпирического риска. Минимизация отступа. Регуляризация.
4. Экспериментальная (эмпирическая) оценка качества обучения. Обучающая, проверочная и тестовая выборки. Метод перекрестного (скользящего) контроля.
5. Метод k ближайших соседей в задачах классификации и восстановления регрессии. Теорема об оценке риска в методе ближайшего соседа. Идея доказательства.
6. Метод наименьших квадратов. Система нормальных уравнений. Псевдорешение.
7. Борьба с переобучением в методе наименьших квадратов. Сокращение числа параметров. Полный перебор всех подмножеств признаков. Жадный (Forward stepwise) алгоритм.
8. Борьба с переобучением в методе наименьших квадратов. Ридж-регрессия (регуляризация). Метод "Лассо".
9. Наивный байесовский классификатор. Сглаживание Лапласа. Использование наивного байесовского классификатора для количественных признаков.
10. Линейный дискриминантный анализ. Квадратичный дискриминантный анализ.
11. Логистическая регрессия. Логистическая функция и softmax
12. Нейронные сети. Алгоритм обучения Back-Propagation.
13. Понятие о глубоком обучении. Сверточные нейронные сети.
14. Машина опорных векторов. Формулировка задачи в виде задачи математического программирования. Двойственная задача. (Случай линейно разделимых и неразделимых классов)
15. Машина опорных векторов. Ядра и спрямляющие пространства.
16. Деревья решений. Алгоритм CART.
17. Случайный лес. Экстремально случайные деревья.
18. Алгоритм XGBoost.
19. Градиентный бустинг деревьев решений.
20. Дилемма "Смещение-разброс". Кривая обучения.
21. Задача понижения размерности. Метод главных компонент.

22. Задача кластеризации. Метод центров тяжести. Метод медоидов.
23. Алгоритм "Ожидание-максимизация".
24. Алгоритм DBSCAN.
25. Алгоритмы иерархической кластеризации.

**7.3.2. Практические задания по дисциплине для самостоятельной подготовки к зачету/экзамену**

**Примерные практические задания к зачету (ПК-1, ПК-2, ПК-3)**

1. Пусть  $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $x \in \mathbb{R}^n$ . Матрицей Якоби называется матрица

$$\frac{\partial g}{\partial x} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \dots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \dots & \frac{\partial g_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \dots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}.$$

В частности, если  $m = 1$  (т. е.  $g(x)$  — скалярная функция векторного аргумента  $x$ ), то  $\partial g / \partial x$  — градиент функции  $g$ .

Доказать, что

- 1) если  $a \in \mathbb{R}^n$ ,  $x \in \mathbb{R}^n$ , то  $\partial(a^T x) / \partial x = a$ ;
- 2) если  $A \in \mathbb{R}^{m \times n}$ ,  $x \in \mathbb{R}^n$ , то  $\partial(Ax) / \partial x = A$ ;
- 3) если  $A \in \mathbb{R}^{n \times n}$ ,  $x \in \mathbb{R}^n$ , то  $\partial(x^T Ax) / \partial x = (A + A^T)x$ ; в частности, если  $A^T = A$ , то  $\partial(x^T Ax) / \partial x = 2Ax$ ;

4) если  $x \in \mathbb{R}^n$ , то  $\partial|x|^2 / \partial x = 2x$ ;

5) если  $g$  — скалярная функция и под  $g(x)$  понимается применение функции  $g$  к каждой компоненте вектора  $x \in \mathbb{R}^n$ , то

$$\partial g(x) / \partial x = \text{diag}(g'(x)),$$

где  $\text{diag}(a)$  — диагональная матрица с диагональю  $a$ ;

6) если  $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $g: \mathbb{R}^m \rightarrow \mathbb{R}^p$ ,  $x \in \mathbb{R}^n$ , то

$$\partial g(h(x)) / \partial x = (\partial g(h(x)) / \partial h)(\partial h(x) / \partial x).$$

2. Пользуясь №1, найдите градиент  $\partial g(\beta) / \partial \beta$  и гессиан  $(\partial^2 g(\beta)) / (\partial \beta^T \partial \beta)$

функции  $g(\beta) = |X\beta - y|^2$ . Выведите отсюда, что решение линейной задачи наименьших квадратов  $\hat{\beta} = \text{argmin} |X\beta - y|^2$  является решением нормальной системы линейных уравнений  $X^T X \beta = X^T y$ .

3. Дана обучающая выборка

|   |   |   |   |   |    |
|---|---|---|---|---|----|
| x | 1 | 1 | 0 | 0 | -1 |
| y | 4 | 4 | 0 | 2 | 6  |

1) изобразить точки;

2) методом наименьших квадратов построить модель вида  $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$ ; построить график этой функции;

3) построить модель того же вида методом ридж-регрессии с параметром регуляции  $\lambda = 1$ ; построить график этой функции.

Замечание: при ручных вычислениях по методу наименьших квадратов рекомендуется составить систему  $X^T X \beta = X^T y$  и решить ее. Регуляризованная система  $(X^T X + \lambda I) \beta = X^T y$ , где  $I$  — единичная матрица.

4. Рассмотрим задачу восстановления регрессии, в которой  $y$  распределен согласно нормальному закону  $N(X\beta, \sigma^2 I)$ , а  $\beta$  имеет априорное распределение  $N(0, \tau I)$ . Найти апостериорное распределение для  $\beta$ . Доказать, что  $\beta^{\text{ridge}}$  есть его математическое ожидание. Найти связь между параметром регуляризации  $\lambda$  и дисперсиями  $\tau$ ,  $\sigma^2$ .

5. Показать, что процедура гребневой регрессии эквивалентна обычному методу наименьших квадратов, примененному к расширенным данным: к централизованной матрице  $X$  дописывается матрица  $(\sqrt{\lambda})I$ , к вектору  $y$  приписывается  $d$  нулей.

6. Показать, как (и объяснить почему) задачу квадратичного программирования в методе лассо

$$\hat{\beta}^{\text{lasso}} = \underset{\beta}{\operatorname{argmin}} \left\{ \sum_{i=1}^N \left( y_i - \beta_0 - \sum_{j=1}^d x_{ij} \beta_j \right)^2 \right\},$$

при условии

$$\sum_{j=1}^d |\beta_j| \leq s$$

можно свести к задаче квадратичного программирования с  $2d+1$  неизвестными и  $2d+1$  линейными ограничениями.

7. Метод использования линейной регрессии в задаче классификации заключается в следующем. Сопоставим каждому классу  $k$  вектор  $(y_1, y_2, \dots, y_k)$ , в котором  $y_k = 1$  а  $y_i = 0$  при  $i \neq k$ . Собрав вместе индикаторные векторы объектов обучающей выборки, получим матрицу  $Y$  размера  $N \times K$ . Пусть  $X$  — матрица размера  $N \times (d+1)$ , первый столбец которой состоит из единиц, а последующие представляют собой векторы из обучающей выборки. Применяя метод наименьших квадратов одновременно к каждому столбцу матрицы  $Y$ , получаем значения  $Y^{\wedge} = X(X^T X)^{-1} X^T Y$ .

Для каждого столбца  $y_k$  матрицы  $Y$  получим свой столбец коэффициентов  $\beta^{\wedge}_k$ . Соберем их в матрицу  $B^{\wedge}$  размера  $(d+1) \times K$ . Имеем

$$B^{\wedge} = (X^T X)^{-1} X^T Y.$$

Объект  $x$  будем классифицировать согласно следующему правилу: Вычислим вектор-строку длину  $K$

$$g(x) = (1, x) B^{\wedge}.$$

Отнесем  $x$  к классу

$$f(x) = \underset{k}{\operatorname{argmax}} g_k(x)$$

Доказать, что

$$\sum_{k=1}^K g_k(x) = 1.$$

Доказать, что в случае  $K = 2$  данный метод эквивалентен решению одной задачи восстановления регрессии. Какой?

8. Дана обучающая выборка

|       |    |   |   |   |   |   |   |   |
|-------|----|---|---|---|---|---|---|---|
| $x_1$ | 0  | 1 | 0 | 2 | 2 | 2 | 4 | 3 |
| $x_2$ | -1 | 0 | 0 | 0 | 1 | 0 | 1 | 2 |
| $y$   | 0  | 0 | 0 | 0 | 0 | 1 | 1 | 1 |

1) Методом линейного дискриминантного анализа для каждого класса построить дискриминантную функцию и записать уравнение разделяющей поверхности.

2) Методом квадратичного дискриминантного анализа построить дискриминантные функции.

Изобразить точки и разделяющие поверхности (кривые)

9. Задача Фишера сводится к максимизации отношения Рэлея

$$\max_a (a^T B a) / (a^T W a).$$

Показать, как эта задача сводится к обобщенной задаче на собственные значения

$$B a = \lambda W a.$$

10. Показать, что оптимальная гиперплоскость, разделяющая два множества, является плоскостью, проходящей через середину отрезка, соединяющего пару ближайших точек из

выпуклой оболочки каждого из классов, и перпендикулярно ему. Указание: рассмотреть задачу, двойственную к задаче определения оптимальной гиперплоскости.

**11.** Показать, что в алгоритме SVM задача

$$\min_{\beta, \beta_0, \xi_i} \frac{1}{2} \|\beta\|^2 + \gamma \sum_{i=1}^N \xi_i,$$

при ограничениях

$$y_i(x_i^\top \beta + \beta_0) \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad (i = 1, 2, \dots, N),$$

эквивалентна задаче

$$\min_{\beta, \beta_0} \sum_{i=1}^N \left[ 1 - y_i (x_i^\top \beta + \beta_0) \right]_+ + \alpha \|\beta\|^2,$$

где  $[\cdot]_+$  означает положительную часть, и  $\alpha = 1/(2\gamma)$ .

**12.** SVM и задача восстановления регрессии. Для восстановления  $\beta, \beta_0$  в модели  $f(x) = x^\top \beta + \beta_0$ . Рассмотрим задачу минимизации функции

$$H(\beta, \beta_0) = \sum_{i=1}^N V(y_i - f(x_i)) + \frac{\alpha}{2} \|\beta\|^2,$$

где

$$V(t) = V_\varepsilon(t) = \begin{cases} 0, & \text{если } |t| < \varepsilon, \\ |t| - \varepsilon & \text{в противном случае.} \end{cases}$$

Доказать, что решение  $\hat{\beta}, \hat{\beta}_0$ , минимизирующее функцию  $\beta H(\beta, \beta_0)$ , можно представить в виде

$$\hat{\beta} = \sum_{i=1}^N (\hat{\alpha}_i^* - \hat{\alpha}_i) x_i, \quad \hat{f}(x) = \sum_{i=1}^N (\hat{\alpha}_i^* - \hat{\alpha}_i) \langle x, x_i \rangle + \hat{\beta}_0,$$

где  $\hat{\alpha}_i$  и  $\hat{\alpha}_i^*$  являются решением следующей задачи квадратичного программирования:

$$\min_{\alpha_i, \alpha_i^*} \left( \varepsilon \sum_{i=1}^N (\alpha_i^* + \alpha_i) - \sum_{i=1}^N y_i (\alpha_i^* - \alpha_i) + \frac{1}{2} \sum_{i,j=1}^N (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) \langle x_i, x_j \rangle \right)$$

при ограничениях

$$0 \leq \alpha_i, \alpha_i^* \leq \frac{1}{\lambda}, \quad \sum_{i=1}^N (\alpha_i^* - \alpha_i) = 0, \quad \alpha_i \alpha_i^* = 0.$$

**13.** Дана обучающая выборка

|       |    |    |    |   |
|-------|----|----|----|---|
| $x_1$ | -1 | -1 | 1  | 1 |
| $x_2$ | -1 | 1  | -1 | 1 |
| $y$   | 1  | -1 | -1 | 1 |

Подобрать ядро и указать соответствующий SVM-классификатор, для которого ошибка на обучающей выборке равна 0.

**14.** Дана обучающая выборка

|       |   |   |   |   |   |   |   |   |   |   |
|-------|---|---|---|---|---|---|---|---|---|---|
| $x_1$ | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 |
| $x_2$ | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $y$   | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |

С помощью наивного байесова классификатора оценить вероятность  $\Pr(Y = 0|X_1 = 1, X_2 = 1)$ ;  $\Pr(Y = 1|X_1 = 1, X_2 = 1)$ .

15. Дана выборка

|       |   |   |    |    |
|-------|---|---|----|----|
| $x_1$ | 4 | 0 | -2 | 2  |
| $x_2$ | 3 | 1 | -3 | -1 |

Найти главные направления и дисперсии по главным компонентам.

Изобразить точки и главные направления.

16. Дана выборка

|       |   |    |    |   |    |
|-------|---|----|----|---|----|
| $x_1$ | 4 | 0  | -1 | 3 | 4  |
| $x_2$ | 2 | -3 | -2 | 1 | 2  |
| $x_3$ | 3 | 2  | 2  | 1 | -3 |

Найти главные направления и дисперсии по главным компонентам.

17. Пусть  $\sigma(z) = 1/(1 + e^{-z})$  (сигмоидальная функция). Проверить, что  $\sigma' = \sigma(1 - \sigma)$ .

18. Пусть в задаче классификации на  $K$  классов  $\{1, 2, \dots, K\}$  последний слой нейронной сети вычисляет softmax-функцию:

$$g_k(s_1, s_2, \dots, s_K) = e^{s_k} / \sum_{\ell=1}^K e^{s_\ell}.$$

В качестве штрафа используется кросс-энтропия (logloss-функция):

$$R^{(i)} = - \sum_{k=1}^K I(y^{(i)} = k) \ln g_k(s_1, s_2, \dots, s_K),$$

где  $g_k(s_1, s_2, \dots, s_K)$  — softmax-функция. Доказать, что

1)  $\partial g_k / \partial s_\ell = g_k \cdot I(k = \ell) - g_\ell$ ;

2)  $\partial R^{(i)} / \partial g_k = - I(y^{(i)} = k) / g_k$ ;

3)  $\partial R^{(i)} / \partial s_\ell = g_\ell - I(\ell = y^{(i)})$ .

19. Предположим, что рассматривается  $K$  задач двухклассовой классификации, в каждой из которых по  $x \in X$  требуется предсказать  $y_k \in \{0, 1\}$  ( $k = 1, 2, \dots, K$ ) (например, требуется определить, присутствует или отсутствует на изображении  $x$  объект  $k$ ). Обучающая выборка составлена из пар  $(x^{(i)}, y^{(i)})$ , где  $y^{(i)} = (y^{(i)}_1, y^{(i)}_2, \dots, y^{(i)}_K) \in \{0, 1\}^K$  ( $i = 1, 2, \dots, N$ ). Для решения такой задачи можно использовать нейронную сеть, последний слой которой вычисляет  $K$  сигмоидальных функций

$$g_k(s_k) = 1/(1 + e^{-s_k}) \quad (k = 1, 2, \dots, K).$$

Пусть в качестве штрафа используется

$$R^{(i)} = - \sum_{k=1}^K (y_k^{(i)} \ln g_k + (1 - y_k^{(i)}) \ln(1 - g_k)).$$

Доказать, что

$$\partial R^{(i)} / \partial s_k = g_k - y_k^{(i)}.$$

20. Нейронная сеть с двумя нелинейными слоями вычисляет функцию  $\text{softmax}(B(\sigma(Ax)))$ .

Пусть в качестве функции потерь используется logloss. Таким образом, на объекте  $x^{(i)}$  потери равны  $\text{logloss}(\text{softmax}(B(\sigma(Ax))))$ . Пользуясь результатом задачи №1, выпишите матричные формулы для алгоритма backpropagation.

21. Пусть рассматривается задача 2-классовой классификации. Доказать, что если известно сколько в выборке представителей каждого из двух классов, то по любым двум показателям из списка TPR, TNR, PPV, NPV определяются остальные два.

22. Пусть рассматривается задача 2-х классовой классификации. Верно ли, что

- 1) если у двух классификаторов на одной и той же выборке совпадают PPV (Precision) и совпадают TPR (Recall или Sensitivity), то будут совпадать TNR (Specificity) и NPV;
- 2) если у двух классификаторов на одной и той же выборке совпадают TNR (Specificity) и совпадают NPV, то будут совпадать PPV(Precision) и TPR (Recall или Sensitivity);
- 3) совпадение ROC кривых(для двух классификаторов на одной и той же выборке) влечет совпадение Precision-Recall кривых и наоборот?

**23.** Рассмотрим задачу восстановления регрессии с квадратичной функцией потерь  $L(y', y) = (y' - y)^2$ . Доказать, что если  $f^*(x) = E(Y | X = x)$  (регрессионная функция). Чему тогда равен средний риск  $R(f^*)$ ?

**24.** Рассмотрим задачу восстановления регрессии с функцией потерь  $L(y', y) = |y' - y|$ . Доказать, что минимум среднему риску доставляет при этом условная медиана  $f(x) = \text{median}(Y | X = x)$ .

**25.** Пусть в задаче классификации с двумя классами  $\{0, 1\}$  используется функция потерь  $L(y', y)$ , такая, что  $L(0, 0) = L(1, 1) = 0$ ,  $L(1, 0) = l_1$ ,  $L(0, 1) = l_0$ . Докажите, что в этом случае байесов классификатор  $f^*(x)$  удовлетворяет условию

$$f(x) = \underset{y \in \{0, 1\}}{\operatorname{argmax}} \ell_y \Pr(y|x).$$

**26.** Выразить байесов классификатор  $f^*(x)$  для классификации с  $K$  классами, если функция потерь равна  $L(y', y) = \ell y' y$  ( $y', y = 1, 2, \dots, K$ ).

**27.** Пусть  $N$  точек распределены случайно равномерно в единичной  $d$ -мерной гиперсфере. Доказать, что медианное расстояние от центра сферы до ближайшей точки равно

$$\rho(d, N) = \sqrt[d]{1 - \sqrt[N]{1/2}}.$$

Найти предел  $\rho(d, N)$  при  $d \rightarrow \infty$ ,  $N = O(d)$ . Какой вывод из этого можно сделать применительно к методу ближайшего соседа при больших  $d$ ?

**28.** Bias-variance trade-off. Рассмотрим задачу восстановления зависимости  $Y = f^*(X)$ , где  $X$  - случайная величина, а  $f^*$  - неизвестная детерминированная функция. Пусть  $x^{(1)}, x^{(2)}, \dots, x^{(N)}$  суть независимые реализации величины  $X$ . В качестве модельной зависимости возьмем функцию  $f(x, D)$ , где  $D = x^{(1)}, x^{(2)}, \dots, x^{(N)}$ . Разложить  $E_D f(x, D) - f^*(x)^2$  в сумму квадрата математического ожидания смещения (bias) и дисперсии (variance).

**29.** Может ли использование коррелированных переменных улучшить качество предсказания? Рассмотрим задачу классификации с двумя классами. Пусть пространство признаков двумерное. Объекты каждого класса имеют нормальное распределение с математическим ожиданием  $(-1, -1)$  и  $(1, 1)$  соответственно и единичной матрицей ковариации каждый. Априорные вероятности классов равны  $1/2$ .

- 1) Вычислите коэффициент корреляции для переменных  $x_1, x_2$ .
- 2) Найти байесов классификатор и вычислить байесову ошибку для усеченной задачи, рассматривая только одну переменную  $x_1$ .
- 3) Найти байесов классификатор и вычислить байесову ошибку для исходной задачи.
- 4) Приводит ли использование второй переменной к уменьшению ошибки?

**28.** Рассмотрим задачу классификации с двумя классами 0 и 1. Пусть пространство признаков двумерное. Объекты каждого класса имеют нормальное распределение с математическим ожиданием  $(0, 0)$  и  $(1, 1)$  соответственно и матрицей ковариации

$$\Sigma = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Априорные вероятности классов равны  $\Pr\{Y = 0\} = 1/3$  и  $\Pr\{Y = 1\} = 2/3$ .

- 1) Найти уравнение разделяющей поверхности байесова классификатора.
- 2) Найти собственное разложение матрицы  $\Sigma$ .

