

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Тольяттинский государственный университет»

Б1.О.07  
(индекс дисциплины)

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

Оценка, оптимизация и управление жизненным циклом моделей машинного обучения

(наименование дисциплины)

по направлению подготовки  
01.04.02 Прикладная математика и информатика

направленность (профиль)  
Искусственный интеллект и машинное обучение в беспилотных мобильных системах и комплексах

Форма обучения: очная

Год набора: 2026

Общая трудоемкость: 6 ЗЕ

**Распределение часов дисциплины по семестрам**

| Семестр                                      | 4          | Итого      |
|--|------------|------------|
| Форма контроля                               | Экзамен    |            |
| Вид занятий                                  |            |            |
| Лекции                                       | 12         | 12         |
| Лабораторные                                 |            |            |
| Практические                                 | 24         | 24         |
| Руководство: курсовые работы (проекты) / РГР |            |            |
| Промежуточная аттестация                     | 0,35       | 0,35       |
| Контактная работа                            | 36,35      | 36,35      |
| Самостоятельная работа                       | 144        | 144        |
| Контроль                                     | 35,65      | 35,65      |
| <b>Итого</b>                                 | <b>216</b> | <b>216</b> |

Рабочую программу составил(и):  
Доцент института цифровых технологий, к.т.н., доцент Сосина Н.А.  

---

(должность, ученое звание, степень, Фамилия И.О.)

---

(должность, ученое звание, степень, Фамилия И.О.)

Рецензирование рабочей программы дисциплины:



Отсутствует



Рецензент

---

(должность, ученое звание, степень, Фамилия И.О.)

Рабочая программа дисциплины составлена на основании ФГОС ВО и учебного плана  
направления подготовки 01.04.02 Прикладная математика и информатика  

---

**Срок действия рабочей программы дисциплины до «31» августа 2028г.**

**УТВЕРЖДЕНО**

На заседании института цифровых технологий  

---

(протокол заседания № 1 от 5 сентября 2025г.)

## 1. Цель освоения дисциплины

Цель освоения дисциплины – сформировать у специалиста компетенции, необходимые для перехода от создания «сырой» модели к ее эксплуатации в промышленной среде как к надежному, управляемому и непрерывно развивающемуся бизнес-активу.

## 2. Место дисциплины в структуре ОПОП ВО

Дисциплины и практики, на освоении которых базируется данная дисциплина: Машинное обучение и нейронные сети, Математическое и компьютерное моделирование динамических систем, Программные платформы и инструменты разработки интеллектуальных систем, Методы обработки и анализа данных для систем искусственного интеллекта

Дисциплины и практики, для которых освоение данной дисциплины необходимо как предшествующее: Производственная практика (научно-исследовательская работа) 3-4., магистерская диссертация.

## 3. Планируемые результаты обучения

| Формируемые и контролируемые компетенции<br>(код и наименование)                                       | Индикаторы достижения компетенций<br>(код и наименование)   | Планируемые результаты обучения   |
|--|---|---|
| ОПК -2. Способен совершенствовать и реализовывать новые математические методы решения прикладных задач | ОПК-2.1. Знает современные и перспективные математические методы, применяемые для решения прикладных задач анализа данных и моделирования | Знать: современные и перспективные математические методы, применяемые для решения прикладных задач анализа данных и моделирования<br>Уметь: современные и перспективные математические методы, применяемые для решения прикладных задач анализа данных и моделирования<br>Владеть: навыками современных и перспективных математических методов, применяемых для решения прикладных задач анализа данных и моделирования |
|  | ОПК-2.2. Умеет адаптировать и развивать математические методы с учетом специфики прикладных задач профессиональной деятельности           | Знать: математические методы с учетом специфики прикладных задач профессиональной деятельности<br>Уметь: адаптировать и развивать математические методы с учетом специфики прикладных задач профессиональной деятельности.<br>Владеть: навыками адаптации математических методов с учетом специфики прикладных задач профессиональной деятельности  |

| <b>Формируемые и контролируемые компетенции</b><br>(код и наименование) | <b>Индикаторы достижения компетенций</b><br>(код и наименование)  | <b>Планируемые результаты обучения</b>  |
|---|---|---|
|   | ОПК-2.3. Владеет навыками реализации и экспериментальной проверки новых математических методов решения прикладных задач | Знать: новые математические методы реализации и экспериментальной проверки для решения прикладных задач<br>Уметь: реализовывать и осуществлять проверку новых математических методов для решения прикладных задач.<br>Владеть: навыками реализации и экспериментальной проверки новых математических методов решения прикладных задач |

#### 4. Структура и содержание дисциплины Оценка, оптимизация и управление жизненным циклом моделей машинного обучения

| Модуль (раздел)  | Вид учебной работы | Наименование тем занятий (учебной работы)  | Семестр | Объем, ч. | Баллы | Интерактив, ч. | Формы текущего контроля (наименование оценочного средства) |
|--|--------------------|--|---------|-----------|-------|----------------|--|
| <b>Модуль1. Оптимизация параметров</b>                 | Лек                | Центральны методы оптимизации: градиентный спуск, стохастический градиентный спуск (SGD), методы второго порядка, учитывающие гессиан для быстрой сходимости. Выпуклая оптимизация минимизирует функцию потерь $L(\theta) = \frac{1}{n} \sum_{i=1}^n \ell(y_i, \hat{y}_i(\theta))$ , где $\theta$ — параметры модели. Регуляризация (L1/L2) и гиперпараметрическая настройка решают ограниченные задачи.   | 4       | 3         |       |                | Выполнение задачи №1                                       |
|  | Пр                 | Решение задачи оптимизации. Нахождение параметров модели, минимизирующих функцию потерь. Решение гладких задач методом градиентного спуска. Методы решения задач оптимизации с ограничениями. .  | 4       | 6         |       |                |  |
|  | Ср                 | Метод градиентного спуска и его вариации. Математическое программирование  | 4       | 36        |       |                |  |
| <b>Модуль2. Инжиниринг признаков и хранение данных</b> | Лек                | Проверка статистических гипотез и анализ данных. Статистический анализ: Расчет корреляции (Пирсона, Спирмена), выявление выбросов, проверка распределений (с помощью критерия Колмогорова-Смирнова) для определения методов нормализации. Преобразование данных: Применение линейной алгебры для масштабирования (StandardScaler, MinMaxScaler) или понижения размерности (PCA — метод главных компонент, SVD — сингулярное разложение). Вычисление признаков: Агрегация временных рядов (средние, скользящие окна), что требует понимания основ анализа сигналов. | 4       | 3         | 20    |                | Выполнение задачи №2                                       |
|  | Пр                 | Проверка статистических гипотез и анализ данных. Статистический анализ. Преобразование данных. Вычисление признаков.   | 4       | 6         |       |                |  |

| Модуль (раздел)                                     | Вид учебной работы | Наименование тем занятий (учебной работы)  | Семестр | Объем, ч.  | Баллы      | Интерактив, ч. | Формы текущего контроля (наименование оценочного средства) |
|---|--------------------|--|---------|------------|------------|----------------|--|
|   | Ср                 | Изучение и анализ методов, позволяющих измерить на сколько хорошо модель решает задачу и насколько устойчиво ее качество на новых данных<br>Проработка лекций, изучение литературы, выполнение задания №1  | 4       | 36         |            |                |  |
| <b>Модуль3. Управление жизненным циклом</b>         | Лек                | Оптимизация ресурсов и инфраструктуры  | 4       | 3          | 20         |                | Выполнение задачи №3                                       |
|   | Пр                 | Теория массового обслуживания (Queuing Theory) для расчета задержек при параллельной обработке запросов к модели   | 4       | 6          |            |                |  |
|   | Ср                 | Проработка лекций, изучение литературы, выполнение задания №2  | 4       | 36         |            |                |  |
| <b>Модуль4. Мониторинг и управление изменениями</b> | Лек                | Количественная оценка изменений. Статистическое распределение входных признаков. Сравнение распределений с помощью расстояния Вассерштейна, дивергенции Кульбака-Лейблера или статистических тестов (тест Колмогорова-Смирнова).<br>Концептуальный дрейф: изменение зависимости между входными данными и целевой переменной. | 4       | 3          | 40         |                | Презентация кейс-задачи                                    |
|   | Пр                 | Анализ ошибок модели во времени, расчет скользящего среднего ошибки (EWMA — экспоненциально взвешенное скользящее среднее) и пороговых значений на основе сигм (правило трёх сигм).  | 4       | 6          |            |                |  |
|   | Ср                 | Проработка лекций, изучение литературы, подготовка презентации   | 4       | 36         |            |                |  |
|   | ПА                 |  | 4       | 0,35       |            | -              |  |
|   | Контроль           |  | 4       | 8,65       |            |                |  |
| Итого:  |                    |  |         | <b>216</b> | <b>100</b> |                |  |

## **5. Образовательные технологии**

В рамках учебного курса предусмотрены следующие образовательные технологии:

- технология традиционного обучения: лекции и практические работы, самостоятельная работа;
- технология проектного обучения: решение кейс-задачи с последующей презентацией.

## **6. Методические указания по освоению дисциплины**

### **6.1. Рекомендации по подготовке к лекционным занятиям**

Изучение дисциплины требует систематического и последовательного накопления знаний, следовательно, пропуски отдельных тем не позволяют глубоко освоить предмет.

В ходе лекционных следует обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации. Задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций.

Студент может дополнить список использованной литературы современными источниками, не представленными в списке рекомендованной литературы, и в дальнейшем использовать собственные подготовленные учебные материалы при написании курсовых и выпускных квалификационных работ.

### **6.2. Рекомендации по подготовке к практическим занятиям**

Для того чтобы практические занятия приносили максимальную пользу, необходимо помнить, что решение задач проводятся по рассмотренному на лекциях материалу и связаны с детальным разбором отдельных вопросов лекционного курса.

Следует помнить, что выполнение каждой работы должно доводиться до окончательного логического ответа, которого требует условие, и по возможности с выводом. Полученный ответ следует проверить способами, вытекающими из существа данной задачи. Полезно также (если возможно) решать несколькими способами и сравнить полученные результаты. Решение задач данного типа нужно продолжать до приобретения твердых навыков в их решении.

### **6.3. Рекомендации по подготовке к экзамену**

Подготовка к экзамену способствует закреплению, углублению и обобщению знаний, получаемых, в процессе обучения, а также применению их к решению практических задач. Готовясь к экзамену, студент ликвидирует имеющиеся пробелы в знаниях, углубляет, систематизирует и упорядочивает свои знания. На экзамене студент демонстрирует то, что он приобрел в процессе обучения по конкретной учебной дисциплине.

Необходимо ориентировать студентов на систематическую подготовку к занятиям в течение семестра, что позволит использовать время экзаменационной сессии для систематизации знаний.

## 7. Оценочные средства

### 7.1. Паспорт оценочных средств

| Семестр | Код контролируемой компетенции<br>(или ее части) | Наименование<br>оценочного средства                               |
|---------|--|---|
| 4       | ОПК-2  | Отчеты по решению задач №1-№3,<br>Презентация решения кейс-задачи |

### 7.2. Типовые задания или иные материалы, необходимые для текущего контроля

Цель выполнения задания - сформировать способность обоснованно разрабатывать, оценивать, оптимизировать и сопровождать модели машинного обучения на всех этапах их жизненного цикла, опираясь на строгие математические методы и обеспечивая надежность, интерпретируемость и эффективность решений.

#### 7.2.1 Типовой пример Кейс -задачи

Рассматривается ML-модель  $M$  с множеством версий  $\{M_t\}_{t=1}^T$ , где  $t$  — дискретные моменты времени (например, дни/месяцы).

Переменные

$x_t \in \{0,1\}$ : индикатор использования модели на шаге  $t$  ( $x_t = 1$  — модель в эксплуатации,  $0$  — отключена).

$\mathbf{p}_t \in \mathbb{R}^d$ : гиперпараметры/параметры модели на шаге  $t$  (если оптимизируются).

$u_t \in \{0,1\}$ : индикатор проведения переобучения/апдейта на шаге  $t$ .

Функция цели

Минимизировать средневзвешенные затраты:

$$\min_{\{x_t, u_t, \mathbf{p}_t\}} \sum_{t=1}^T [c_t^{\text{inf}} x_t + c_t^{\text{train}} u_t + c_t^{\text{data}} x_t + \lambda(\varepsilon - Q(t; \mathbf{p}_t))^+],$$

где:

$c_t^{\text{inf}}$  — стоимость инференса на шаге  $t$  (вычисление, пропускная способность, облачные ресурсы);

$c_t^{\text{train}}$  — стоимость переобучения/перерасчёта;

$c_t^{\text{data}}$  — стоимость хранения/обработки данных;

$Q(t; \mathbf{p}_t)$  — метрика качества модели в момент  $t$  (например, F1, AUC, MAE);

$\varepsilon$  — минимально допустимое качество;

$(\cdot)^+ = \max(\cdot, 0)$  штрафует отклонение качества ниже порога;

$\lambda \geq 0$  — вес штрафа за падение качества.

Ограничения

Баланс жизни модели:

$$x_t \leq \sum_{s=t-n}^t u_s, \forall t \in [n+1, T],$$

где  $n$  — минимальная «свежесть» модели: модель может быть включена, только если в последние  $n$  шагов была хотя бы одна перекачка/переобучение.

Минимальное качество:

$$Q(t; \mathbf{p}_t) \geq \varepsilon, \forall t \in [1, T]: x_t = 1.$$

Ресурсные ограничения:

$$\sum_{t=1}^T (c_t^{\text{inf}} x_t + c_t^{\text{train}} u_t) \leq C_{\text{budget}},$$

где  $C_{\text{budget}}$  — общий бюджет жизненного цикла.

Дискретность обновлений (если не хочется «переобучать каждый день»):

$$\sum_{t=1}^T u_t \leq U_{\text{max}},$$

где  $U_{\text{max}}$  — максимальное число разрешённых обновлений за весь цикл.

### Задача участников кейса

Минимизировать суммарные затраты за жизненный цикл при одновременном удовлетворении ограничений на качество и надёжность.

#### №1. Прогнозирование качества:

обучаются дополнительные модели (ML или регрессия) для прогноза  $F1_t$ , как функции от времени  $t$  и стиля дрейфа данных (например, по Kullback–Leibler или KS-статистике между новым и старым распределением).

#### № 2. Дискретная оптимизация:

задача может решаться как динамическое программирование (если время дискретно и не слишком большое) или методом ветвей и границ для булевых переменных  $x_t, u_t$ ; для больших  $T$  допускается релаксация  $x_t, u_t \in [0, 1]$  и решение как линейной или смешанной целочисленной задачи.

#### №3. Итеративная адаптация:

после развертывания собираются реальные метрики мониторинга; модель перестраивается периодически: старые решения корректируются по новым данным, это соответствует онлайн-оптимизации или реинфорс-адаптации стратегии обновлений.

#### Формат кейса для учебной задачи

Дано:

$T = 30$  дней,  $c^{\text{inf}} = 10$ ,  $c^{\text{train}} = 100$ ,  $\lambda = 1000$ ,  $\varepsilon = 0,7$ .

$F1_t$  без переобучения:  $F1_t = 0,9 - 0,01t$ .

Требуется:

выбрать моменты  $t$ , в которые устанавливать  $u_t = 1$ , чтобы минимизировать целевую функцию при  $x_t = 1$  на всех шагах.

Пояснение: записать  $F1_t$  после каждого переобучения как ступенчатую функцию (после  $u_t = 1$  качество сбрасывается к 0,9 и снова падает); перебрать небольшое число возможных расписаний обновлений или решить, как задачу булевой оптимизации с помощью линейного решателя.

### Форма отчета по задачам №1-№3

В отчет по задаче должны быть включены следующие пункты:

- краткие теоретические сведения;
- описание хода выполнения работы;
- результаты выполненной работы.

| Формы<br>текущего контроля | Критерии и нормы оценки  |
|----------------------------|--|
| Отчет по задачам №1-№3     | <p>«отлично (20баллов)» – задание выполнено в полном объеме без замечаний</p> <p>«хорошо (12-19 баллов)» – задание выполнено в полном объеме, присутствуют замечания по выполнению задания</p> <p>«удовлетворительно (8-11 баллов)» – задание выполнено не в полном объеме, присутствуют несущественные замечания</p> <p>«неудовлетворительно (менее 7 баллов)» – задание выполнено не в полном объеме, присутствуют замечания по выполнению задания</p> |

### Форма отчета по выполнению кейс-задачи

Презентация кейс-задачи (заданий №1-№3), включает:

1. прогноз качества модели
2. метод решения,
3. описание стратегии обновлений

Оценка проводится по следующим критериям:

1. Наличие всей существенной информации по работе
2. Точность и полнота предоставляемых сведений
3. Непротиворечивость приводимой информации
4. Правильность интерпретаций и выводов, которые сделаны по результатам работы
5. Степень достижения студентом поставленной цели
6. Обоснованность применяемого решения
7. Грамотность (содержательная) используемых формулировок

### Процедура оценивания

#### Оценка «отлично (40 баллов)» (Отлично)

**Содержание:** Тема раскрыта полностью, информация структурирована, приведены актуальные данные и примеры. Цель работы достигнута.

**Оформление:** Идеальный баланс текста и картинок. Единый стиль, читаемый шрифт, качественные изображения. Отсутствуют ошибки и опечатки.

**Защита:** Выступающий свободно владеет материалом, не читает с листа, уверенно отвечает на вопросы, уложился в регламент.

#### Оценка «хорошо (29-39 баллов)» (Хорошо)

**Содержание:** Тема раскрыта, но есть незначительные неточности (фактические или логические) либо не хватает глубины анализа.

**Оформление:** Дизайн хороший, но есть мелкие замечания: на одном слайде слишком много текста, неудачно подобран цвет фона или шрифта. Присутствуют 1-2 грамматические ошибки.

**Защита:** Выступающий знает тему, но немного «плавает» в деталях, иногда подглядывает в текст. На вопросы отвечает, но неуверенно.

#### Оценка «удовлетворительно (19-28 баллов)» (Удовлетворительно)

**Содержание:** Тема раскрыта поверхностно, информация взята из одного источника, нет собственных выводов. Логика повествования нарушена.

**Оформление:** Слайды перегружены текстом, используется много разных цветов и шрифтов, картинки не по теме или плохого качества. Есть орфографические ошибки.

**Защита:** Выступление сводится к чтению текста со слайдов. Выступающий не может ответить на вопросы по теме.

#### Оценка «неудовлетворительно (менее 18 баллов)» (Неудовлетворительно)

**Содержание:** Тема не раскрыта, информация не соответствует заданию или отсутствует.

**Оформление:** Полное отсутствие дизайна (черно-белый текст, «рванные» слайды), большое количество ошибок, нечитаемый текст.

**Защита:** Выступающий не понимает, о чем говорит, не может ответить на вопросы или отказывается отвечать.

### 7.2.2 Типовые тестовые задания

#### Модуль 1. Теоретические основы оптимизации

**Вопрос 1.** Почему на практике в алгоритме стохастического градиентного спуска (SGD) рекомендуется использовать **маленький** темп обучения (learning rate)?

- А) Это гарантирует нахождение глобального минимума за одну эпоху
- В) Это позволяет алгоритму работать только с категориальными признаками
- С) Это позволяет избежать "выстреливания" (расходимости) и достичь более точной сходимости к минимуму
- D) Это увеличивает скорость вычислений за счет векторизации

**Правильный ответ: С.** Слишком большой learning rate может привести к тому, что градиентный спуск начнет "перескакивать" через минимум, и значение функции потерь будет колебаться или расти. Маленький шаг обеспечивает более стабильную и точную сходимость.

**Вопрос 2.** Допустим, мы минимизируем функцию потерь, которая является суммой ошибки модели и регуляризационного члена:  $C(w) = \text{DataLoss}(w) + \lambda * \text{Regularization}(w)$ . Изначально  $\lambda = 2$ , и глобальный минимум равен 12. Если мы изменим  $\lambda$  на 20 и снова найдем глобальный минимум, что произойдет со значением функции потерь  $C(w)$ ?

- А) Значение увеличится
- В) Значение уменьшится
- С) Значение останется неизменным
- D) Недостаточно информации

**Правильный ответ: D.** Мы штрафуем модель сильнее ( $\lambda$  больше), следовательно, оптимизатор выберет другие веса  $w$ . Однако без знания конкретного вида  $\text{DataLoss}$  и  $\text{Regularization}$  нельзя однозначно сказать, изменится ли итоговая сумма  $C(w)$  (общее значение функции) и в какую сторону

#### Модуль 2. Инжиниринг признаков и хранение данных

**Вопрос 3.** В задаче бинарной классификации модель правильно идентифицирует 95% спам-писем, но при этом ошибочно помечает множество легитимных писем как спам. Какая метрика, скорее всего, будет низкой?

- А) Полнота (Recall) для класса "спам"
- В) Точность (Precision) для класса "спам"
- С) Specificity
- D) Accuracy

**Правильный ответ: В.** Точность (Precision) измеряет долю действительно спам-писем среди всех писем, помеченных моделью как спам. Большое количество ложноположительных срабатываний (False Positives) приводит к снижению точности.

**Вопрос 4.** Что представляет собой F1-мера (F1-Score)?

- А) Среднее арифметическое точности и полноты
- В) Среднее гармоническое точности и полноты
- С) Среднее геометрическое точности и полноты
- D) Доля правильных ответов модели

**Правильный ответ: В.** F1-мера является средним гармоническим точности (Precision) и полноты (Recall) и используется, когда необходимо сбалансировать эти две метрики, особенно на несбалансированных данных .

**Вопрос 5.** В каких случаях метрика «Ассигасу» (доля правильных ответов) перестает быть информативной и может вводить в заблуждение?

- А) При очень большом размере выборки
- В) При сильном дисбалансе классов в выборке
- С) При использовании ансамблевых методов
- D) При стандартизации признаков

**Правильный ответ: В.** На несбалансированных данных (например, 95% объектов — класс "0", 5% — класс "1") модель может предсказывать только класс "0" и показывать точность 95%, будучи при этом совершенно бесполезной для выявления объектов класса "1" .

**Вопрос 6.** Какая метрика позволяет оценить, насколько модель лучше случайного угадывания, и часто отображается в виде графика отношения истинно-положительных к ложноположительным срабатываниям?

- А) Lift-диаграмма (Lift Chart)
- В) ROC-кривая (Receiver Operating Characteristic curve)
- С) Матрица ошибок (Confusion Matrix)
- D) График остатков

**Правильный ответ: В.** ROC-кривая показывает компромисс между True Positive Rate и False Positive Rate, а площадь под ней (AUC-ROC) количественно измеряет способность модели различать классы.

### **Модуль3. Управление жизненным циклом**

**Вопрос 7.** Что происходит, когда модель чрезмерно сложна (например, имеет слишком много параметров или высокую степень полинома), хорошо работает на обучающей выборке, но показывает низкую эффективность на валидационной или тестовой выборке?

- А) Смещение (Bias)
- В) Переобучение (Overfitting)
- С) Недообучение (Underfitting)
- D) Дисперсия (Variance)

**Правильный ответ: В.** Это классическое определение переобучения — модель выучивает шумы и случайности в обучающих данных вместо общих закономерностей.

**Вопрос 8.** Если модель находится в состоянии переобучения (overfitting), какие действия могут помочь улучшить её качество на тестовых данных? (Выберите все верные варианты)

- А) Уменьшить регуляризацию ( $\lambda$ )
- В) Увеличить регуляризацию ( $\lambda$ )
- С) Уменьшить сложность модели (например, степень полинома)
- D) Добавить новые признаки

**Правильный ответ: В, С.** Увеличение регуляризации штрафует сложность модели, а уменьшение количества признаков или порядка полинома напрямую снижает сложность, помогая бороться с переобучением.

**Вопрос 9.** Разложение ошибки на смещение и дисперсию описывается формулой:  $\text{Ошибка} = \text{Смещение}^2 + \text{Дисперсия} + \text{Неустраняемая ошибка}$ . Какая из моделей, как правило, характеризуется **высоким смещением (high bias)** и **низкой дисперсией (low variance)**

- А) Линейная регрессия на нелинейных данных
- В) Глубокое дерево решений
- С) Метод k-ближайших соседей с  $k=1$
- D) Случайный лес из 1000 деревьев

**Правильный ответ: А.** Простая модель (линейная) на сложных данных неспособна уловить закономерности, что приводит к систематической ошибке (высокое смещение). При этом её результат стабилен при смене выборки (низкая дисперсия). Глубокое дерево, наоборот, имеет низкое смещение, но высокую дисперсию.

### **Модуль 3. Жизненный цикл: от данных до мониторинга**

**Вопрос 10.** Какие данные наиболее подходят для **обучения** прогностических моделей (building predictive models)?

- А) Валидационные (Validation)
- В) Тестовые (Test)
- С) Обучающие (Training)
- D) Исходные (Original)

**Правильный ответ: С.** Обучающий набор данных напрямую используется алгоритмами машинного обучения для подбора коэффициентов (весов) модели.

**Вопрос 11.** Какое действие является наиболее подходящим **финальным шагом** после успешной оценки модели и достижения желаемого качества?

- А) Сбор большего количества данных для дальнейшего улучшения
- В) Развертывание модели в реальном приложении (деплой)
- С) Запуск проекта с нуля с новым алгоритмом
- D) Построение новой модели для той же задачи

**Правильный ответ: В.** Смысл разработки модели — её практическое применение. После того как модель проверена и удовлетворяет требованиям, её жизненный цикл переходит в стадию эксплуатации (deployment).

**Вопрос 12.** Если в модели линейной регрессии присутствуют **номинальные (категориальные) признаки** (например, цвет: красный, зеленый, синий), что необходимо сделать перед обучением?

- А) Ничего, линейная регрессия работает с ними напрямую
- В) Удалить их, так как они бесполезны
- С) Преобразовать в числовой формат с помощью one-hot-кодирования (dummy-кодирование)
- D) Отсортировать их по алфавиту

**Правильный ответ: С.** Линейные модели работают только с числовыми данными. Категориальные признаки необходимо преобразовывать, создавая бинарные фиктивные переменные (dummy variables) для каждой категории.

**Вопрос 13.** Метод главных компонент (PCA) используется на этапе Feature Engineering. Какова его основная цель в контексте жизненного цикла модели?

- А) Увеличить точность модели за счет создания новых признаков
- В) Снизить размерность данных, уменьшив вычислительную сложность и риск переобучения
- С) Заменить пропущенные значения средними
- D) Визуализировать работу градиентного спуска

**Правильный ответ: В.** PCA уменьшает количество признаков, сохраняя большую часть дисперсии (информации) данных, что упрощает модель и может помочь в борьбе с "проклятием размерности".

#### **Модуль 4. Мониторинг и управление изменениями**

**Вопрос 14.** Какова основная цель перекрестной проверки (кросс-валидации), например, k-fold cross-validation?

- А) Обучить модель максимально быстро
- В) Оценить обобщающую способность модели и подобрать гиперпараметры
- С) Увеличить размер обучающей выборки
- D) Очистить данные от выбросов

**Правильный ответ: В.** Кросс-валидация используется для более стабильной и надежной оценки того, как модель будет работать на новых, невиданных ранее данных, а также для выбора гиперпараметров без переобучения на тестовой выборке.

**Вопрос 15.** В чем главный недостаток метода отложенной выборки (Holdout) по сравнению с k-блочной кросс-валидацией?

- А) Holdout не может использоваться для задач классификации
- В) Holdout требует значительно больше вычислительных ресурсов
- С) Holdout чувствителен к способу разбиения данных на обучающий и тестовый наборы
- D) Holdout не позволяет обучать линейные модели

**Правильный ответ: С.** Результат оценки качества при использовании holdout-метода может сильно варьироваться в зависимости от того, какие именно объекты попали в

тестовую выборку. Кросс-валидация усредняет результат по нескольким разбиениям, снижая эту зависимость.

**Вопрос 16.** Для чего нужен **валидационный** набор данных (validation set)?

- А) Для окончательной проверки модели перед внедрением
- В) Для настройки гиперпараметров и выбора модели
- С) Для непосредственного обучения весов модели
- D) Для увеличения разнообразия данных

**Правильный ответ: В.** Валидационный набор используется в процессе разработки для настройки параметров и выбора лучшей модели, в то время как тестовый набор используется только один раз для финальной оценки качества

| Формы текущего контроля | Критерии и нормы оценки   |
|-------------------------|---|
| тест                    | 10 баллов – на все вопросы получен правильный ответ<br>0-9 баллов – количеству правильных ответов |

### Темы письменных работ

Учебным планом не предусмотрено.

## 7.3. Оценочные средства для промежуточной аттестации по итогам освоения дисциплины

### 7.3.1. Вопросы к промежуточной аттестации

#### Семестр 4

#### Вопросы к экзамену

#### Теория вероятностей и статистика

1. Сформулируйте и докажите формулу полной вероятности и формула Байеса.
2. Понятие случайной величины, функции распределения и плотности.
3. Математическое ожидание, дисперсия, ковариация: свойства и интерпретации.
4. Закон больших чисел и центральная предельная теорема: роль в машинном обучении.
5. Оценки параметров: смещённость, состоятельность, эффективность.
6. Метод максимального правдоподобия (MLE) и метод моментов.
7. Доверительные интервалы и проверка статистических гипотез.
8. Переобучение как статистическое явление (variance vs bias).

#### 2. Линейная алгебра

9. Векторные пространства, базис, размерность.
8. Собственные значения и собственные векторы: применение в PCA.
9. Разложения матриц: LU, QR, SVD.
10. Условность матриц и численная устойчивость.
11. Нормы векторов и матриц (L1, L2, спектральная).
12. Псевдообратная матрица и решение переопределённых систем.

13.Геометрическая интерпретация линейных моделей.

### **3. Оптимизация**

14. Постановка задачи оптимизации в машинном обучении.

15. Необходимые и достаточные условия экстремума.

16. Градиент и гессиан: интерпретация.

17. Метод градиентного спуска и его сходимость.

18. Стохастический градиентный спуск (SGD).

19. Методы ускорения (Momentum, Adam).

20. Ограниченная оптимизация и множители Лагранжа.

21.Выпуклость функций и её значение для обучения моделей.

### **4. Функции потерь и метрики качества**

22. Свойства функции потерь: выпуклость, гладкость.

23. Квадратичная функция потерь и её связь с нормальным распределением.

24. Логистическая функция потерь и кросс-энтропия.

25. Метрики классификации: Accuracy, Precision, Recall, F1-score.

26. ROC-кривая и AUC: математическая интерпретация.

27.Метрики регрессии: MSE, MAE,  $R^2$ .

28. Смещение и разброс оценок.

### **5. Регуляризация и обобщающая способность**

29.Переобучение и недообучение: математическая формализация.

30.L1- и L2-регуляризация.

31.Связь регуляризации с априорными распределениями (байесовский подход).

32.Кросс-валидация: виды и математическое обоснование.

33.VC-размерность и теория обучения.

34.Обобщающая ошибка и её оценки.

### **6. Байесовские методы**

35.Апостериорное распределение и априорные предположения.

37.Байесовская линейная регрессия.

38.МАР-оценка и её связь с регуляризацией.

39.Вариационный вывод и EM-алгоритм.

40.Марковские цепи и MCMC.

### **7. Управление жизненным циклом моделей (MLOps)**

41.Формализация жизненного цикла ML-модели.

42.Метрики деградации модели (data drift, concept drift).

43.Онлайн- и офлайн-оценка качества.

44. А/В тестирование моделей: статистическая основа.

45. Мониторинг моделей и статистические критерии отклонения.

46. Версионирование моделей и воспроизводимость.

### **8. Теория информации**

47.Энтропия энтропия Шеннона.

48. Дивергенция Кульбака–Лейблера.

49. Кросс-энтропия и её связь с функцией потерь.

50. Взаимная информация.

### **9. Численные методы и вычислительная устойчивость**

51. Ошибки округления и накопление ошибок.

52. Условие задачи и устойчивость алгоритмов.

53. Итерационные методы решения систем.

54. Аппроксимация функций.

### **10. Продвинутое темы**

55. Теория оптимального управления в ML.

56. Многокритериальная оптимизация.

57. Стохастические процессы в обучении моделей.

58. Дифференцируемое программирование.
59. Автоматическое дифференцирование.
60. Статистическая валидация моделей.
61. Оценка устойчивости (Robustness).
62. Стресс-тестирование модели.
63. Обнаружение и mitigation смещений (bias) в данных и алгоритмах.
64. Интерпретируемые модели (Linear, Tree-based)
65. Роль объяснимости в процессе приемки модели.
66. Оптимизация признакового пространства.
67. Понятие Feature Store (онлайн/офлайн), управление версиями признаков.
68. Оптимизация гиперпараметров (HPO).
69. Масштабирование HPO в распределенных средах.
70. Сжатие моделей (Model Optimization)

### 7.3.2. Критерии и нормы оценки

| Семестр | Форма проведения промежуточной аттестации | Критерии и нормы оценки |                         |
|---------|---|-------------------------|-------------------------|
| 3       | Экзамен (по накопительному рейтингу)      | «отлично»               | рейтинговый балл 85-100 |
|         |   | «хорошо»                | рейтинговый балл 70-84  |
|         |   | «удовлетворительно»     | рейтинговый балл 55-69  |
|         |   | «неудовлетворительно»   | рейтинговый балл 0-54   |

## 8. Учебно-методическое и информационное обеспечение дисциплины

### 8.1. Обязательная литература

| № п/п | Авторы, составители | Заглавие (заголовок)   | Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.) | Год издания | Количество в научной библиотеке / Наименование ЭБС <sup>1</sup> |
|-------|---------------------|--|---|-------------|---|
| 1.    | Бовырин А.В.        | Разработка мультимедийных приложений с использованием библиотек OpenCV и IPP: учебное пособие / А. В. Бовырин, П. Н. Дружков, В. Л. Ерухимов [и др.]. – 4-е изд. (эл.). – Москва: ИНТУИТ: Ай Пи Ар Медиа, 2024. – 515 с. – URL: <a href="https://www.iprbookshop.ru/133976.html">https://www.iprbookshop.ru/133976.html</a> (дата обращения: 02.11.2023). – Режим доступа: Цифровой образовательный ресурс IPR SMART. – ISBN 978-5-4497-2481-6. – Текст : электронный. URL: <a href="https://www.iprbookshop.ru/133976.html">https://www.iprbookshop.ru/133976.html</a><br>Сигла хранения:эбс-IPRbooks | учебное пособие   | 2024        | PRbooks   |
| 2.    | Гвоздева В. А.      | Интеллектуальные технологии в беспилотных системах : учебник / В. А. Гвоздева. – 2-е изд., доп. – Москва : ИНФРА-М, 2026. – 197 с. : ил. – (Высшее образование). – URL: <a href="https://znanium.ru/catalog/product/2217548">https://znanium.ru/catalog/product/2217548</a> (дата обращения: 31.10.2025). – Режим доступа: Электронно-библиотечная система– ISBN 978-5-16-110783-6 . –   | учебник   | 2026        | "ZNANIUM".  |

<sup>1</sup> Указывается количество экз. для печатных изданий, для электронных изданий – наименование ЭБС.

| №<br>п/п | Авторы, составители   | Заглавие (заголовок)   | Тип (учебник, учебное<br>пособие, учебно-<br>методическое пособие,<br>практикум, др.) | Год издания | Количество в<br>научной<br>библиотеке /<br>Наименование<br>ЭБС <sup>1</sup> |
|----------|-----------------------|--|---|-------------|---|
|          |                       | Текст : электронный.<br>URL: <a href="https://znanium.ru/catalog/product/2217548">https://znanium.ru/catalog/product/2217548</a>   |   |             |   |
| 3.       | <b>Богатырев С.Ю</b>  | Машинное обучение в финансах : учебник для магистратуры / С. Ю. Богатырев, А. А. Помулев, А. В. Затевахина [и др.] ; под общей редакцией С. Ю. Богатырева. – Москва : Прометей, 2024. – 224 с. : ил. – URL: <a href="https://znanium.ru/catalog/product/2144368">https://znanium.ru/catalog/product/2144368</a> (дата обращения: 04.09.2025). – Режим доступа: Электронно-библиотечная система "ZNANIUM". – ISBN 978-5-00172-572-5 . – Текст : электронный. URL: <a href="https://znanium.ru/catalog/product/2144368">https://znanium.ru/catalog/product/2144368</a> | учебное пособие   | 2024        | "ZNANIUM".  |
| 4.       | <b>Кудинов, Ю. И.</b> | Основы современной информатики: учебное пособие / Ю. И. Кудинов, Ф. Ф. Пащенко. – Изд. 6-е, стер. – Санкт-Петербург: Лань, 2024. – 256 с.: ил. – (Высшее образование). – URL: <a href="https://e.lanbook.com/book/392393">https://e.lanbook.com/book/392393</a> (дата обращения: 26.02.2024). – Режим доступа: Электронно-библиотечная система "Лань". – ISBN 978-5-507-47572-8. – Текст: электронный. URL: <a href="https://e.lanbook.com/book/392393">https://e.lanbook.com/book/392393</a>  | учебное пособие   | 2024        | "Лань".   |

| <b>№<br/>п/п</b> | <b>Авторы, составители</b> | <b>Заглавие (заголовок)</b>  | <b>Тип (учебник, учебное<br/>пособие, учебно-<br/>методическое пособие,<br/>практикум, др.)</b> | <b>Год издания</b> | <b>Количество в<br/>научной<br/>библиотеке /<br/>Наименование<br/>ЭБС<sup>1</sup></b> |
|------------------|----------------------------|--|---|--------------------|---|
|                  | <b>Майтак Р. В.</b>        | Python, Django, Data Science : учебное пособие / Р. В. Майтак, А. В. Протодяконов, П. А. Пылов. – Москва ; Вологда : Инфра-Инженерия, 2025. – 516 с. : ил. – URL: <a href="https://znanium.ru/catalog/product/2225323">https://znanium.ru/catalog/product/2225323</a> (дата обращения: 05.09.2025). – Режим доступа: Электронно-библиотечная система "ZNANIUM". – ISBN 978-5-9729-1006-9 . – Текст : электронный. URL: <a href="https://znanium.ru/catalog/product/2225323">https://znanium.ru/catalog/product/2225323</a> | учебное пособие   | 2025               | "ZNANIUM".  |
|                  | <b>Трушков, А. С.</b>      | Исследование операций [Электронный ресурс]: учеб. пособие. Т. 1. Линейное программирование / А. С. Трушков. - Санкт-Петербург: Лань, 2022. - 288 с. ISBN 978-5-8114-8282-5.  | учеб. пособие   | 2022               | ЭБС "Лань"  |
|                  | <b>Акулич И. Л.</b>        | Математическое программирование в примерах и задачах : учебное пособие / И. Л. Акулич. – Изд. 5-е, стер. – Санкт-Петербург : Лань, 2024. – 347 с. : ил. – (Высшее образование). – URL: <a href="https://e.lanbook.com/book/360428">https://e.lanbook.com/book/360428</a> (дата обращения: 24.05.2024). – Режим доступа: Электронно-библиотечная система "Лань". – ISBN 978-5-507-47317-5. – Текст : электронный. URL: <a href="https://e.lanbook.com/book/360428">https://e.lanbook.com/book/360428</a>                    | Учебное пособие   | 2024               | "Лань"  |
|                  | <b>Голубева Н. В.</b>      | Математическое моделирование систем и процессов : учебное пособие / Н. В.  | Учебное пособие   | 2024               | "Лань"  |

| №<br>п/п | Авторы, составители | Заглавие (заголовок)   | Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.) | Год издания | Количество в научной библиотеке / Наименование ЭБС <sup>1</sup> |
|----------|---------------------|--|---|-------------|---|
|          |                     | Голубева. – Изд. 4-е, испр. и доп. – Санкт-Петербург : Лань, 2024. – 244 с. : ил. – (Высшее образование). – URL: <a href="https://e.lanbook.com/book/393023">https://e.lanbook.com/book/393023</a> (дата обращения: 05.03.2024). – Режим доступа: Электронно-библиотечная система "Лань". – ISBN 978-5-507-48455-3. – Текст : электронный.<br><br>URL: <a href="https://e.lanbook.com/book/393023">https://e.lanbook.com/book/393023</a> |   |             |   |
|          |                     |  |   |             |   |
|          |                     |  |   |             |   |
|          |                     |  |   |             |   |

## 8.2. Дополнительная литература

| №<br>п/п | Авторы, составители | Заглавие (заголовок)                 | Тип (учебник, учебное пособие, учебно-методическое пособие, практикум, др.) | Год издания | Количество в научной библиотеке / Наименование ЭБС |
|----------|---------------------|--------------------------------------|---|-------------|--|
| 1        | Евдошенко О.И.      | Проектирование информационных систем | Учебно-методическое пособие   | 2022        | ЭБС "IPRbooks"                                     |
| 2        | Золотов С. Ю.       | Проектирование информационных систем | Учебное пособие   | 2013        | ЭБС "IPRbooks"                                     |

### 8.3. Перечень профессиональных баз данных и информационных справочных систем

| № пп | Наименование   | Ссылка  |
|------|--|---|
| 1    | Springer Nature (Полнотекстовая коллекция журналов)                                      | <a href="https://www.springernature.com/gp/products">https://www.springernature.com/gp/products</a> |
| 2    | Springer eBooks (Полнотекстовая коллекция электронных книг издательства Springer Nature) | <a href="https://link.springer.com/">https://link.springer.com/</a>                                 |
| 3    | «Кодекс»   | <a href="https://kodeks.ru/">https://kodeks.ru/</a>   |
| 4    | ELIBRARY.RU (электронная библиотека научных публикаций)                                  | <a href="http://elibrary.ru">http://elibrary.ru</a>   |
| 5    | "Гарант"   | <a href="https://www.garant.ru/">https://www.garant.ru/</a>   |
| 6    | "КонсультантПлюс"  | <a href="https://www.consultant.ru/">https://www.consultant.ru/</a>                                 |
| 7    | Техэксперт   | <a href="https://cntd.ru/">https://cntd.ru/</a>   |

### 8.4. Перечень программного обеспечения

| № п/п | Наименование ПО                                      | Реквизиты договора<br>(дата, номер, срок действия)  |
|-------|--|---|
| 1     | WinPro 10 RUS Upgrd OLP NL Acdmc                     | Договор № 757 от 04.07.2018, срок действия - бессрочно;<br>Контракт № 1653 от 14.12.2018, срок действия – бессрочно |
| 2     | Office Stdandard 2013 Russian OLP NL AcademicEdition | Контракт № 690 от 19.05.2015, срок действия - бессрочно)  |

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

| № п/п | Наименование оборудованных учебных кабинетов, лабораторий, мастерских и др. объектов для проведения практических и лабораторных занятий, помещений для самостоятельной работы обучающихся<br>(номер аудитории)  | Перечень основного оборудования   |
|-------|---|---|
| 1     | Учебная аудитория для проведения занятий лекционного типа. Учебная аудитория для проведения занятий семинарского типа. Учебная аудитория для курсового проектирования (выполнения курсовых работ). Учебная аудитория для проведения групповых и индивидуальных консультаций. Учебная аудитория для проведения занятий текущего контроля и промежуточной аттестации. (УЛК-413) | Стол� ученические двухместные (моноблок), стол преподавательский, стул, доска аудиторная (меловая), проектор. |

| №<br>п/п | <b>Наименование оборудованных учебных кабинетов, лабораторий, мастерских и др. объектов для проведения практических и лабораторных занятий, помещений для самостоятельной работы обучающихся (номер аудитории)</b> | <b>Перечень основного оборудования</b>   |
|----------|--|--|
| 2        | Помещение для самостоятельной работы обучающихся. (УЛК-105)  | Столы, стулья, стеллажи (в т.ч. выставочные) с книгами, компьютеры, мобильные рабочие места. |
| 3        | Помещение для самостоятельной работы обучающихся. (УЛК-406)  | Столы компьютерные, стулья, микрокомпьютеры raspberry pi 32 bit..                            |