

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Баламирзоев Назим Лиодинович  
Должность: Ректор  
Дата подписания: 25.03.2026 16:00:31  
Уникальный программный ключ:  
5cf0d6f89e80f49a334f6a4ba58e91f3326b9926



**МИНОБРНАУКИ РОССИИ**

Федеральное государственное бюджетное образовательное учреждение  
высшего образования

**Институт кибербезопасности и цифровых технологий**

**Региональный партнёр**

**ФГБОУ ВО**

**«Дагестанский государственный технический университет»**



**X**

**Б1.В.02 Основы обучаемых алгоритмов**

Направление подготовки 09.03.01 «Информатика и вычислительная техника»

Направленность (профиль подготовки): «Прикладной искусственный интеллект»

Квалификация выпускника бакалавр

Форма обучения очная

Махачкала 2023

# 1. ПАСПОРТ фонда оценочных средств

## по дисциплине Б1.В.02 Основы обучаемых алгоритмов

### 1.1. Результаты обучения по дисциплине:

Код	Наименование компетенции	Индикатор достижения компетенции	В результате освоения дисциплины обучающийся должен:	Другая дисциплина (дисциплины) / практика, участвующая в формировании компетенции
ПК-2	Способен использовать системы искусственного интеллекта в решении задач анализа, прогнозирования, планирования, синтеза и принятия решений	ПК-2.1 (ПК-3.1 РЭУ) Выбирает методы решения задач с использованием систем искусственного интеллекта	Знать: основные обучаемые алгоритмы Уметь: использовать основные обучаемые алгоритмы Владеть: программными средствами реализации обучаемых алгоритмов	Системы искусственного интеллекта Методы решения оптимизационных задач Производственная (проектно-технологическая) практика
		ПК-2.2 (ПК-3.2 РЭУ) Решает задачи с использованием систем искусственного интеллекта	Знать: возможности одной из основных библиотек машинного обучения. Уметь: выбирать функции библиотеки машинного обучения для решения задачи. Владеть: программными средствами библиотеки машинного обучения.	Нейронные сети в решении практических задач Основы компьютерного зрения Методы обработки естественного языка Интеллектуальные методы обработки сигналов Производственная (проектно-технологическая) практика
ПК-4	Способен использовать одну или несколько сквозных цифровых субтехнологий искусственного интеллекта	ПК-4.1 (ПК-8.1 РЭУ). Решает прикладные задачи и участвует в реализации проектов в области сквозной цифровой субтехнологии "Компьютерное зрение"	Знать: базовую архитектуру и алгоритмы обучения нейронных сетей. Уметь: выбирать архитектуру и алгоритмы обучения полносвязных нейронных сети. Владеть: программными средствами реализации полносвязных нейронных сетей.	Основы компьютерного зрения Производственная (технологическая) практика
		ПК-4.2 (ПК-8.2 РЭУ) Решает прикладные задачи и участвует в реализации проектов в области сквозной цифровой субтехнологии «Обработка	Знать: базовую архитектуру и алгоритмы обучения нейронных сетей. Уметь: выбирать архитектуру и алгоритмы обучения	Методы обработки естественного языка Производственная (технологическая) практика

		естественного языка»	полносвязных нейронных сети. Владеть: программными средствами реализации полносвязных нейронных сетей.	
ПК-6	Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта	ПК-6.1 (ПК-6.1 РЭУ) Осуществляет поиск данных в открытых источниках, специализированных библиотеках и репозиториях	Знать: возможности одной из основных библиотек машинного обучения. Уметь: проводить поиск данных и функций в одной из основных библиотек машинного обучения и открытых источниках. Владеть: поиском данных в открытых источниках.	Обработка больших данных Производственная (технологическая) практика
		ПК-6.2 (ПК-6.2 РЭУ) Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения	Знать: основы подготовки данных для машинного обучения. Уметь: выбирать алгоритмы подготовки данных для машинного обучения. Владеть: программными средствами подготовки данных для машинного обучения.	Основы компьютерного зрения Методы обработки естественного языка Производственная (технологическая) практика

## 1.2. Программа оценивания контролируемой компетенции:

№ п/п	Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции / индикатора	Наименование оценочного средства
1.	Введение в обучаемые алгоритмы	ПК-2/ ПК-2.1	Собеседование
2.	Подготовка данных для машинного обучения	ПК-2/ ПК-2.1 ПК-2/ ПК-2.2 ПК 6/ПК-6.1 ПК 6/ПК-6.2	Собеседование
3.	Регрессия	ПК-2/ ПК-2.1 ПК-2/ ПК-2.2	Собеседование
4.	Классификация и кластеризация	ПК-2/ ПК-2.1 ПК-2/ ПК-2.2	Собеседование
5.	Ансамбли моделей машинного обучения	ПК-2/ ПК-2.1 ПК-2/ ПК-2.2	Собеседование
6.	Многослойный перцептрон	ПК-2/ ПК-2.1 ПК-2/ ПК-2.2 ПК-4/ПК-4.1 ПК-4/ПК-4.2	Собеседование
Форма промежуточной аттестации в 3 семестре — зачет с оценкой			

### 3. Контрольные вопросы и задания

#### 3.1. Вопросы к экзамену по дисциплине "Основы обучаемых алгоритмов"

1. Основные понятия машинного обучения: модель алгоритмов, объекты и признаки, метод обучения, функция потерь и функционал качества, принцип минимизации эмпирического риска, обобщающая способность.
2. Что такое функция потерь?
3. Примеры прикладных задач, решаемых с помощью машинного обучения.
4. Основные возможности библиотеки Scikit learn.
5. Виды данных для машинного обучения.
6. Как заполняются пропущенные значения в данных?
7. Что такое одномерные и многомерные заполнители?
8. Очистка данных.
9. Кодирование качественных данных.
10. Нормализация и стандартизация.
11. Методы отбора признаков.
12. Декорреляция данных.
13. Понижение размерности данных методом главных компонент.
14. Понятие линейной регрессии.
15. Построение линейной регрессии.
16. Регуляризация линейной регрессии.
17. Оценка качества регрессионной модели.
18. Логистическая регрессия и задача классификации.
19. Теорема Байеса и наивный Байесовский классификатор.
20. Метод ближайших соседей.
21. Понятие дерева решений.
22. Построение дерева решений.
23. Регуляризация деревьев решений.
24. Понятие метода опорных векторов.
25. Задача кластеризации. Кластеризация методом  $k$ -ближайших соседей.
26. Иерархические алгоритмы кластеризации.
27. Алгоритм кластеризации DBSCAN.
28. Перекрестная проверка моделей.
29. Метрики качества задач классификации.
30. ROC-кривая.
31. Ансамбли классификаторов с голосованием.
32. Алгоритм бэггинг.
33. Алгоритм AdaBoost.
34. Градиентный бустинг.
35. Биологический и искусственный нейроны.
36. Математическая модель искусственного нейрона.
37. Функции активации.
38. Структура полносвязных нейронных сетей.
39. Алгоритм обратного распространения обучения нейронных сетей.
40. Градиентные алгоритмы обучения нейронных сетей первого порядка: градиентный спуск, градиентный спуск с импульсом, метод Нестерова.
41. Идея градиентных алгоритмов с адаптивной скоростью обучения.
42. Регуляризация в обучении нейронных сетей.

### 3.2. Вопросы для проверки уровня обученности ЗНАТЬ и УМЕТЬ

1. Что такое машинное обучение?
2. Возможно ли построить модель машинного обучения, хорошо работающую на любых данных?
3. Чем отличаются обучение с учителем и обучение без учителя?
4. Что такое обучение с подкреплением?
5. Что такое прецедент в машинном обучении?
6. Перечислите типы признаков.
7. Объясните элементы матрицы признаков.
8. Приведите примеры бинарных, качественных и количественных признаков.
9. Приведите примеры задач классификации и регрессии.
10. В чем отличие функционала качества от функции потерь?
11. Что такое Anaconda?
12. Перечислите основные достоинства Jupyter Notebook.
13. Как запустить Jupyter Notebook в заданной папке?
14. Что такое ячейки и ядро Jupyter Notebook?
15. Что может содержать ячейка Jupyter Notebook?
16. Приведите примеры "магических команд" Jupyter Notebook.
17. Как корректно закрыть Jupyter Notebook?
18. Как в Jupyter Notebook устанавливаются дополнительные пакеты?
19. Как производится обновление Anaconda?
20. Как в Google Colab создать блокнот?
21. Как в Google Colab сохранить блокнот на своем компьютере?
22. Какова основная идея "одномерного" заполнителя?
23. Какова основная идея "многомерного" заполнителя?
24. Что из следующего верно/неверно и почему?
  - "Одномерный" заполнитель способен работать с данными, представленными в целых числах.
  - "Одномерный" заполнитель неспособен работать с данными, представленными в вещественных числах.
  - "Одномерный" заполнитель неспособен работать с данными, представленным в строках.
  - "Многомерный" заполнитель не способен работать, если количество итераций  $\max\_iter = 1$ .
25. Зачем необходима нормализация данных?
26. Для чего необходимо отбирать признаки при подготовке данных?
27. Какие существуют подходы для отбора наиболее информативных признаков?
28. Объясните идею метода главных компонент.
29. Что такое линейная регрессия?
30. Как находятся коэффициенты линейной регрессии?
31. С какой целью применяется регуляризация в регрессии?
32. Что такое гребневой регрессии?
33. Что такое лассо регрессия?
34. Как оценивается качество регрессионных моделей?
35. Что такое логистическая регрессия?
36. Какова основная идея Байесовского классификатора?
37. Как Байесовский классификатор определяет наиболее вероятный класс?
38. Почему Байесовский классификатор называют "наивным"?
39. Какие распределения признаков используются?
40. Почему метод ближайших соседей относится к непараметрическим методам?

41. Какие метрики расстояния используются в классификации методом ближайших соседей?
42. Назовите основные преимущества деревьев решений перед другими моделями.
43. Назовите основные недостатки деревьев решений.
44. Что такое чистота узла дерева?
45. Как выбираются атрибуты ветвления?
46. Что такое индекс Джини?
47. Что такое критерий уменьшения энтропии?
48. Как производится регуляризация деревьев решений?
49. Какие показатели применяются для оценки загрязненности узла?
50. Назовите основные преимущества деревьев решений перед другими моделями.
51. Назовите основные недостатки деревьев решений.
52. Что такое чистота узла дерева?
53. Как выбираются атрибуты ветвления?
54. Что такое индекс Джини?
55. Что такое критерий уменьшения энтропии?
56. Как производится регуляризация деревьев решений?
57. Какие показатели применяются для оценки загрязненности узла?
58. Чем задача кластеризации отличается от задачи классификации?
59. Какие метрики используются в кластеризации?
60. Сравните рассмотренные алгоритмы кластеризации.
61. Как можно оценить качество кластеризации?
62. Как строится алгоритм DBSCAN?
63. Алгоритм DBSCAN требует задания количества кластеров?
64. Что такое иерархические алгоритмы кластеризации?
65. Как можно выбрать оптимальное количество кластеров?
66. Перечислите метрики классификаторов.
67. В чем недостаток метрики "точность"?
68. Классификатор должен определять больных пациентов. Дайте содержательную трактовку метрик классификатора.
69. Как строится ROC кривая?
70. Как строится кривая полноты-точности?
71. Что такое нейронная сеть?
72. Структура полносвязной нейронной сети.
73. Основные функции активации нейронов, используемые в нейронных сетях?
74. В чем недостаток сигмоидальных функций активации и как его можно преодолеть?
75. Почему важна дифференцируемость функции активации?
76. Опишите математически слой полносвязной нейронной сети.
77. Что такое функция потерь?
78. Какие функции потерь используются при решении задач классификации и регрессии?
79. По какой формуле изменяются веса сети в алгоритме градиентного спуска?
80. Чем градиентный спуск с импульсом отличается от обычного градиентного спуска?
81. Что такое алгоритм импульса Нестерова?
82. Что такое алгоритм обратного распространения ошибки?
83. Какие приемы преодоления переобучения используются в Scikit Learn?

## 4. Задания для проверки обученности ВЛАДЕТЬ

### 1. Подготовка данных для машинного обучения

1. Используя код, представленный ниже, проверить какое значение `strategy` дает меньшее значение ошибки (в качестве показателя ошибки применяется метрика Минковского). Нужно проверить параметры заполнения "mean", "median" и "most\_frequent".

```
import numpy as np
from sklearn.impute import SimpleImputer
from scipy.spatial.distance import minkowski

# Функция вычисления расстояния Минковского между массивами
def calc_error (a, b):
    # Преобразование массива в строку в строчном порядке
    a_flatten = np.matrix(a).flatten()
    b_flatten = np.matrix(b).flatten()
    # Вычисление округленного расстояния Минковского при p=2
    return round(minkowski(a_flatten, b_flatten, 2), 4)

imp = SimpleImputer(missing_values=np.nan, strategy="mean")
# Настройка алгоритма заполнения
imp.fit([[ 2, 1],
        [ 3, np.nan],
        [np.nan, 2],
        [ 1, 0.5]])

X = [[ 2.5, np.nan],
     [ 4.1, np.nan],
     [np.nan, 2],
     [ 1, np.nan]]
# Заполнение матрицы с пропусками
a = imp.transform(X)
b = [[2.5, 1.25],
     [4.1, 2.05],
     [4, 2],
     [1, 0.5]]

print("mean", calc_error(a, b))
a
```

*Указания к выполнению задания 1*

Расстояние (метрика) Минковского между двумя векторами  $x$  и  $y$  в  $n$ -мерном пространстве определяется следующим образом

$$d(x, y) = \left( \sum_{i=1}^n |x_i - y_i| \right)^{1/p}.$$

При  $p=1$  получается расстояние городских кварталов (манхэттенское расстояние), при  $p=2$  — евклидова метрика.

В приведенном коде пропуски заполняются средним значением по столбцу. Массивы представляются как векторы построчно (сначала первая строка, затем вторая и т.д.). Вычисляется ошибка между элементами заполненного массива  $a$  и некоторого массива  $b$  с из-

вестными элементами. Учтите, что в примере заполнитель настраивается по одному массиву, а заполняется другой массив X.

2. Используя код, представленный ниже, заполнить пропуски в признаках и вычислить значение ошибки. Полученное значение ошибки больше или меньше, чем в предыдущем задании? Почему?

```
import numpy as np
from sklearn.experimental import enable_iterative_imputer
from sklearn.impute import IterativeImputer

imp = IterativeImputer(max_iter=25, random_state=0)
imp.fit([[ 2, 1],
        [ 3, np.nan],
        [np.nan, 2],
        [ 1, 0.5]])
X = [[ 2.5, np.nan],
     [ 4.1, np.nan],
     [np.nan, 2],
     [ 1, np.nan]]
a = imp.transform(X)
b = [[2.5, 1.25],
     [4.1, 2.05],
     [ 4, 2],
     [ 1, 0.5]]
print(calc_error(a, b))
```

## 2. Регрессия

1. Используя входные значения X и целевые значения y обучить обычную линейную регрессию (LinearRegression в Scikit-learn):

```
X = [[-2.1], [-1.1], [0], [1.1], [2.1]]
y = [-2, -1, 0, 1, 2]
```

На полученной модели оценить качество модели используя метрики:

- Значение среднеквадратичной ошибки
- Значение корня среднеквадратичной ошибки

Для вычисления среднеквадратичной ошибки можно воспользоваться кодом ниже:

```
from sklearn.metrics import mean_squared_error as sklearn_mse
def mse(outputs, targets):
    return sklearn_mse(outputs, targets)
```

На основе вычисленных метрик оценить качество полученной модели.

2. Используя данные, представленные ниже обучить обыкновенную линейную регрессию и оценить качество построенной модели, основываясь на значении среднеквадратичной ошибки:

```
X = [[0.1], [0.3], [0.1], [0.2]]
y = [1, 3, 1, 2]
```

Попробуйте обучить регрессию с регуляризацией (Lasso или Ridge), рекомендуем начинать обучение со значения  $\alpha=0.1$ , постепенно уменьшая значение, если потребуется. Регуляризованные модели оказалось обучить проще или сложнее?

3. Возьмите набор чисел  $[-7, -5, -3, -1, 0, 1, 3, 5, 7]$ . Обучите модель логистической регрессии, способную различать положительные и отрицательные числа. Сможете ли Вы создать подобную модель для обычной линейной регрессии или линейной регрессии с регуляризатором?

## 2. Наивный Байесовский классификатор

1. Используя набор "Ирисы Фишера", поэкспериментируйте с Байесовским классификатором. Установите соответствие между номером класса в `sklearn` и латинским названием цветка.

2. Используя наивный Байесовский классификатор и набор данных `Breast Cancer Wisconsin` [6], проведите диагностику рака молочной железы.

## 3. Метод ближайших соседей

Используя один из наборов данных [6], например, `Breast Cancer Wisconsin` проведите классификацию.

## 4. Деревья решений

Используя один из наборов данных, например, `Breast Cancer Wisconsin` проведите классификацию с использованием дерева решений. Проведите эксперименты с деревьями разной глубины.

## 5. Метод опорных векторов

1. Используя классификатор с Гауссовым ядром, исследуйте влияние на результаты классификации параметра классификатора.

2. Проведите эти же эксперименты для классификаторов с линейным, полиномиальным и сигмоидальным ядрами.

## 6. Методы кластеризации

Используя набор данных "Ирисы Фишера" из библиотеки `scikit-learn` и учитывая только длину и ширину лепестков, постройте модель кластеризации данных различными алгоритмами. Постройте диаграммы рассеяния данных для всех алгоритмов

## 7. Оценка классификаторов

1. Используя пример двухклассового классификатора с добавлением шума в признаки, постройте ROC-кривую. Исследуйте влияние количества случайных признаков на характеристики классификатора.

2. По указанию преподавателя для одного из набора данных из библиотеки `datasets` постройте классификатор и оцените его характеристики с использованием ROC-кривой и кривой полноты-точности.

### **8. Ансамбли моделей**

Используя набор данных `breast cancer` (рак молочной железы) библиотеки `scikit-learn`, реализуйте различные ансамбли.

### **9. Нейронные сети**

1. На примере набора данных "Ирисы Фишера" исследуйте влияние структуры сети (числа слоев и нейронов), функции активации скрытых слоев и алгоритмов обучения на метрики качества классификатора.

2. Используя набор данных `breast cancer` (рак молочной железы) библиотеки `scikit-learn`, реализуйте нейросетевой классификатор.

### **5. Описание показателей и критериев оценивания с указанием шкалы оценивания для очной формы обучения (с применением балльно-рейтинговой системы):**

Оценка по дисциплине складывается из текущего рейтинга и рейтинга на зачете.

Текущий рейтинг складывается из оценок, полученных при защите лабораторных работ. Для лабораторных работ 1–16 максимальное число баллов за работу равно 3. Для работ 17–18 максимальное число баллов за работу равно 6. Студент считается сдавшим лабораторные работы, если он сдал все работы и набрал не менее 36 баллов.

Рейтинг на зачете определяется следующим образом:

Ответы на первый и второй вопросы — до 10 баллов, выполнение задания — до 10 баллов, дополнительные вопросы в рамках курса до 10 баллов.

*Оценивание ответов на первый, второй и дополнительные вопросы:*

*9-10 баллов выставляется, если студент демонстрирует полное понимание вопросов, правильность ответов, полное и логически последовательное изложение материала.*

*7-8 баллов выставляется, если студент демонстрирует: значительное понимание вопросов, правильность, но недостаточную полноту ответов на заданные теоретические вопросы; допущение неточности ответа;*

*5-6 баллов выставляется, если студент демонстрирует: понимание вопросов, по существу излагает материал, но не усвоил его деталей, есть погрешности в ответах; допускает неточности, недостаточно правильные формулировки, нарушения логической последовательности в изложении материала;*

*Менее 5 баллов выставляется, если студент демонстрирует: непонимание вопросов; студент не знает значительной части материала, не ответил на дополнительные вопросы или отказался от ответов на вопросы и задания.*

*Оценивание задания:*

*9-10 баллов выставляется, если студент демонстрирует полное понимание заданий, правильность ответов; полное, точное и логически последовательное изложение материала;*

*7-8 баллов выставляется, если студент демонстрирует:* значительное понимание заданий, правильность, но недостаточную полноту ответов на заданные задания; допущение неточности ответа;

*5-6 баллов выставляется, если студент демонстрирует:* понимание заданий, основные этапы задания выполнены, но есть погрешности в ответах

*Менее 5 баллов выставляется, если студент демонстрирует:* непонимание заданий; основные шаги задания не выполнены или выполнены неправильно, не ответил на дополнительные вопросы или отказался от ответов на вопросы и задания.

Минимальный балл рейтинга на зачете в соответствии с положением о рейтинге равен 24, максимальный балл — 40.

В итоге по курсу, суммируя итоги текущего рейтинга и рейтинга на зачете оценка "зачтено" выставляется обучающемуся, если он набрал не менее 60 баллов.

**Сведения о дополнениях и изменениях, внесенных в ФОС дисциплины**

Учебный год	Решение кафедры (№ протокола, дата)	Внесенные в ФОС дополнения и измене- ния	Подпись заведующего кафедрой