

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ  
ИМЕНИ В.Ф. УТКИНА»

Кафедра «Электронные вычислительные машины»

## **ОЦЕНОЧНЫЕ МАТЕРИАЛЫ**

по дисциплине

**«Прикладные проекты машинного обучения»**

Направление подготовки

09.03.01 Информатика и вычислительная техника

Направленность (профиль) подготовки

**«Программно-аппаратное обеспечение вычислительных  
комплексов и систем искусственного интеллекта»**

Квалификация выпускника – бакалавр

Форма обучения – очная

## **1 ОБЩИЕ ПОЛОЖЕНИЯ**

Оценочные материалы – это совокупность учебно-методических материалов (контрольных заданий, описаний форм и процедур), предназначенных для оценки качества освоения обучающимися данной дисциплины как части основной профессиональной образовательной программы.

Цель – оценить соответствие знаний, умений и уровня приобретенных компетенций, обучающихся целям и требованиям основной профессиональной образовательной программы в ходе проведения текущего контроля и промежуточной аттестации.

Основная задача – обеспечить оценку уровня сформированности общекультурных, общепрофессиональных и профессиональных компетенций, приобретаемых обучающимся в соответствии с этими требованиями.

Контроль знаний проводится в форме текущего контроля и промежуточной аттестации.

Текущий контроль успеваемости проводится с целью определения степени усвоения учебного материала, своевременного выявления и устранения недостатков в подготовке обучающихся и принятия необходимых мер по совершенствованию методики преподавания учебной дисциплины (модуля), организации работы обучающихся в ходе учебных занятий и оказания им индивидуальной помощи.

К контролю текущей успеваемости относятся проверка знаний, умений и навыков, приобретенных обучающимися в ходе выполнения индивидуальных заданий на практических занятиях и лабораторных работах. При оценивании результатов освоения практических занятий и лабораторных работ применяется шкала оценки «зачтено – не зачтено». Количество лабораторных и практических работ и их тематика определена рабочей программой дисциплины, утвержденной заведующим кафедрой.

Результат выполнения каждого индивидуального задания должен соответствовать всем критериям оценки в соответствии с компетенциями, установленными для заданного раздела дисциплины.

Промежуточный контроль по дисциплине осуществляется проведением экзамена и теоретического зачета.

Форма проведения экзамена – письменный ответ по утвержденным экзаменационным билетам, сформулированным с учетом содержания учебной дисциплины. В экзаменационный билет включается два теоретических вопроса и одна задача. После выполнения письменной работы обучаемого производится ее оценка преподавателем и, при необходимости, проводится теоретическая беседа с обучаемым для уточнения экзаменационной оценки.

## **2 ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ ОЦЕНИВАНИЯ КОМПЕТЕНЦИЙ**

Сформированность каждой компетенции (или ее части) в рамках освоения данной дисциплины оценивается по трехуровневой шкале:

- 1) пороговый уровень является обязательным для всех обучающихся по завершении освоения дисциплины;
- 2) продвинутый уровень характеризуется превышением минимальных характеристик сформированности компетенций по завершении освоения дисциплины;
- 3) эталонный уровень характеризуется максимально возможной выраженностью компетенций и является важным качественным ориентиром для самосовершенствования.

### Уровень освоения компетенций, формируемых дисциплиной:

#### Описание критериев и шкалы оценивания тестирования:

| Шкала оценивания                 | Критерий  |
|----------------------------------|---|
| 3 балла<br>(эталонный уровень)   | Уровень усвоения материала, предусмотренного программой: процент верных ответов на тестовые вопросы от 85 до 100% |
| 2 балла<br>(продвинутый уровень) | Уровень усвоения материала, предусмотренного программой: процент верных ответов на тестовые вопросы от 70 до 84%  |
| 1 балл<br>(пороговый уровень)    | Уровень усвоения материала, предусмотренного программой: процент верных ответов на тестовые вопросы от 50 до 69%  |
| 0 баллов                         | Уровень усвоения материала, предусмотренного программой: процент верных ответов на тестовые вопросы от 0 до 49%   |

#### Описание критериев и шкалы оценивания теоретического вопроса:

| Шкала оценивания                 | Критерий   |
|----------------------------------|--|
| 3 балла<br>(эталонный уровень)   | выставляется студенту, который дал полный ответ на вопрос, показал глубокие систематизированные знания, смог привести примеры, ответил на дополнительные вопросы преподавателя |
| 2 балла<br>(продвинутый уровень) | выставляется студенту, который дал полный ответ на вопрос, но на некоторые дополнительные вопросы преподавателя ответил только с помощью наводящих вопросов                    |
| 1 балл<br>(пороговый уровень)    | выставляется студенту, который дал неполный ответ на вопрос в билете и смог ответить на дополнительные вопросы только с помощью преподавателя                                  |
| 0 баллов                         | выставляется студенту, который не смог ответить на вопрос  |

#### Описание критериев и шкалы оценивания практического задания:

| Шкала оценивания                 | Критерий  |
|----------------------------------|---|
| 3 балла<br>(эталонный уровень)   | Задача решена верно   |
| 2 балла<br>(продвинутый уровень) | Задача решена верно, но имеются неточности в логике решения               |
| 1 балл<br>(пороговый уровень)    | Задача решена верно, с дополнительными наводящими вопросами преподавателя |
| 0 баллов                         | Задача не решена  |

На промежуточную аттестацию (экзамен) выносятся тест, два теоретических вопроса и 2 задачи. Максимально студент может набрать 15 баллов. Итоговый суммарный балл студента, полученный при прохождении промежуточной аттестации, переводится в традиционную форму по системе «отлично», «хорошо», «удовлетворительно» и «неудовлетворительно».

**Оценка «отлично»** выставляется студенту, который набрал в сумме 15 баллов (выполнил все задания на эталонном уровне). Обязательным условием является выполнение всех предусмотренных в течение семестра практических заданий.

**Оценка «хорошо»** выставляется студенту, который набрал в сумме от 10 до 14 баллов при условии выполнения всех заданий на уровне не ниже продвинутого. Обязательным условием является выполнение всех предусмотренных в течение семестра практических заданий.

**Оценка «удовлетворительно»** выставляется студенту, который набрал в сумме от 5 до 9 баллов при условии выполнения всех заданий на уровне не ниже порогового.

Обязательным условием является выполнение всех предусмотренных в течение семестра практических заданий.

**Оценка «неудовлетворительно»** выставляется студенту, который набрал в сумме менее 5 баллов или не выполнил всех предусмотренных в течение семестра практических заданий.

### 3 ПАСПОРТ ОЦЕНОЧНЫХ МАТЕРИАЛОВ ПО ДИСЦИПЛИНЕ

| Контролируемые разделы (темы) дисциплины          | Код контролируемой компетенции (или её части) | Вид, метод, форма оценочного мероприятия |
|---|---|--|
| Раздел 1. Генеративные модели данных              | ПК-2, ПК-11, ПК-16                            | Экзамен                                  |
| Раздел 2. Кастомные архитектуры и слои            | ПК-2, ПК-16                                   | Экзамен                                  |
| Раздел 3. Компьютерное зрение и видеоаналитика    | ПК-7, ПК-11, ПК-14                            | Экзамен                                  |
| Раздел 4. Обучение и эксплуатация нейронных сетей | ПК-7, ПК-11, ПК-14                            | Экзамен                                  |
| Раздел 5. Продвинутое методы обучения             | ПК-7, ПК-11, ПК-14, ПК-16                     | Экзамен                                  |
| Раздел 6. Промежуточная аттестация                | ПК-2, ПК-7, ПК-11, ПК-16                      | Экзамен                                  |

### 4 ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ

#### Промежуточная аттестация в форме экзамена

| Код компетенции | Результаты освоения ОПОП<br>Содержание компетенций             |
|-----------------|--|
| ПК-2            | Способен проектировать и разрабатывать программное обеспечение |

ПК-2.1. Проектирует и разрабатывает программное обеспечение

ПК-2.2. Применяет современные инструментальные средства при разработке программного обеспечения

#### Типовые тестовые вопросы:

1. Какой принцип ООП описывает скрытие внутренней реализации объекта от внешнего мира?

- а) Наследование
- б) Полиморфизм
- в) Инкапсуляция**
- г) Абстракция

2. Какой инструмент НЕ относится к системам контроля версий?

- а) Git
- б) SVN
- в) Maven**
- г) Mercurial

3. Инструмент MLflow primarily используется для:

- а) Визуализации данных
- б) Управления жизненным циклом ML-моделей**
- в) Разработки интерфейсов

г) Тестирования безопасности

4. Какая из перечисленных парадигм программирования НЕ относится к ООП?

а) Наследование

б) Полиморфизм

**в) Функциональное программирование**

г) Инкапсуляция

5. Какой из перечисленных форматов НЕ используется для сериализации данных?

а) JSON

б) YAML

**в) SQL**

г) Protocol Buffers

6. Паттерн «Стратегия» (Strategy) позволяет:

**а) Изменять алгоритмы выполнения операции**

б) Создавать сложные объекты

в) Управлять состоянием объекта

г) Организовывать иерархии классов

7. Какова основная цель генеративной модели, такой как Variational Autoencoder (VAE), в отличие от дискриминативной модели?

а) Точно классифицировать данные по заданным меткам.

**б) Научиться реалистичному распределению исходных данных для генерации новых образцов.**

в) Находить оптимальную разделяющую гиперплоскость между классами.

г) Снижать размерность данных только для задачи визуализации.

8. В чем ключевое архитектурное различие между стандартным автоэнкодером (AE) и вариационным автоэнкодером (VAE), которое позволяет VAE генерировать новые данные?

а) VAE использует сверточные слои, а AE — полносвязные.

б) VAE имеет больший скрытый слой (латентный код), чем AE.

**в) VAE кодирует входные данные в распределение (например, гауссово), а не в фиксированную точку.**

г) VAE использует более сложную функцию потерь только для реконструкции.

9. Какой из перечисленных подходов является ПРАВИЛЬНЫМ способом создания собственного слоя в TensorFlow/Keras, который имеет обучаемые веса?

а) Унаследоваться от класса `tf.Module` и реализовать метод `__call__`.

**б) Унаследоваться от класса `tf.keras.layers.Layer` и реализовать метод `build` (для создания весов) и `call` (для прямой передачи).**

в) Создать функцию Python, которая использует `tf.function` для компиляции.

г) Использовать встроенный класс `tf.keras.layers.Lambda` для любого слоя с обучаемыми параметрами.

10. При реализации Generative Adversarial Network (GAN) с помощью TensorFlow/Keras, почему для Discriminator (Дискриминатора) и Generator (Генератора) обычно создают две ОТДЕЛЬНЫЕ модели, вместо одной объединенной?

**а) Это позволяет использовать разные оптимизаторы для каждой модели, что критически важно для стабильного обучения.**

б) Объединенная модель не сможет скомпилироваться из-за конфликта функций потерь.

в) Так требуется меньше видеопамати.

г) Архитектура GAN не позволяет создать их как одну модель.

### Типовые теоретические вопросы

1. Вопрос: Что такое «инкапсуляция» в ООП?

Ответ: Принцип объединения данных и методов в единый объект с ограничением доступа к

внутренним данным.

2. Вопрос: Что такое CI/CD?

Ответ: Практика автоматической сборки, тестирования и развертывания кода.

3. Вопрос: Что такое «чистый код»?

Ответ: Код, который легко читать, понимать и поддерживать.

4. Вопрос: Что такое «полиморфизм»?

Ответ: Возможность объектов с одинаковой спецификацией иметь различную реализацию.

5. Вопрос: Для чего используется Git?

Ответ: Для контроля версий и совместной разработки.

6. Вопрос: Что такое «микросервисная архитектура»?

Ответ: Архитектура, где приложение состоит из небольших независимых сервисов.

7. Вопрос: Для чего используется Jupyter Notebook?

Ответ: Интерактивная среда для разработки и визуализации данных.

8. Вопрос: Что такое «тестирование программного обеспечения»?

Ответ: Процесс проверки соответствия программы требованиям.

| Код компетенции | Результаты освоения ОПОП<br>Содержание компетенций   |
|-----------------|--|
| ПК-7            | Способен применять и (или) разрабатывать алгоритмы, методы и технологии компьютерного зрения |

ПК-7.1. Применяет (проводя выбор и эксперименты) известные алгоритмы и библиотеки компьютерного зрения, предобученные глубокие нейросетевые модели для прикладных задач анализа изображений и видеопотока, при необходимости дообучая и валидируя на собственных наборах данных

ПК-7.2. Определяет стек технологий, методов и алгоритмов для построения продуктов с компьютерным зрением (системы видеоаналитики, поисковые системы по изображениям и т.д.)

ПК-7.3. Имплементирует известные алгоритмы, архитектуры и модели компьютерного зрения на реальных данных, строит пайплайны обучения моделей и развертывания сервисов компьютерного зрения в продуктивной среде

#### Типовые тестовые вопросы:

- Какой алгоритм НЕ используется для детекции объектов?
  - YOLO
  - SSD
  - U-Net**
  - Faster R-CNN
- Для чего используется операция пулинга в CNN?
  - Увеличение размера feature map
  - Уменьшение размерности и выделение основных признаков**
  - Улучшение цветопередачи
  - Сжатие изображения
- Метод SIFT в компьютерном зрении используется для:
  - Классификации изображений
  - Выделения и описания ключевых точек**
  - Сегментации
  - Генерации изображений
- Что такое «аугментация данных»?
  - Удаление шумов
  - Искусственное расширение набора данных**
  - Сжатие изображений
  - Нормализация данных
- Архитектура ResNet ввела концепцию:
  - Сверточных слоев

- б) Skip-connections**
  - в) Attention механизмов
  - г) Рекуррентных сетей
6. Метрика IoU используется для оценки:
- а) Качества классификации
  - б) Точности детекции объектов**
  - в) Скорости работы модели
  - г) Эффективности сегментации
7. Что такое «transfer learning»?
- а) Обучение с нуля
  - б) Использование предобученной модели**
  - в) Самостоятельное обучение
  - г) Обучение без учителя
8. Какой формат разметки используется для детекции объектов?
- а) CSV
  - б) COCO JSON**
  - в) XML
  - г) TXT
9. OpenCV primarily предназначена для:
- а) Работы с базами данных
  - б) Обработки изображений и видео**
  - в) Разработки веб-приложений
  - г) Анализа текстов
10. Что такое «семантическая сегментация»?
- а) Классификация всего изображения
  - б) Присвоение класса каждому пикселю**
  - в) Обнаружение объектов
  - г) Отслеживание движения
11. Алгоритм SURF является:
- а) Методом классификации
  - б) Детектором ключевых точек**
  - в) Методом сегментации
  - г) Алгоритмом сжатия
12. DataLoader в PyTorch используется для:
- а) Визуализации данных
  - б) Пакетной загрузки данных**
  - в) Оптимизации модели
  - г) Разметки данных
13. Что такое «non-maximum suppression»?
- а) Метод аугментации
  - б) Алгоритм отбора bounding box'ов**
  - в) Тип пулинга
  - г) Метод регуляризации
14. Функция потерь Cross-Entropy используется в задачах:
- а) Регрессии
  - б) Классификации**
  - в) Кластеризации
  - г) Детекции
15. Batch Normalization используется для:
- а) Увеличения размера батча
  - б) Стабилизации обучения**
  - в) Ускорения инференса
  - г) Уменьшения модели

### Типовые теоретические вопросы

1. Вопрос: Что такое операция свертки?

Ответ: Линейная операция применения фильтра к изображению для извлечения признаков.

2. Вопрос: В чем разница между семантической и инстанс-сегментацией?  
 Ответ: Семантическая не различает экземпляры, инстанс-различает.
3. Вопрос: Что такое Transfer Learning?  
 Ответ: Использование предобученной модели для решения новой задачи.
4. Вопрос: Для чего используется Non-Maximum Suppression?  
 Ответ: Для отбора наилучших bounding box'ов в детекции объектов.
5. Вопрос: Что такое Data Augmentation?  
 Ответ: Искусственное расширение набора данных через преобразования.
6. Вопрос: Опишите архитектуру CNN.  
 Ответ: Последовательность сверточных, пулинговых и полносвязных слоев.
7. Вопрос: Что такое «feature map»?  
 Ответ: Карта признаков, полученная после применения свертки.
8. Вопрос: Для чего используется OpenCV?  
 Ответ: Библиотека для обработки изображений и компьютерного зрения.
9. Вопрос: Что такое «optical flow»?  
 Ответ: Определение движения объектов между кадрами видео.
10. Вопрос: Опишите метрику IoU.  
 Ответ: Отношение площади пересечения к площади объединения bounding box'ов.
11. Вопрос: Что такое «batch normalization»?  
 Ответ: Нормализация активаций внутри слоя для стабилизации обучения.
12. Вопрос: Для чего используется YOLO?  
 Ответ: Для реального времени детекции объектов.
13. Вопрос: Что такое «image segmentation»?  
 Ответ: Разделение изображения на смысловые области.
14. Вопрос: Опишите работу сверточного слоя.  
 Ответ: Применение фильтров с обучаемыми весами к входному изображению.
15. Что такое «model deployment»?  
 Ответ: Развертывание обученной модели в production-среде.

| Код компетенции | Результаты освоения ОПОП<br>Содержание компетенций  |
|-----------------|---|
| ПК-11           | Способен применять фундаментальные принципы и методы машинного обучения включая подготовку данных оценку качества моделей и работу с признаками |

ПК-11.1. Различает основные типы задач машинного обучения и применяет на практике принципы их решения

ПК-11.2. Применяет методы предварительной обработки данных и работы с признаками

#### Типовые тестовые вопросы:

1. Какая задача НЕ относится к обучению с учителем?
  - а) Классификация
  - б) Регрессия
  - в) Кластеризация**
  - г) Прогнозирование
2. Что такое «переобучение»?
  - а) Модель не обучается
  - б) Модель слишком подстраивается под обучающие данные**
  - в) Модель слишком простая
  - г) Модель не сходится
3. Какая метрика НЕ используется для регрессии?
  - а) MSE
  - б) MAE
  - в) Accuracy**
  - г)  $R^2$



4. Метод главных компонент (PCA) используется для:

- а) Классификации
- б) Снижения размерности**
- в) Кластеризации
- г) Регрессии

5. Что такое «decision tree»?

- а) Линейная модель
- б) Древоподобная структура решений**
- в) Метод кластеризации
- г) Ансамблевый метод

6. Random Forest является:

- а) Одиночной моделью
- б) Ансамблевым методом**
- в) Методом кластеризации
- г) Линейной моделью

7. Регуляризация L1 (Lasso) приводит к:

- а) Увеличению весов
- б) Обнулению части весов**
- в) Увеличению сложности
- г) Уменьшению точности

8. Что такое «grid search»?

- а) Метод визуализации
- б) Поиск оптимальных гиперпараметров**
- в) Алгоритм кластеризации
- г) Метод снижения размерности

9. Метод опорных векторов (SVM) может использоваться для:

- а) Только классификации
- б) Классификации и регрессии**
- в) Только регрессии
- г) Кластеризации

10. Что такое «precision» в бинарной классификации?

- а) Доля правильных предсказаний
- б) Точность среди предсказанных положительных**
- в) Полнота предсказаний
- г) F1-мера

11. Gradient Boosting является:

- а) Методом кластеризации
- б) Ансамблевым методом**
- в) Линейной моделью
- г) Методом снижения размерности

12. Что такое «feature engineering»?

- а) Выбор модели
- б) Создание и преобразование признаков**
- в) Визуализация данных
- г) Разметка данных

13. Какая функция активации НЕ используется в глубоком обучении?

- а) ReLU

- б) Sigmoid
- в) Linear (без активации)**
- г) Tanh

### Типовые теоретические вопросы

1. Вопрос: Что такое «обучение с учителем» (supervised learning)?  
 Ответ: Тип машинного обучения, при котором модель обучается на размеченных данных, где каждому примеру соответствует правильный ответ.
2. Вопрос: В чем разница между классификацией и регрессией?  
 Ответ: Классификация предсказывает дискретные метки классов, а регрессия — непрерывные числовые значения.
3. Вопрос: Что такое «переобучение» (overfitting) и как с ним бороться?  
 Ответ: Когда модель слишком сложна и запоминает шум в данных, а не общие закономерности. Методы борьбы: регуляризация, увеличение данных, упрощение модели, кросс-валидация.
4. Вопрос: Для чего используется кросс-валидация?  
 Ответ: Для более надежной оценки обобщающей способности модели и подбора гиперпараметров без использования тестовой выборки.
5. Вопрос: Опишите принцип работы метода k-ближайших соседей (KNN).  
 Ответ: Алгоритм классифицирует объект based on классы k ближайших к нему объектов в пространстве признаков.
6. Вопрос: Что такое «регуляризация» в машинном обучении?  
 Ответ: Техника добавления штрафа за сложность модели к функции потерь для предотвращения переобучения.
7. Вопрос: В чем разница между L1 и L2 регуляризацией?  
 Ответ: L1 (Lasso) обнуляет некоторые веса, выполняя отбор признаков, а L2 (Ridge) уменьшает все веса, но не обнуляет их.
8. Вопрос: Что такое «дерево решений» (decision tree)?  
 Ответ: Модель, которая принимает решения путем последовательного разделения данных на подмножества based on значений признаков.
9. Вопрос: Как работает метод случайного леса (Random Forest)?  
 Ответ: Ансамблевый метод, который строит множество деревьев на случайных подвыборках данных и признаков, а затем усредняет их предсказания.
10. Вопрос: Что такое «градиентный бустинг» (Gradient Boosting)?  
 Ответ: Ансамблевый метод, который последовательно строит модели, где каждая следующая модель исправляет ошибки предыдущих.
11. Вопрос: Для чего используется метод главных компонент (PCA)?  
 Ответ: Для снижения размерности данных путем проецирования их в пространство меньшей размерности с сохранением максимальной дисперсии.
12. Вопрос: Что такое «метрики качества» в машинном обучении?  
 Ответ: Численные показатели, которые оценивают производительность модели на тестовых данных.
13. Вопрос: В чем разница между precision и recall?  
 Ответ: Precision — доля истинно положительных среди всех предсказанных положительных, recall — доля истинно положительных среди всех реально положительных.
14. Вопрос: Что такое «confusion matrix» (матрица ошибок)?  
 Ответ: Таблица, которая показывает количество правильных и неправильных предсказаний модели для каждого класса.
15. Вопрос: Для чего используется метод k-средних (k-means)?  
 Ответ: Для кластеризации данных — разбиения на k групп based on близости объектов к центроидам кластеров.

| Код компетенции | Результаты освоения ОПОП<br>Содержание компетенций                    |
|-----------------|---|
| ПК-14           | Способен применять алгоритмы обучения на нестандартных объемах данных |

ПК-14.1. Обосновывает способы и варианты применения алгоритмов обучения на нестандартных объемах данных в задачах ИИ, включая их преобразование и адаптацию к специфике задачи

ПК-14.2. Применяет методы повышения устойчивости, надежности, безопасности алгоритмов обучения на нестандартных объемах данных для проверки разведочных гипотез и подготовки данных к применению современных методов ИИ

**Типовые тестовые вопросы:**

1. Какая задача НЕ относится к обучению с учителем?

- а) Классификация
- б) Регрессия
- в) Кластеризация**
- г) Прогнозирование

2. Что такое «несбалансированные данные»?

- а) Данные с пропусками
- б) Неравномерное распределение классов**
- в) Данные с шумом
- г) Высокомерные данные

3. Метод SMOTE используется для:

- а) Удаления шума
- б) Генерации примеров миноритарного класса**
- в) Снижения размерности
- г) Нормализации данных

4. Что такое «active learning»?

- а) Обучение без учителя
- б) Стратегия выбора наиболее информативных данных**
- в) Обучение с подкреплением
- г) Глубокое обучение

5. Какая проблема характерна для работы с текстовыми данными?

- а) Низкая размерность
- б) Высокая разреженность**
- в) Отсутствие признаков
- г) Необходимость аугментации

6. Метод Isolation Forest используется для:

- а) Классификации
- б) Обнаружения аномалий**
- в) Кластеризации
- г) Регрессии

7. Что такое «transfer learning» в контексте NLP?

- а) Обучение с нуля
- б) Использование предобученных языковых моделей**
- в) Ручное кодирование
- г) Статистические методы

8. Какая техника НЕ используется для работы с несбалансированными данными?

- а) Взвешивание классов
- б) Стратифицированная выборка
- в) Увеличение размера мажоритарного класса**
- г) Изменение порога классификации

9. Метод t-SNE используется для:

- а) Классификации
- б) Визуализации многомерных данных**
- в) Кластеризации
- г) Регрессии

10. Что такое «few-shot learning»?

- а) Обучение на большом наборе данных
- б) Обучение на небольшом количестве примеров**
- в) Обучение без разметки
- г) Обучение с подкреплением

11. Какая проблема характерна для временных рядов?

- а) Отсутствие тренда
- б) Автокорреляция**
- в) Низкая размерность
- г) Необходимость кластеризации

12. Метод MICE используется для:

- а) Обнаружения выбросов
- б) Заполнения пропущенных значений**
- в) Снижения размерности
- г) Нормализации

13. Что такое «domain adaptation»?

- а) Адаптация модели к новому распределению данных**
- б) Изменение архитектуры модели
- в) Обучение с подкреплением
- г) Аугментация данных

14. Какая метрика НЕ подходит для несбалансированных данных?

- а) F1-score
- б) Precision-Recall AUC
- в) Accuracy**
- г) Matthews correlation coefficient

15. Что такое «covariate shift»?

- а) Изменение распределения целевой переменной
- б) Изменение распределения признаков**
- в) Изменение размера данных
- г) Сдвиг во времени

### Типовые теоретические вопросы

1. Вопрос: Что такое «несбалансированные данные»?

Ответ: Данные с неравномерным распределением классов, когда один класс значительно преобладает над другими.

2. Вопрос: Для чего используется метод SMOTE?

Ответ: Для генерации синтетических примеров миноритарного класса и балансировки набора данных.

3. Вопрос: Что такое «Active Learning» (Активное обучение)?

Ответ: Стратегия машинного обучения, при которой модель выбирает наиболее информативные данные для разметки, чтобы уменьшить общий объем ручной разметки.

4. Вопрос: Опишите проблему «разреженности данных».

Ответ: Ситуация, когда большинство признаков в данных имеют нулевые значения, что характерно для текстовых данных после векторизации.

5. Вопрос: Что такое «Data Drift» (Концептуальный дрейф)?

Ответ: Изменение статистических свойств и распределения входных данных с течением

времени, приводящее к снижению качества модели.

6. Вопрос: Для чего используется метод Isolation Forest?

Ответ: Для обнаружения аномалий путем изоляции наблюдений в случайных подпространствах признаков.

7. Вопрос: Что такое «Few-Shot Learning» (Обучение с малым количеством примеров)?

Ответ: Подход, позволяющий модели быстро обучаться на очень небольшом количестве размеченных примеров для нового класса.

8. Вопрос: Опишите метод MICE для обработки пропущенных значений.

Ответ: Многократное вменение с помощью цепных уравнений — итеративный метод, который предсказывает пропуски на основе других признаков.

9. Вопрос: Что такое «Domain Adaptation» (Адаптация домена)?

Ответ: Задача адаптации модели, обученной на исходном домене (распределении данных), для эффективной работы на целевом домене с другим распределением.

10. Вопрос: Для чего используется метрика F1-score вместо ассигасы на несбалансированных данных?

Ответ: F1-score является гармоническим средним precision и recall и лучше отражает качество модели на миноритарном классе, тогда как ассигаса может вводить в заблуждение.

11. Вопрос: Что такое «Covariate Shift» (Ковариатный сдвиг)?

Ответ: Изменение распределения входных переменных (признаков) при неизменном распределении выходной переменной.

12. Вопрос: Опишите технику «Stratified Sampling» (Стратифицированная выборка).

Ответ: Метод разделения данных, при котором в обучающей и тестовой выборках сохраняется исходное распределение классов.

13. Вопрос: Что такое «Transfer Learning» в контексте NLP?

Ответ: Использование языковых моделей (например, BERT, GPT), предобученных на больших корпусах текстов, для решения конкретных задач с минимальным дообучением.

14. Вопрос: Для чего используется метод t-SNE?

Ответ: Для визуализации высокоразмерных данных путем проецирования их в пространство меньшей размерности (2D или 3D) с сохранением локальных расстояний между точками.

15. Вопрос: Что такое «Class Weighting» (Взвешивание классов)?

Ответ: Техника борьбы с дисбалансом, при которой в функции потерь примерам миноритарного класса присваивается больший вес, чтобы модель уделяла им больше внимания.

### Типовые теоретические вопросы

| Код компетенции | Результаты освоения ОПОП<br>Содержание компетенций  |
|-----------------|---|
| ПК-16           | Способен проводить фронтальные исследования в области архитектур, алгоритмов МО, оптимизации и математики |

ПК-16.1. Разрабатывает фундаментальные основы и новые алгоритмы машинного обучения

ПК-16.2. Разрабатывает новые архитектуры глубоких нейросетей

### Типовые тестовые вопросы:

1. Что такое «нейросетевой поиск архитектур» (NAS)?

а) Ручной подбор архитектуры

**б) Автоматизированный поиск оптимальной архитектуры**

в) Обучение готовых архитектур

г) Визуализация архитектур

2. Какая архитектура НЕ является трансформером?

а) BERT

б) GPT

**в) ResNet**

г) T5

3. Что такое «attention mechanism»?
- а) Метод регуляризации
  - б) Механизм взвешивания важности элементов**
  - в) Функция активации
  - г) Метод оптимизации
4. GAN (Generative Adversarial Network) состоит из:
- а) Двух генераторов
  - б) Генератора и дискриминатора**
  - в) Двух дискриминаторов
  - г) Кодировщика и декодера
5. Что такое «reinforcement learning»?
- а) Обучение с учителем
  - б) Обучение через взаимодействие со средой**
  - в) Обучение без учителя
  - г) Полуавтоматическое обучение
6. Метод Batch Normalization служит для:
- а) Увеличения размера батча
  - б) Стабилизации распределения активаций**
  - в) Ускорения инференса
  - г) Уменьшения модели
7. Что такое «transformer architecture»?
- а) Сверточная сеть
  - б) Архитектура на основе attention**
  - в) Рекуррентная сеть
  - г) Автокодировщик
8. Какая функция потерь используется в GAN?
- а) Cross-Entropy
  - б) Adversarial loss**
  - в) MSE
  - г) MAE
9. Что такое «meta-learning»?
- а) Обучение одной задачи
  - б) Обучение учиться**
  - в) Обучение без данных
  - г) Обучение с подкреплением
10. Архитектура Vision Transformer (ViT) применяется для:
- а) Обработки текста
  - б) Обработки изображений**
  - в) Обработки звука
  - г) Работы с табличными данными
11. Что такое «neural architecture search»?
- а) Ручной подбор параметров
  - б) Автоматический поиск архитектуры**
  - в) Обучение модели
  - г) Визуализация графа
12. Какая задача НЕ решается с помощью GAN?
- а) Генерация изображений
  - б) Стилизация**

**в) Классификация**

г) Увеличение данных

13. Что такое «policy gradient» в RL?

а) Метод оптимизации

**б) Алгоритм обучения с подкреплением**

в) Метод регуляризации

г) Функция потерь

14. Архитектура U-Net используется для:

а) Классификации

**б) Сегментации**

в) Детекции

г) Генерации

**Типовые теоретические вопросы**

1. Вопрос: Что такое «Neural Architecture Search» (NAS)?

Ответ: Процесс автоматизированного поиска оптимальной архитектуры нейронной сети для конкретной задачи и набора данных.

2. Вопрос: Опишите принцип работы «Attention Mechanism» (Механизма внимания).

Ответ: Механизм, который позволяет модели динамически взвешивать и комбинировать информацию из разных частей входной последовательности, фокусируясь на наиболее релевантных элементах.

3. Вопрос: Из каких компонентов состоит Generative Adversarial Network (GAN)?

Ответ: Из двух нейронных сетей: Генератора, который создает синтетические данные, и Дискриминатора, который отличает реальные данные от сгенерированных.

4. Вопрос: Что такое «Reinforcement Learning» (Обучение с подкреплением)?

Ответ: Раздел машинного обучения, в котором агент обучается принимать решения путем взаимодействия со средой и получая награду за правильные действия.

5. Вопрос: Опишите архитектурный блок «Transformer».

Ответ: Архитектура, основанная на механизме самовнимания (self-attention), которая позволяет обрабатывать последовательности данных без использования рекуррентных или сверточных слоев.

6. Вопрос: Для чего используется «Batch Normalization» (Пакетная нормализация)?

Ответ: Для стабилизации и ускорения обучения глубоких нейронных сетей путем нормализации активаций предыдущего слоя для каждого мини-батча.

7. Вопрос: Что такое «Meta-Learning» (Мета-обучение)?

Ответ: Подход «обучение учиться», при котором модель развивает способность быстро адаптироваться к новым задачам на основе небольшого количества примеров.

8. Вопрос: Опишите принцип «Policy Gradient» в обучении с подкреплением.

Ответ: Метод, который оптимизирует политику агента напрямую, поднимаясь в направлении градиента ожидаемой награды.

9. Вопрос: Что такое «Vision Transformer» (ViT)?

Ответ: Архитектура, которая применяет модель Transformer к задачам компьютерного зрения, разбивая изображение на патчи и обрабатывая их как последовательность токенов.

10. Вопрос: В чем заключается «Exploration-Exploitation Trade-off» (Дилемма «исследование-использование»)?

Ответ: Фундаментальная дилемма в RL, где агент должен балансировать между исследованием новых действий (для сбора информации) и использованием известных действий (для максимизации награды).

11. Вопрос: Что такое «Diffusion Models» (Диффузионные модели)?

Ответ: Генеративные модели, которые создают данные путем постепенного удаления шума из начального случайного распределения, обучаясь обратному процессу добавления шума к данным.

12. Вопрос: Опишите концепцию «Self-Supervised Learning» (Самообучение).

Ответ: Подход, при котором модель обучается, извлекая сигналы для обучения непосредственно из неразмеченных входных данных, создавая для себя задачи (например,

предсказание скрытой части входных данных).

13. Вопрос: Для чего используется «Graph Neural Network» (GNN)?

Ответ: Для обработки данных, представленных в виде графов, позволяя модели учитывать не только признаки узлов, но и связи между ними.

14. Вопрос: Что такое «Neural Ordinary Differential Equations» (Neural ODEs)?

Ответ: Модель, которая рассматривает глубину сети как непрерывную переменную, описываемую обыкновенным дифференциальным уравнением, что позволяет создавать более эффективные и емкие архитектуры.

15. Вопрос: Опишите принцип работы «Siamese Neural Network» (Сиамской нейронной сети).

Ответ: Архитектура, состоящая из двух или более идентичных подсетей, которые разделяют веса, используется для вычисления схожести между входными образцами (например, для верификации лиц).

### Типовые теоретические вопросы:

- 1) Опишите архитектуры генерации текста: от RNN до Transformer-based моделей (GPT, BERT, T5).
- 2) В чём преимущества fine-tuning предобученных моделей для задач генерации текста?
- 3) Какие методы используются для генерации и редактирования изображений (GAN, Diffusion, Neural Style Transfer)?
- 4) Как осуществляется перенос стиля в реальном времени на видео с использованием OpenCV и PyTorch?
- 5) Какие типы кастомных слоёв нейронных сетей вы знаете? Приведите примеры.
- 6) Как реализуется слой линейной регрессии с обучаемыми параметрами?
- 7) Для каких задач применяются спектральные преобразования и вейвлет-слои?
- 8) Как проектировать и обучать кастомные слои в PyTorch/TensorFlow?
- 9) Опишите методы трекинга объектов в видеопотоке: BlobTracking, SORT, DeepSORT.
- 10) Как реализовать систему подсчёта людей в видеопотоке с использованием YOLO и DeepSORT?
- 11) Какие подходы используются для предсказания следующего кадра видео (например, на основе ConvLSTM)?
- 12) Как решаются задачи стабилизации и синтеза движений в видео?
- 13) Какие методы сжатия и ускорения моделей машинного обучения вы знаете?
- 14) Что такое MLOps? Опишите ключевые компоненты: контейнеризация, мониторинг, управление экспериментами.
- 15) Как развернуть OCR-систему для распознавания текста на упаковках товаров?
- 16) Какие методы активного обучения применяются для обнаружения дефектов на изображениях?
- 17) Какие продвинутое методы аугментации данных используются для борьбы с переобучением?
- 18) Как решаются задачи идентификации по лицу, определения пола, возраста, направления взгляда?
- 19) В чём особенности обучения на несбалансированных данных?
- 20) Какие методы используются для повышения устойчивости и надежности моделей на нестандартных данных?
- 21) Какие основные типы задач машинного обучения вы знаете? Приведите примеры алгоритмов для каждого типа.
- 22) Как выбрать подходящий алгоритм машинного обучения для конкретной бизнес-задачи?
- 23) Опишите процесс построения end-to-end пайплайна от данных до продакшена.
- 24) Какие инструменты и фреймворки вы использовали для развертывания ML-моделей? (Docker, MLflow, TensorFlow Serving и др.)
- 25) Как вы проводите валидацию моделей и интерпретируете метрики качества?

Оператор ЭДО ООО "Компания "Тензор"

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ

СОГЛАСОВАНО

ФГБОУ ВО "РГРТУ", РГРТУ, Костров Борис Васильевич,  
Заведующий кафедрой ЭВМ

04.12.25 16:35 (MSK)

Простая подпись