МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Рязанский государственный радиотехнический университет»

Кафедра «Безопасность жизнедеятельности и экология»

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

по дисциплине

«Автоматизированное машинное обучение»

Направление

09.03.04 Программная инженерия

Профиль

Программная инженерия систем искусственного интеллекта
Программное обеспечение систем искусственного интеллекта

Квалификация выпускника – бакалавр

Форма обучения – очная

Рязань

1. ОБЩИЕ ПОЛОЖЕНИЯ

Оценочные материалы – это совокупность учебно-методических материалов и процедур, предназначенных для оценки качества освоения обучающимися данной дисциплины как части основной образовательной программы.

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

Основная задача – обеспечить оценку уровня сформированности компетенций и индикаторов их достижения, приобретаемых обучающимися в соответствии с требованиями ФГОС ВО.

Контроль знаний обучающихся проводится в форме текущего контроля (лабораторные и практические работы) и промежуточной аттестации (зачёт).

2. ПЕРЕЧЕНЬ КОМПЕТЕНЦИЙ С УКАЗАНИЕМ ЭТАПОВ ИХ ФОРМИРОВАНИЯ

При освоении дисциплины формируются следующие компетенции:

- **ПК-1** – способен проводить передовые исследования в области архитектур, алгоритмов МО, оптимизации и математики (FC-1);
  - **ПК-1.1** – разрабатывает фундаментальные основы и новые алгоритмы машинного обучения (FC-1.1).
- **ПК-4** – способен определять требования к наборам данных для решения задач машинного обучения, проводить разметку и анализ наборов данных, оценивать качество данных, обеспечивать непрерывную интеграцию данных (BD-2);
  - **ПК-4.1** – определяет требования к наборам и качеству данных для решения задач машинного обучения (BD-2.1);
  - **ПК-4.2** – работает с данными, в том числе собирает данные из разрозненных источников, проверяет данные на корректность (BD-2.2);
  - **ПК-4.3** – применяет инструменты и практики непрерывной интеграции данных (DataOps) (BD-2.3).
- **ПК-8** – способен осуществлять поиск, сбор, очистку и предварительный анализ данных (BD-1);
  - **ПК-8.1** – обосновывает способы и варианты применения методов предварительного анализа данных в задачах ИИ (BD-1.2);
  - **ПК-8.3** – применяет методы понижения размерности для первичной интерпретации и визуализации многомерных данных (BD-1.4);
  - **ПК-8.4** – отбирает признаки данных, значимые для исследования (BD-1.5).

- **ПК-11** – способен применять фундаментальные принципы и методы машинного обучения, включая подготовку данных, оценку качества моделей и работу с признаками (ML-2);
  - **ПК-11.1** – различает основные типы задач машинного обучения и применяет на практике принципы их решения (ML-2.1);
  - **ПК-11.2** – применяет методы предварительной обработки данных и работы с признаками (ML-2.2).
- **ПК-15** (индикаторы):
  - **ПК-15.1** – разрабатывает и отлаживает прикладные решения с использованием языка Python (PL-1.1);
  - **ПК-15.2** – осуществляет выбор инструментов разработки на Python (PL-1.2).
- **ПК-22** – способен разрабатывать и (или) применять методы повышения устойчивости, надёжности, безопасности алгоритмов МО (ML-5);
  - **ПК-22.1** – обосновывает способы и варианты применения методов повышения устойчивости, надёжности, безопасности алгоритмов МО в задачах ИИ (ML-5.1);
  - **ПК-22.2** – применяет методы повышения устойчивости, надёжности, безопасности алгоритмов МО для проверки разведочных гипотез и подготовки данных (ML-5.2).
- **ПК-23** (индикаторы):
  - **ПК-23.1** – обосновывает способы и варианты применения алгоритмов автоматического машинного обучения в задачах ИИ (ML-7.1);
  - **ПК-23.2** – применяет методы повышения устойчивости, надёжности, безопасности алгоритмов автоматического ML (ML-7.2).
- **ПК-24** – способен проводить эксперименты на данных, формулировать гипотезы исследования, строить (обучать, дообучать) модели ИИ с оценкой их качества и анализом ошибок, обеспечивать воспроизводимость и масштабируемость исследований на данных (LC-2);
  - **ПК-24.1** – проводит эксперименты с моделями ИИ, оценивает их качество (точность, производительность) (LC-2.1);
  - **ПК-24.2** – проводит эксперименты на данных и визуализирует результаты с применением технологий анализа данных, методов и алгоритмов МО (LC-2.2).

Указанные компетенции формируются в соответствии со следующими этапами:

- формирование и развитие теоретических знаний (лекционные занятия, самостоятельная работа);
- приобретение и развитие практических умений (лабораторные и практические работы, самостоятельная работа);
- закрепление знаний и навыков (выполнение и защита лабораторных работ, подготовка к зачёту).

3. ПОКАЗАТЕЛИ И КРИТЕРИИ ОЦЕНИВАНИЯ КОМПЕТЕНЦИЙ

Сформированность каждой компетенции оценивается по трёхуровневой шкале:

- **пороговый уровень** – обязательный для всех обучающихся; соответствует освоению более 40% приведённых знаний, умений и навыков;
- **продвинутый уровень** – соответствует освоению более 60% знаний, умений и навыков;
- **эталонный уровень** – соответствует освоению более 80% знаний, умений и навыков.

При освоении менее 40% компетенция считается несформированной.

Таблица 1. Показатели достижения индикаторов компетенций

Компетенция	Индикаторы (краткое содержание)	Этап	Оценочное средство
ПК-1	**ПК-1.1** – знать принципы разработки современных моделей МО; уметь разрабатывать модели, опираясь на передовой опыт; владеть навыками разработки с учётом актуальных конференций.	6	Зачёт
ПК-4	**ПК-4.1** – знать принципы применения статистических критериев; уметь выбирать и обосновывать требования к данным; владеть навыками.	6	Зачёт
ПК-4	**ПК-4.2** – знать способы и инструменты сбора данных и проверки корректности; уметь собирать данные из разрозненных источников; владеть навыками.	6	Зачёт
ПК-4	**ПК-4.3** – знать способы, инструменты и практики непрерывной интеграции данных; уметь применять их; владеть навыками.	6	Зачёт
ПК-8	**ПК-8.1** – знать способы и варианты применения методов разведочного анализа данных; уметь применять; владеть навыками с преобразованием и адаптацией.	6	Зачёт

Компетенция	Индикаторы (краткое содержание)	Этап	Оценочное средство
ПК-8	**ПК-8.3** – знать методы линейного и нелинейного снижения размерности; уметь применять и выбирать подходящие; владеть навыками визуализации.	6	Зачёт
ПК-8	**ПК-8.4** – знать методы отбора признаков; уметь применять и проверять их значимость; владеть навыками.	6	Зачёт
ПК-11	**ПК-11.1** – знать методы решения задач МО с учётом данных и контекста; уметь настраивать базовые модели и оценивать; владеть навыками.	6	Зачёт
ПК-11	**ПК-11.2** – знать методы инженерии признаков; уметь применять для отбора, создания и преобразования признаков; владеть навыками.	6	Зачёт
ПК-15	**ПК-15.1** – знать Python и библиотеки NumPy, SciPy, Matplotlib, Pandas, Keras, TensorFlow; уметь создавать программы для задач ИИ; владеть навыками.	6	Зачёт
ПК-15	**ПК-15.2** – знать подходы к разработке собственных компонентов для библиотек МО; уметь разрабатывать; владеть навыками.	6	Зачёт
ПК-22	**ПК-22.1** – знать методы повышения устойчивости и надёжности моделей; уметь обосновывать выбор методов; владеть навыками оценки рисков.	6	Зачёт
ПК-22	**ПК-22.2** – знать продвинутые методы дообучения моделей при подготовке	6	Зачёт

Компетенция	Индикаторы (краткое содержание)	Этап	Оценочное средство
	данных; уметь применять; владеть навыками.		
ПК-23	**ПК-23.1** – знать инструменты анализа специфики задачи с учётом современных трендов (AutoML, MLOps); уметь выбирать AutoML-инструменты; владеть навыками.	6	Зачёт
ПК-23	**ПК-23.2** – знать продвинутые методы повышения устойчивости моделей AutoML; уметь применять; владеть навыками.	6	Зачёт
ПК-24	**ПК-24.1** – знать инструменты автоматизации проведения экспериментов; уметь применять; владеть навыками.	6	Зачёт
ПК-24	**ПК-24.2** – знать инструменты оценивания качества данных; уметь применять с учётом специфики моделей ИИ; владеть навыками.	6	Зачёт

4. ПАСПОРТ ОЦЕНОЧНЫХ МАТЕРИАЛОВ ПО ДИСЦИПЛИНЕ

Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции (или её части)	Наименование оценочного средства
Тема 1. Введение в AutoML. Подготовка данных	ПК-1.1, ПК-4.1, ПК-4.2, ПК-4.3, ПК-8.1, ПК-15.1, ПК-24.1, ПК-24.2	Зачёт
Тема 2. Разведочный анализ данных	ПК-8.1, ПК-8.2, ПК-8.3, ПК-15.1, ПК-24.1, ПК-24.2	Зачёт
Тема 3. Конструирование признаков	ПК-4.3, ПК-8.3, ПК-8.4, ПК-11.2, ПК-15.1, ПК-22.1, ПК-22.2	Зачёт

Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции (или её части)	Наименование оценочного средства
Тема 4. Генерация модели. Оптимизация гиперпараметров	ПК-1.1, ПК-11.1, ПК-15.1, ПК-24.1, ПК-24.2, ПК-23.1, ПК-23.2	Зачёт
Тема 5. Оценка модели	ПК-1.1, ПК-11.1, ПК-15.1, ПК-22.1, ПК-22.2	Зачёт
Тема 6. AutoML сервисы	ПК-11.1, ПК-15.2, ПК-23.1, ПК-23.2	Зачёт
Тема 7. Конвейерная разработка модели: финализация	ПК-1.1, ПК-4.2, ПК-11.1, ПК-15.1	Зачёт

5. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ

5.1. Промежуточная аттестация (зачёт)

Зачёт проводится в форме устного ответа по билетам, включающим теоретический вопрос и практическое задание (задачу). Также может использоваться тестирование.

5.1.1. Тестовые вопросы закрытого типа (20 вопросов)

1. Что такое автоматизированное машинное обучение (AutoML)?
а) Процесс ручного подбора моделей и гиперпараметров;
б) Процесс автоматизации выбора моделей, гиперпараметров и предобработки данных;
в) Метод обучения с подкреплением;
г) Алгоритм кластеризации.
Правильный ответ: б)

2. Какой этап не относится к типичному конвейеру AutoML?
а) Сбор и предобработка данных;
б) Генерация моделей и оптимизация гиперпараметров;
в) Написание кода вручную без использования библиотек;
г) Оценка модели и развертывание.
Правильный ответ: в)

3. Что означает аббревиатура TPOT?
а) Tree-based Pipeline Optimization Tool;
б) Total Performance Optimization Technique;
в) Time-based Processing Optimization Tool;

г) Tensor Processing Optimization Technology.
Правильный ответ: а)

4. Для чего применяются скейлеры при подготовке данных?
а) Для удаления выбросов;
б) Для приведения признаков к единому масштабу;
в) Для заполнения пропусков;
г) Для генерации новых признаков.
Правильный ответ: б)

5. Что такое ранняя остановка (early stopping)?
а) Остановка обучения при достижении максимального числа итераций;
б) Остановка обучения при ухудшении качества на валидационной выборке;
в) Остановка после первой эпохи;
г) Остановка при достижении нулевой ошибки на обучении.
Правильный ответ: б)

6. Что такое дрейф данных (data drift)?
а) Изменение распределения данных с течением времени;
б) Ошибка в коде;
в) Переобучение модели;
г) Метод оптимизации.
Правильный ответ: а)

7. Какой метод не является методом оптимизации гиперпараметров?
а) Поиск по сетке (Grid Search);
б) Случайный поиск (Random Search);
в) Байесовская оптимизация;
г) Метод главных компонент (PCA).
Правильный ответ: г)

8. Какой инструмент является библиотекой для AutoML?
а) Scikit-learn;
б) Auto-sklearn;
в) Matplotlib;
г) NumPy.
Правильный ответ: б)

9. Что такое конструирование признаков (feature engineering)?
а) Удаление ненужных столбцов;
б) Создание новых признаков на основе существующих;
в) Разделение данных на train и test;
г) Обучение модели.
Правильный ответ: б)

10. Что такое суррогатная модель (surrogate model) в AutoML?
а) Модель, заменяющая сложную целевую функцию для ускорения оптимизации;

б) Модель, которая всегда дает точный результат;
в) Модель, не требующая данных;
г) Модель, работающая только с изображениями.
Правильный ответ: а)

11. Какова основная цель AutoML?
а) Полностью заменить человека в разработке ИИ;
б) Снизить порог входа и автоматизировать рутинные задачи ML;
в) Увеличить время обучения;
г) Усложнить архитектуру моделей.
Правильный ответ: б)

12. В чём отличие поиска по сетке от случайного поиска?
а) Поиск по сетке перебирает все комбинации, случайный — выбирает случайные;
б) Случайный поиск перебирает все комбинации, поиск по сетке — случайные;
в) Они полностью идентичны;
г) Поиск по сетке быстрее случайного.
Правильный ответ: а)

13. Какой AutoML-сервис предоставляется компанией Google?
а) IBM Watson;
б) AzureML;
в) Google Cloud AutoML;
г) MLJAR.
Правильный ответ: в)

14. Что включает в себя предварительная обработка данных?
а) Только обучение модели;
б) Очистку, заполнение пропусков и масштабирование;
в) Только визуализацию;
г) Только развертывание.
Правильный ответ: б)

15. Что означает «оптимизация с учётом ресурсов» (resource-aware) в AutoML?
а) Ограничение времени и памяти при поиске моделей;
б) Использование только облачных ресурсов;
в) Отказ от оптимизации;
г) Увеличение размера модели.
Правильный ответ: а)

16. Зачем нужна визуализация в разведочном анализе данных?
а) Для поиска закономерностей, выбросов и аномалий;
б) Для ускорения обучения;
в) Для уменьшения размера данных;
г) Для шифрования данных.
Правильный ответ: а)

17. Что такое кросс-валидация?
а) Обучение модели на всех данных;
б) Разбиение данных на k частей с поочередным использованием каждой как тестовой;
в) Удаление выбросов;
г) Генерация признаков.
Правильный ответ: б)

18. Какова цель отбора признаков (feature selection)?
а) Удаление неинформативных признаков для улучшения модели;
б) Добавление всех возможных признаков;
в) Замена модели;
г) Изменение формата данных.
Правильный ответ: а)

19. Что такое конвейер (pipeline) в машинном обучении?
а) Последовательность шагов обработки данных и обучения модели;
б) Отдельный алгоритм;
в) Тип нейронной сети;
г) Метрика качества.
Правильный ответ: а)

20. Для чего используется мониторинг модели после развертывания?
а) Для обнаружения деградации качества и дрейфа данных;
б) Для повторного обучения;
в) Для удаления модели;
г) Для изменения гиперпараметров.
Правильный ответ: а)

5.1.2. Тестовые вопросы открытого типа (10 вопросов)

1. Дайте определение AutoML и опишите его основные компоненты.

2. Перечислите и охарактеризуйте основные этапы подготовки данных в конвейере AutoML.

3. Сравните методы оптимизации гиперпараметров: поиск по сетке, случайный поиск и байесовскую оптимизацию.

4. Что такое конструирование признаков? Перечислите основные методы и приведите примеры.

5. Объясните суть суррогатного метода (surrogate method) в контексте AutoML.

6. Опишите функциональные возможности инструментов AutoML (на примере Auto-sklearn или TPOT).

7. Какие проблемы возникают при развертывании AutoML-моделей в производственную среду?

8. Как оценивается качество модели в AutoML? Перечислите основные метрики и подходы.

9. Что такое дрейф данных и как осуществляется его мониторинг?

10. Сравните облачные AutoML-сервисы (Google Cloud AutoML, AzureML) с открытыми библиотеками (Auto-sklearn, TPOT).

5.1.3. Типовые практические задания (задачи) к зачёту

1. Выполнить предварительную обработку данных (заполнение пропусков, масштабирование) для заданного набора данных на Python.
2. Провести разведочный анализ данных (EDA) и визуализировать результаты (гистограммы, boxplot, корреляционная матрица).
3. Применить метод отбора признаков (например, SelectKBest или RFE) и обосновать выбор признаков.
4. Использовать GridSearchCV для подбора гиперпараметров классификатора Random Forest.
5. Применить библиотеку Optuna для байесовской оптимизации гиперпараметров модели.
6. Построить базовый конвейер AutoML с использованием библиотеки Auto-sklearn на заданном датасете.
7. Сравнить результаты автоматического подбора моделей с помощью TPOT и ручного подбора.
8. Обучить модель и сохранить её для развертывания (с использованием joblib или pickle).
9. Смоделировать дрейф данных и продемонстрировать подходы к его обнаружению.
10. Использовать облачный AutoML-сервис (или его локальный аналог) для обучения модели и оценки её качества.

5.1.4. Вопросы к зачёту (30 вопросов)

1. Понятие автоматизированного машинного обучения (AutoML). Основные предпосылки возникновения.
2. Основные этапы конвейера AutoML.
3. Сбор и предварительная обработка данных в AutoML.
4. Обработка пропущенных значений и выбросов.
5. Масштабирование и нормализация данных.
6. Разведочный анализ данных: цели и методы.
7. Анализ количественных данных.
8. Анализ категориальных данных.
9. Графические методы в разведочном анализе данных.
10. Конструирование признаков: извлечение, отбор и генерация.

11. Методы отбора признаков.
12. Линейное снижение размерности: PCA, LDA.
13. Нелинейное снижение размерности: t-SNE, UMAP.
14. Генерация моделей машинного и глубокого обучения в AutoML.
15. Оптимизация гиперпараметров: поиск по сетке и случайный поиск.
16. Байесовская оптимизация гиперпараметров.
17. Эволюционные и популяционные алгоритмы оптимизации.
18. Обучение с подкреплением в AutoML.
19. Инструменты AutoML: TPOT, Auto-WEKA, Auto-sklearn.
20. Инструменты AutoML: Auto-Keras, MLBox, TransmogrifAI, Optuna.
21. Оценка модели: методы низкой точности (low fidelity).
22. Суррогатные методы оценки моделей.
23. Ранняя остановка (early stopping) в AutoML.
24. Оптимизация с учётом ресурсов (resource-aware).
25. AutoML-сервисы: Google Cloud AutoML, IBM Watson.
26. AutoML-сервисы: MLJAR, AzureML.
27. Сравнительный анализ возможностей AutoML-сервисов.
28. Финализация и развертывание модели.
29. Мониторинг и обслуживание моделей AutoML.
30. Обнаружение дрейфа данных и концептуального дрейфа.

6. КРИТЕРИИ ОЦЕНИВАНИЯ

6.1. Шкала оценивания тестирования

Шкала	Критерий
3 балла (эталонный уровень)	Процент верных ответов от 85 до 100%
2 балла (продвинутый уровень)	Процент верных ответов от 75 до 84%
1 балл (пороговый уровень)	Процент верных ответов от 60 до 74%
0 баллов	Процент верных ответов от 0 до 59%

6.2. Шкала оценивания теоретического вопроса

Шкала	Критерий
3 балла (эталонный уровень)	Дан полный ответ, показаны глубокие систематизированные знания, приведены примеры, даны ответы на дополнительные вопросы.
2 балла (продвинутый уровень)	Дан полный ответ, но на некоторые дополнительные вопросы ответы даны с помощью наводящих вопросов.
1 балл (пороговый уровень)	Дан неполный ответ, на дополнительные вопросы ответы даны только с помощью преподавателя.
0 баллов	Студент не смог ответить на вопрос.

6.3. Шкала оценивания практического задания

Шкала	Критерий
3 балла (эталонный уровень)	Задание решено верно, получены корректные результаты, сделаны выводы.
2 балла (продвинутый уровень)	Задание решено верно, но имеются незначительные технические неточности.
1 балл (пороговый уровень)	Задание решено верно с дополнительными наводящими вопросами преподавателя.
0 баллов	Задание не решено.

6.4. Итоговая оценка

На зачёт выносится тестовое задание, практическое задание и теоретический вопрос. Максимальный балл – 9.
Оценка «зачтено» выставляется при сумме баллов не менее 5 и выполнении всех предусмотренных лабораторных и практических работ.
Оценка «не зачтено» – при сумме менее 5 баллов или наличии несданных лабораторных работ.

Оператор ЭДО ООО "Компания "Тензор"
ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ
СОГЛАСОВАНО **ФГБОУ ВО "РГРТУ", РГРТУ,** Овечкин Геннадий Владимирович, Заведующий кафедрой ВПМ **05.11.26** 14:46 (MSK) Простая подпись