

Контрольные и тестовые вопросы по ПР1

«Знакомство с инструментом RapidMiner» по вариантам с ответами

Вариант 1

1. Что из нижеперечисленного наилучшим образом описывает основную концепцию RapidMiner Studio?
А) Язык программирования для автоматизации обработки данных
В) Графический интерфейс для разработки программ на Python
С) Платформа для визуального проектирования аналитических процессов без программирования
D) Среда анализа временных рядов в Excel

Ответ: С

2. Для чего используется панель «Operators» в интерфейсе RapidMiner?
А) Для отображения входных данных
В) Для выполнения математических вычислений
С) Для выбора и добавления компонентов в аналитический процесс
D) Для управления лицензией

Ответ: С

3. Какую функцию выполняет инструмент TurboPrep?
А) Объединение моделей
В) Подготовка данных без необходимости ручного программирования
С) Тестирование гипотез
D) Создание визуализаций вручную

Ответ: В

4. Что произойдет, если не определить целевую переменную в процессе обучения модели в AutoModel?
А) Автоматически выберется случайный признак
В) Модель не будет построена
С) RapidMiner выполнит классификацию по всем признакам
D) Построится модель без обучения

Ответ: В

5. В чем преимущество AutoModel по сравнению с ручным построением процессов?
- A) Он требует больше вычислительных ресурсов
 - B) Он исключает возможность настройки
 - C) Он ускоряет процесс построения и оценки моделей для начинающих
 - D) Он применим только к текстовым данным

Ответ: C

6. Какой тип задач можно решить с помощью AutoModel?
- A) Только кластеризация
 - B) Регрессия, классификация, кластеризация
 - C) Только построение графиков
 - D) Только экспорт моделей

Ответ: B

7. Что отображается в разделе «Weights» после построения модели?
- A) Нагрузка на память
 - B) Значения корреляции между признаками
 - C) Значимость признаков в построенной модели
 - D) Список всех ошибок модели

Ответ: C

8. Какие основные возможности предоставляет RapidMiner при работе с данными?

Ответ: Загрузка данных, предварительная обработка, визуализация, построение моделей машинного обучения, интерпретация результатов и экспорт моделей.

9. В чем заключается преимущество работы через визуальный процесс (workflow)?

Ответ: Позволяет создавать и настраивать анализ без программирования, упрощает отладку и повторное использование процессов.

10. Назовите минимум три этапа, на которые разделяется типовой анализ в RapidMiner.

Ответ: Загрузка и подготовка данных, построение модели, оценка модели (включая визуализацию и интерпретацию результатов).

Вариант 2

1. Какой элемент интерфейса RapidMiner отвечает за запуск аналитического процесса?
A) Кнопка Run
B) Панель Operators
C) Закладка Statistics
D) TurboPrep

Ответ: А

2. Какую роль выполняет блок “Result” в графическом процессе RapidMiner?
A) Фильтрует данные
B) Визуализирует результат выполнения цепочки
C) Создает модель
D) Сохраняет файл

Ответ: В

3. Что происходит при использовании AutoModel с задачей классификации?
A) Проводится анализ корреляций
B) Запускаются все доступные модели одновременно
C) Происходит автоматический подбор модели, обучение и сравнение метрик
D) Построение модели невозможно без скрипта

Ответ: С

4. Что такое оператор в RapidMiner?
A) Сервис по предсказанию данных
B) Инструмент для загрузки интерфейсов
C) Базовая функциональная единица, обрабатывающая данные
D) Компонент для визуализации

Ответ: С

5. Какую функцию выполняет раздел “Statistics”?
A) Сохраняет отчёты
B) Показывает описательную статистику по выбранному набору данных
C) Проводит машинное обучение
D) Осуществляет вывод текстов

Ответ: В

6. Почему важно использовать нормализацию в RapidMiner перед обучением модели?
- A) Для устранения выбросов
 - B) Для ускорения визуализации
 - C) Для приведения значений к единому масштабу и корректной работы алгоритмов
 - D) Для создания бинарных переменных

Ответ: С

7. Какой компонент отвечает за устранение пропущенных значений в TurboPrep?
- A) Merge
 - B) Cleanse
 - C) Generate
 - D) Transform

Ответ: В

8. Почему RapidMiner подходит для начинающих в анализе данных?

Ответ: Благодаря интуитивному интерфейсу, отсутствию необходимости программировать, встроенным шаблонам и автоматизации работы с моделями.

9. Какие типы данных можно визуализировать в RapidMiner Charts?

Ответ: Числовые и категориальные, включая распределения, корреляции, плотности, сравнение классов.

10. Что обеспечивает функция “Commit Cleanse” в TurboPrep?

Ответ: Подтверждает и сохраняет все действия по очистке и преобразованию данных, применённые в процессе предобработки.

Вариант 3

1. Какой элемент в RapidMiner используется для графического соединения операторов в процессе анализа?

- A) Port Connector
- B) Process Layout

- C) Line Editor
- D) Link Manager

Ответ: А

2. Что произойдет, если в RapidMiner оставить незаполненными входные порты у операторов?
- A) Процесс не будет запущен
 - B) Порты будут проигнорированы
 - C) Операторы самостоятельно заполняют данные
 - D) RapidMiner автоматически назначит роль

Ответ: А

3. Какой вклад вносит AutoModel в процесс обучения моделей в RapidMiner?
- A) Только предоставляет график точности
 - B) Создает предобученную нейросеть
 - C) Позволяет быстро протестировать несколько моделей и сравнить их метрики
 - D) Позволяет изменить интерфейс

Ответ: С

4. Для чего используется вкладка "Design" в RapidMiner?
- A) Отображает только логи
 - B) Предназначена для настройки модели в коде
 - C) Служит для построения и редактирования процессов анализа
 - D) Используется для построения визуализаций

Ответ: С

5. Как RapidMiner обрабатывает бинарные переменные в AutoModel?
- A) Удаляет их
 - B) Преобразует в числовые
 - C) Анализирует как категориальные признаки
 - D) Игнорирует

Ответ: С

6. Что делает панель «Log» в RapidMiner?
- A) Записывает все действия пользователя
 - B) Показывает визуализации

- C) Отображает сообщения, связанные с выполнением процесса
- D) Предоставляет список операторов

Ответ: C

7. Что означает предупреждение «No value for label attribute» в AutoModel?
- A) Нет целевой переменной
 - B) Все признаки категориальны
 - C) Признаки одинаковые
 - D) Недостаточно данных для нормализации

Ответ: A

8. Почему важно проверять входные данные до запуска AutoModel?

Ответ: Наличие пропущенных значений, некорректных типов данных и несоответствий может повлиять на выбор моделей и их результаты.

9. Назови два преимущества визуального подхода к построению анализа в RapidMiner.

Ответ: Высокая наглядность процесса; легкость в отладке и модификации цепочки без написания кода.

10. Какие действия доступны в TurboPrep на этапе «Transform»?

Ответ: Преобразование типов, создание новых признаков, переименование столбцов, вычисление выражений на основе существующих данных.

Вариант 4

1. Что произойдет при попытке запустить AutoModel без предварительного анализа данных?
- A) Сначала будет выполнена автоматическая очистка
 - B) Запуск невозможен — будет ошибка
 - C) RapidMiner предложит рекомендации по улучшению
 - D) Откроется визуализатор

Ответ: B

2. Какая вкладка используется для построения процессов вручную?

- A) Data
- B) Design

- C) Results
- D) View

Ответ: В

3. Что позволяет сделать раздел «Results» после завершения процесса?
- A) Сравнивать операторы
 - B) Просматривать и анализировать выходные данные моделей
 - C) Строить нейросети
 - D) Формировать обучающую выборку

Ответ: В

4. Какой из элементов не является оператором в RapidMiner?
- A) Normalize
 - B) Filter Examples
 - C) Create Process
 - D) Cross Validation

Ответ: С

5. Какую задачу решает компонент «Select Attributes»?
- A) Строит модель
 - B) Выполняет предсказание
 - C) Оставляет или исключает нужные признаки из анализа
 - D) Проводит кросс-валидацию

Ответ: С

6. В чём преимущество выбора модели вручную в AutoModel?
- A) Позволяет задать параметры нормализации
 - B) Исключает необходимость визуализации
 - C) Даёт больше контроля над этапами обучения
 - D) Отключает систему логирования

Ответ: С

7. Какой тип задач чаще всего решается с помощью AutoModel?
- A) Графовая кластеризация
 - B) Классификация и регрессия
 - C) Сравнение форматов файлов
 - D) Генерация отчётов

Ответ: В

8. Какие типы ошибок могут быть обнаружены на этапе AutoModel перед обучением модели?

Ответ: Отсутствие целевой переменной, пропущенные значения, неподходящие типы признаков, несбалансированные классы.

9. Какую роль играет предварительная визуализация данных в TurboPrep?

Ответ: Позволяет выявить выбросы, пропуски, несоответствия в данных до начала моделирования, что повышает качество анализа.

10. Почему RapidMiner популярен среди начинающих аналитиков?

Ответ: Интуитивный интерфейс, визуальное программирование, наличие AutoModel и встроенных инструментов подготовки данных без необходимости писать код.