

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

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

Кафедра «Экономика, менеджмент и организация производства»

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ
Б1.О.24 «ИНСТРУМЕНТАЛЬНЫЕ СИСТЕМЫ РАБОТЫ С
ДАННЫМИ»

Направление подготовки
01.03.05 Статистика

Направленность (профиль) подготовки
«Экономика данных»

Квалификация выпускника – бакалавр

Форма обучения – очная

1. ОБЩИЕ ПОЛОЖЕНИЯ

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

Цель – оценить соответствие знаний, умений и владений, приобретенных обучающимся в процессе изучения дисциплины, целям и требованиям ОПОП в ходе проведения промежуточной аттестации.

Промежуточный контроль по дисциплине осуществляется путем проведения экзамена. Форма проведения зачета – тестирование и выполнение практических заданий. При необходимости, проводится теоретическая беседа с обучаемым для уточнения оценки. Выполнение заданий на лабораторных работах в течение семестра и заданий на самостоятельную работу является обязательным условием для допуска к зачету.

2. ПАСПОРТ ОЦЕНОЧНЫХ МАТЕРИАЛОВ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Контролируемые разделы (темы) дисциплины (результаты по разделам)	Код контролируемой компетенции (или её части)	Наименование оценочного средства
Тема 1. Основы работы в Loginom	ОПК-3.3, ОПК-4.2	Зачет
Тема 2. Подготовка данных	ОПК-3.3, ОПК-4.2	Зачет
Тема 3. Визуализация данных	ОПК-3.3, ОПК-4.2	Зачет

3. ОПИСАНИЕ ПОКАЗАТЕЛЕЙ И КРИТЕРИЕВ ОЦЕНИВАНИЯ КОМПЕТЕНЦИЙ

Сформированность каждой компетенции 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.

Описание критериев и шкалы оценивания промежуточной аттестации

а) описание критериев и шкалы оценивания тестирования:

За каждый вопрос назначается максимально 2 балла в соответствии со следующим правилом:

- 2 балла – ответ на тестовый вопрос полностью правильный;
- 1 балл – отчет на тестовый вопрос частично правильный (выбраны не все правильные варианты, указаны частично верные варианты);
- 0 баллов – ответ на тестовый вопрос полностью не верный.

б) описание критериев и шкалы оценивания практических заданий:

Шкала оценивания	Критерий
5 баллов	Задание выполнено верно
3 балла	Задание выполнено верно, но имеются неточности в ответе
1 балл	Задание выполнено верно, с дополнительными наводящими вопросами преподавателя
0 баллов	Задание не выполнено

На зачет выносятся 10 тестовых вопросов и 4 практических задания. Студент может набрать максимум 20 баллов. Итоговый суммарный балл студента, полученный при прохождении

промежуточной аттестации, переводится в традиционную форму по системе «зачтено» / «не зачтено».

Шкала оценивания	Критерий	
зачтено	10 – 20 баллов	Обязательным условием является выполнение всех предусмотренных в течение семестра заданий (на лабораторных работах и при самостоятельной работе)
Не зачтено	0 – 9 баллов	Студент не выполнил всех предусмотренных в течение семестра текущих заданий (на лабораторных работах и при самостоятельной работе)

4. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ ИЛИ ИНЫЕ МАТЕРИАЛЫ

4.1. Промежуточная аттестация

Коды компетенций	Результаты освоения ОПОП Содержание компетенций
ОПК-3.3	Готовит статистические материалы для докладов, публикаций и других аналитических материалов

а) типовые тестовые вопросы закрытого типа:

- Подход, при котором для решения задач выбираются готовые математические модели с известными параметрами, характерен для:
 - вычислительной математики
 - разведочного анализа
 - аналитики данных
- Укажите наиболее точное определение бизнес-аналитики, которое сегодня подходит к аналитике данных в целом?
 - Инструменты и приложения для поиска, анализа, моделирования и доставки информации, необходимой для принятия решений.
 - Технология в области программного и аппаратного обеспечения, интегрирует, организует, Q управляет и анализирует данные, характеризующиеся четырьмя характеристиками: объемом, разнообразием, изменчивостью и скоростью.
 - Междисциплинарное направление информационных технологий, включающие все аспекты работы с данными с целью извлечения из них полезных знаний.
- Какой столбец набора данных можно с большой вероятностью считать информативным для решения задачи анализа?
 - Столбец, относительно которого выдвинута гипотеза о его влиянии на результат
 - Столбец, содержащий только одно уникальное значение
 - Столбец, содержащий различные значения признака
 - Столбец с полностью уникальными значениями
- Укажите особенности настольных пакетов и библиотек с алгоритмами Data Mining.
 - Отсутствие развитых инструментов интеграции данных
 - Обработка больших объемов данных
 - Большое разнообразие алгоритмов анализа
 - Тесная интеграция с промышленными источниками данных
- Какие особенности малокодовых (low-code) систем в аналитике данных?
 - Поток обработки данных представляет собой графическую диаграмму
 - При реализации логики часто требуется писать код на языке программирования
 - Высокий порог входа для пользователя
 - Ограниченная функциональность
- Отметьте неверное высказывание по тематике Big Data.
 - Объем данных в сотни терабайт или петабайт не позволяет легко хранить и управлять

- ими с помощью традиционных реляционных баз данных
- Б) Большинство всех данных Big Data являются структурированными
- В) Хранение огромного количества данных, описывающих некоторые легко наблюдаемые события, не всегда приводит к выгодному пониманию реальности
- Д) Big Data, в контексте данных, обычно хранятся и организуются в распределенных файловых системах
- 7) Какие способы защиты от копирования применяются в коммерческих редакциях Loginom?
- А) Аппаратный и программный ключ Sentinel
- Б) Текстовый файл с лицензией
- В) Аппаратный и программный ключ Guardant
- 8) Как быстро создать независимую копию узла в рамках одного сценария с такими же настройками и связями?
- А) Командой Клонировать узел
- Б) Командой Копировать узел
- В) Командой Создать производный компонент...
- 9) При каких настройках узла Узел 1 в Модуль 1 им можно будет воспользоваться в Модуль 2? Пакет состоит из Модуль 1 и Модуль 2. Выберите все варианты, комбинация которых даст желаемый результат.
- А) Модификатор доступа у Узел 1 — открытый
- Б) Модификатор доступа у Модуль 1 — закрытый В) Модификатор доступа у Модуль 1 — внутренний
- Г) Модификатор доступа у Узел 1 — внутренний
- Д) Модификатор доступа у Узел 1 — закрытый
- Е) Модификатор доступа у Модуль 1 — открытый
- 10) Сценарий в Loginom представляет собой:
- А) древовидную структуру с ветвями
- Б) направленный граф
- В) нет верного варианта
- 11) При импорте текстового файла у части столбцов конвертация типов данных происходит с потерями. Что может помочь исправить ситуацию?
- А) Изменить вид данных
- Б) Изменить десятичный разделитель
- В) Изменить метку
- Г) Уменьшить значение параметра Кол-во строк для анализа
- Д) Изменить тип данных
- 12) Узел Калькулятор. Отметьте неверные утверждения.
- А) Функция GETVAR() позволяет прочитать значение переменной пакета, системы, приложения
- Б) Функция IFF() возвращает строковое значение
- В) Промежуточное поле не будет передано в выходной набор данных
- Г) Для разделения параметров в функциях калькулятора используется символ /
- 13) Узел Калькулятор. Для использования поля в выражении необходимо указать:
- А) Идентификатор поля
- Б) Имя поля
- В) Метку поля
- Г) Порядковый номер поля в наборе
- 14) При добавлении нового узла в сценарий автосинхронизация его портов:
- А) включена
- Б) не включена
- 15) Какие типы данных совместимы между собой в Loginom?
- А) Строковый и логический
- Б) Целый и вещественный
- В) Дата и Дата/время
- Г) Никакие несовместимы
- 16) Какие метаданные столбцов не принимаются во внимание алгоритмом автоматического связывания?
- А) Тип данных

- Б) Метка столбца
 - В) Вид данных
 - Г) Имя столбца
- 17) Какая группа переменных имеет наивысший приоритет в сценарии?
- А) Пользователя
 - Б) Системы
 - В) Пакета
 - Г) Сессии
- 18) В каком компоненте используется агрегация по полям?
- А) Таблица в переменные
 - Б) Переменные в таблицу
 - В) В обоих
- 19) Какой из компонентов никогда не имеет входных портов?
- А) Любой узел из группы Импорт
 - Б) Выполнение узла
 - В) Узел-ссылка
 - Г) Условие
- 20) Как в сценарии воспользоваться данными узла, который находится в другом модуле текущего пакета?
- А) Добавить ссылку на другой модуль через соответствующую команду в дереве пакетов
 - Б) Использовать компонент Узел-ссылка
 - В) Такой возможности нет

б) типовые практические задания:

Задача 1.

Создайте пакет в Loginom и импортируйте файл Товары.txt. Исключите из набора записи, в которых для товара отсутствует артикул. Сколько строк осталось в файле?

Задача 2.

Создайте пакет в Loginom и импортируйте файл Продажи.txt. С помощью функций компонента Калькулятор рассчитайте, сколько месяцев прошло от даты последней продажи каждого товара до 01.04.2018. Отсортируйте набор данных по количеству месяцев по убыванию. Какой код имеет тридцать седьмой товар?

Задача 3.

Создайте пакет в Loginom и импортируйте файл Численность_населения.txt. Выполните равномерное квантование двумя разными способами. Какое количество групп получилось в каждом случае? Почему это количество отличается/совпадает?

Коды компетенций	Результаты освоения ОПОП Содержание компетенций
ОПК-4.2	Применяет современный информационно-коммуникационные технологии для сбора, хранения, обработки и визуализации больших объемов данных

а) типовые тестовые вопросы закрытого типа:

1. На какой вопрос необходимо ответить при планировании стратегии визуализации:
 - 1) **Какие визуализаторы нам потребуются?**
 - 2) Что мы хотим показать?
 - 3) **Сколько визуализаторов нам потребуется?**
 - 4) **Какую цветовую гамму выбрать?**
2. В каких случаях не рекомендуется использовать круговую диаграмму?
 - 1) Если общая сумма элементов не равна 100%;
 - 2) Если показателей слишком много;
 - 3) **Если элементы нельзя упорядочить логически;**
 - 4) **Если визуализируемые данные являются многомерными.**
3. При сравнительном анализе информации графики должны быть:

- 1) **разнообразными;**
 - 2) **единообразными.**
4. Мозг воспринимает картину целиком и быстрее делает выводы по диаграммам если подписи:
- 1) **расположены непосредственно на графике (диаграмме);**
 - 2) **расположены под графиком (диаграммой);**
 - 3) **расположены над графиком (диаграммой);**
 - 4) **расположены слева от графика (диаграммы);**
 - 5) **расположены справа от графика (диаграммы).**
5. Какой тип визуализации предназначен для анализа и обработки набора данных, с целью обнаружения закономерностей в них:
- 1) **презентационная;**
 - 2) **исследовательская;**
 - 3) **концептуальная;**
 - 4) **метафорическая.**
6. Какой тип визуализации используется для сравнения бизнес-моделей компании:
- 1) **стратегическая;**
 - 2) **исследовательская;**
 - 3) **концептуальная;**
 - 4) **разведочная.**
7. Чтобы не перегружать график большим количеством информации, оптимальное количество разных типов данных или категорий не должно превышать:
- 1) **4-5;**
 - 2) **7-9;**
 - 3) **10-12.**
8. Из каких систем начала развиваться бизнес-аналитика
- 1) **баз данных;**
 - 2) **экспертных систем;**
 - 3) **систем компьютерной математики;**
 - 4) **систем управления производством;**
 - 5) **систем поддержки принятия решений;**
 - 6) **систем бухгалтерского учёта.**
9. В чем заключается задача анализа рыночной корзины:
- 1) **выявление товаров, чаще всего покупаемых совместно;**
 - 2) **выявление наиболее популярных товаров;**
 - 3) **выявление наименее популярных товаров;**
 - 4) **выявление целевых групп клиентов.**
10. Что понимается под транзакцией в задаче анализа рыночной корзины:
- 1) **покупка товаров за одно посещение супермаркета;**
 - 2) **покупка набора товаров по одному чеку;**
 - 3) **покупка товара по клиентской карте торговой сети.**
11. Что позволяет XYZ-анализ:
- 1) **классифицировать ресурсы компании в зависимости от характера их потребления;**
 - 2) **классифицировать ресурсы компании в зависимости от требуемой точности прогнозирования потребности в них;**
 - 3) **сегментировать клиентов компании по уровню лояльности;**
 - 4) **производить анализ товарных запасов.**
12. Какие из перечисленных действий относятся к процедуре профайлинга данных:
- 1) **обнаружение пропусков в данных.**
 - 2) **обнаружение зависимостей и закономерностей в данных;**
 - 3) **обнаружение дубликатов;**
 - 4) **обнаружение выбросов и аномалий;**
 - 5) **агрегирование данных.**
13. Задача распределения объектов на группы на основе сходства их характеристик, при условии, что группы априори заданы, называется:
- 1) **классификация;**
 - 2) **кластеризация;**
 - 3) **регрессия;**

4) ассоциация.

14. Группа векторов в пространстве признаков, такая что, расстояние между членами группы меньше, чем до любого объекта из другой группы, называется:

- 1) кластер;
- 2) класс;
- 3) выборка;
- 4) совокупность.

15. Для решения какой задачи классификации используется самоорганизующаяся карта признаков:

- 1) классификации;
- 2) регрессии;
- 3) ассоциации;
- 4) кластеризации.

16. В каком из элементов структуры дерева решений располагаются решающие правила:

- 1) узел;
- 2) лист;
- 3) корневой узел;
- 5) ствол.

17. Когда обучающий пример считается распознанным в задаче классификации:

- 1) когда предсказанный и наблюдаемый классы совпадают;
- 2) когда предсказанный и наблюдаемый классы не совпадают.

б) типовые тестовые вопросы открытого типа:

1. Интерактивная аналитическая панель с графическим интерфейсом, где на одном экране расположены все ключевые метрики, показатели цели или процессов называется _____ - **(Дашборд)**.

2. _____ тип визуализации представляет данные в виде образов - дерева, пирамиды, иных понятных систем и конструкций **(Метафорический)**.

3. Чем _____ график - тем быстрее отображаемая информация доходит до мозга **(Проще)**.

4. Для демонстрации частей целого наилучшим образом подходит _____ диаграмма **(Круговая)**.

5. _____ диаграмма содержит круговые оси, берущие начало от одной точки отсчета координат **(Радиальная)**.

6. Величина, показывающая, насколько часто условие и следствие ассоциативного правила встречаются совместно в транзакционной базе данных, называется _____ **(Поддержка)**.

7. _____ ассоциативного правила определяется как отношение частоты появления условия и следствия к частоте появления только следствия **(Достоверность)**.

8. _____ правило количественно описывает силу связи между двумя и более событиями **(Ассоциативное)**.

9. В ассоциативном правиле "Из А следует В", событие А называется _____, а событие В _____. **(Условие, Следствие)**.

10. Ассоциативные правила, для которых значения поддержки или достоверности превышают определенный заданный пользователем порог, называются _____ правилами **(Сильными)**.

в) типовые практические задания:

1. Разработать структуру дашборда для представления данных о ходе бизнес-процесса организации (покупки, продажи, услуги и т.д.).

2. Произвести визуализацию набора данных с помощью круговой диаграммы.

3. Разработать визуализатор для анализа активности клиентов компании в течение года.

4. Разработать визуализатор для представления колебаний курсов валют в течение года.

5. Разработать визуализатор взаимодействия структурных подразделений организации с использованием диаграммы связей.

6. Рассчитать значения поддержки и достоверности ассоциативных правил для

транзакционной базы данных.

7. Используя алгоритм поиска ассоциативных правил для анализа рыночной корзины определить, какие два товара наиболее часто продаются вместе.

8. Для заданного набора данных построить модель линейной регрессии и проинтерпретировать его коэффициенты. Сделать вывод о характере влияния независимых переменных на зависимую.

9. Рассчитать коэффициент корреляции Пирсона для входных и выходных переменных аналитической модели. На основе полученных результатов сделать вывод о характере влияния входных переменных на выходные и целесообразности их использования в анализе.

10. Произвести загрузку набора данных по заданию преподавателя в аналитическую платформу.

11. Используя аналитическую платформу по выбору, построить классификатор с использованием дерева решений. Проинтерпретировать сформированные решающие правила, оценить качество классификации с использованием таблицы сопряженности.

12. Используя аналитическую платформу по выбору, построить модель численного предсказания на основе нейросети. Обосновать выбор конфигурации сети, алгоритма обучения и его параметров. Оценить точность модели на основе диаграммы рассеяния.

13. Используя аналитическую платформу по выбору, построить модель кластеризации на основе самоорганизующихся карт признаков. Выполнить содержательную интерпретацию кластеров и оценить качество кластеризации.

14. Построить модель бинарной классификации на основе логистической регрессии. Произвести выбор дискриминационного порога. Оценить точность классификации с помощью матрицы ошибок.

15. Составить список аналитических платформ, предлагаемых в настоящее время на рынке. Сравнить их функциональность, доступность и стоимость и сделать рекомендации по их применению в той или иной предметной области.

Оператор ЭДО ООО "Компания "Тензор"

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ

СОГЛАСОВАНО

ФГБОУ ВО "РГРТУ", РГРТУ, Евдокимова Елена Николаевна, Заведующий
кафедрой ЭМОП

Простая подпись