

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»

МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

**Б1.О.30 «ОПЕРАТИВНЫЙ И ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ
ДАННЫХ»**

Направление подготовки
27.03.05 Инноватика

Направленность (профиль) подготовки
«Технологическое предпринимательство»

Квалификация выпускника – бакалавр

1. ПРимерные планы практических занятий

Тема 1. Общая характеристика методов оперативного и интеллектуального анализа данных

Цель: Познакомить студентов с основными методами оперативного и интеллектуального анализа данных, их применением в различных областях и особенностями реализации. В ходе занятия студенты получают представление о методах анализа данных, что поможет им в дальнейшем изучении предмета и применении полученных знаний в практической деятельности.

Вопросы для обсуждения:

1. Какие основные методы оперативного анализа данных существуют, и в чем их отличия от методов интеллектуального анализа?
2. Каковы примеры применения методов интеллектуального анализа данных в бизнесе и науке?
3. Какие технологии и инструменты используются для реализации методов анализа данных?

Задание:

1. Задание 1: Подготовьте краткий обзор (1-2 страницы) одного из методов оперативного анализа данных (например, OLAP, ETL или Data Warehousing). Опишите его основные характеристики, преимущества и недостатки.
2. Задание 2: Выберите метод интеллектуального анализа данных (например, машинное обучение, кластеризация или ассоциативные правила) и подготовьте презентацию (5-7 слайдов), в которой опишите его принцип работы, примеры применения и инструменты для реализации.

Задания для самостоятельной работы:

Изучите конспект лекции по теме, основную рекомендованную литературу, дополнительные источники:

1. Towards Data Science - Introduction to Data Analysis - статья, описывающая основы анализа данных и различные подходы.
2. KDnuggets - Data Mining and Data Analysis Tutorials - ресурс с множеством статей и учебных материалов по методам анализа данных и машинного обучения.

Тема 2. Аналитические платформы Data Mining

Цель: Познакомить студентов с основными аналитическими платформами для Data Mining, их функциональными возможностями, преимуществами и недостатками, а также применением в реальных проектах. В ходе занятия студенты получают представление о различных аналитических платформах для Data Mining и научатся применять их для решения практических задач анализа данных.

Вопросы для обсуждения:

1. Какие основные аналитические платформы для Data Mining существуют на рынке и в чем их ключевые особенности?
2. Как выбрать подходящую платформу для решения конкретной задачи анализа данных?
3. Какие примеры успешного применения аналитических платформ Data Mining можно привести из практики?

Задание:

1. Задание 1: Проведите исследование одной из популярных аналитических платформ для Data Mining (например, RapidMiner, KNIME или Weka). Подготовьте краткий отчет (1-2 страницы), в котором опишите ее функциональные возможности, интерфейс, преимущества и недостатки.
2. Задание 2: Создайте простую модель анализа данных с использованием выбранной вами аналитической платформы. Подготовьте презентацию (5-7 слайдов), в которой опишите процесс создания модели, использованные данные и полученные результаты.

Задания для самостоятельной работы:

Изучите конспект лекции по теме, основную рекомендованную литературу, дополнительные источники.

1. RapidMiner Documentation - официальная документация RapidMiner, содержащая множество примеров и руководств по использованию платформы.
2. KNIME Hub - ресурс с коллекцией узлов и рабочих процессов KNIME, а также учебными материалами и примерами использования платформы.

Тема 3. Методы Data Mining

Цель: Познакомить студентов с основными методами Data Mining, их применением в различных областях, а также научить выбирать подходящие методы для решения конкретных задач анализа данных. В ходе занятия студенты познакомятся с основными методами Data Mining, научатся применять их на практике и понимать, как выбирать подходящие методы для решения аналитических задач.

Вопросы для обсуждения:

1. Какие основные методы Data Mining существуют и в чем их отличия?
2. Как выбрать подходящий метод Data Mining для решения конкретной задачи?
3. В каких областях и задачах наиболее эффективно применяются методы Data Mining?

Задание:

1. Задание 1: Подготовьте краткий обзор (1-2 страницы) трех различных методов Data Mining (например, кластеризация, регрессия, ассоциативные правила). Включите в обзор описание каждого метода, его применение и примеры реальных задач, где он может быть использован.
2. Задание 2: Выберите одну из изученных тем и разработайте простую модель с использованием одного из методов Data Mining (например, кластеризация с помощью K-средних). Используйте доступные наборы данных (например, Iris или Titanic) и подготовьте презентацию (5-7 слайдов), в которой опишите процесс анализа, использованные данные и полученные результаты.

Задания для самостоятельной работы:

Изучите конспект лекции по теме, основную рекомендованную литературу, дополнительные источники.

1. KDnuggets - ресурс с множеством статей, руководств и учебных материалов по Data Mining и машинному обучению.
2. Towards Data Science on Medium - платформа с публикациями на темы Data Science, включая методы Data Mining, примеры проектов и исследования.

Тема 4. Организация освоения производства нового продукта.

Цель: Познакомить студентов с основными этапами проведения интеллектуального анализа данных, а также научить их правильно интерпретировать и представлять результаты анализа. В ходе занятия студенты овладеют навыками проведения интеллектуального анализа данных, научатся правильно интерпретировать полученные результаты и представлять их в понятной форме для различных аудиторий.

Вопросы для обсуждения:

1. Какие основные этапы включает в себя процесс интеллектуального анализа данных?
2. Каковы ключевые аспекты интерпретации результатов анализа и какие ошибки следует избегать?

3. Как визуализация данных может помочь в интерпретации результатов и донесении их до целевой аудитории?

Задание:

1. Задание 1: Составьте пошаговый план (1-2 страницы) проведения интеллектуального анализа данных на примере конкретной задачи (например, анализ продаж в магазине). Укажите все этапы, начиная с определения проблемы и заканчивая интерпретацией результатов. Обоснуйте выбор каждого этапа.

2. Задание 2: Выберите набор данных (например, набор данных о клиентах или продажах) и проведите анализ, используя один из методов (например, регрессионный анализ или кластеризацию). Подготовьте отчет (3-4 страницы), в котором опишите процесс анализа, результаты и их интерпретацию. Включите графики и таблицы для визуализации данных.

Задания для самостоятельной работы:

Изучите конспект лекции по теме, основную рекомендованную литературу, дополнительные источники.

1. Towards Data Science on Medium - платформа с множеством статей по интеллектуальному анализу данных, включая этапы анализа и интерпретацию результатов.

2. DataCamp - онлайн-курс по аналитике данных, который включает разделы о проведении анализа и интерпретации данных.

II. РЕКОМЕНДАЦИИ ПО ПОДГОТОВКЕ ДОКЛАДОВ

Доклад – это краткое публичное устное изложение результатов индивидуальной учебно-исследовательской деятельности студента, представляет собой сообщение о сути вопроса или исследования применительно к заданной тематике. Доклады направлены на более глубокое самостоятельное изучение обучающимися лекционного материала или рассмотрения вопросов для дополнительного изучения. Данный метод обучения используется в учебном процессе при проведении практических занятий в форме семинаров. Его задачами являются:

- формирование навыков самостоятельной работы, работы с источниками литературы, их систематизация;
- развитие навыков логического мышления;
- углубление теоретических знаний по проблеме исследования;
- развитие навыков публичного выступления, умения уверенно использовать научной терминологией.

Доклад должен представлять аргументированное изложение определенной темы, быть структурирован (по главам, разделам, параграфам) и включать разделы: введение, основная часть, заключение. В ходе доклада должны быть сделаны ссылки на использованные источники. В зависимости от тематики доклада он может иметь мультимедийное сопровождение, в ходе доклада могут быть приведены иллюстрации, таблицы, схемы, макеты, документы и т. д.

Примерные темы докладов:

1. Основные методы и алгоритмы интеллектуального анализа данных.
2. Применение машинного обучения в оперативном анализе данных.
3. Сравнение методов кластеризации: K-means, иерархическая кластеризация и DBSCAN.
4. Роль визуализации данных в интерпретации результатов анализа.
5. Анализ временных рядов: методы и их применение в бизнесе.
6. Этика и безопасность при работе с большими данными.
7. Применение нейронных сетей в интеллектуальном анализе данных.
8. Анализ социальных сетей: методы и инструменты.
9. Использование предиктивной аналитики для повышения эффективности бизнеса.
10. Обработка естественного языка (NLP) в анализе текстовых данных.
11. Системы поддержки принятия решений на основе анализа данных.
12. Инструменты для анализа больших данных: Hadoop, Spark и другие.

13. Кейс-стадии успешного применения аналитики в различных отраслях.
14. Методы оценки качества данных: важность для анализа.
15. Будущее интеллектуального анализа данных: тренды и прогнозы.

III. ВОПРОСЫ К ЗАЧЕТУ ПО ДИСЦИПЛИНЕ

1. Что такое интеллектуальный анализ данных и каковы его основные цели?
2. Какие этапы включает процесс анализа данных?
3. В чем разница между описательной и предсказательной аналитикой?
4. Каковы основные методы машинного обучения?
5. Что такое переобучение и как его избежать?
6. Каковы ключевые аспекты визуализации данных?
7. Что такое кластеризация и для чего она используется?
8. Каковы преимущества и недостатки использования нейронных сетей?
9. Что такое обработка естественного языка (NLP) и какие задачи она решает?
10. Как можно оценить качество данных перед их анализом?
11. В чем разница между структурированными и неструктурированными данными?
12. Какие инструменты используются для работы с большими данными?
13. Каковы основные принципы работы с временными рядами?
14. Какова роль этики в анализе данных?
15. Что такое система поддержки принятия решений и как она функционирует?
16. Каковы основные метрики для оценки качества модели машинного обучения?
17. Что такое регрессионный анализ и в каких случаях он применяется?
18. Как осуществлять мониторинг и обновление моделей аналитики?
19. Какие существуют методы обработки пропущенных значений в данных?
20. Каковы основные тренды в области анализа данных на сегодняшний день?