МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯРОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»(ФГБОУ ВО «ВГТУ», ВГТУ)

Факультет информационных технологий и компьютерной безопасности

Кафедра компьютерных интеллектуальных технологий проектирования

ОТЧЁТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №1

По дисциплине: «Искусственный интеллект в программно-информационных системах»

Тема: «Установка и настройка программного окружения»

Выполнил работу студент группы змИИВТ-231:                            Енин М. М.

подпись, дата

Принял:                                                                                             Короленко В.В.

подпись, дата

Воронеж 2024

Содержание

1. Содержание 2
2. Введение 3
3. Установка программ 6
4. Данные для работы 7
5. Заключение 7
6. Контрольные вопросы 8
7. Литература 30

# **ВВЕДЕНИЕ**

**Jupyter Notebook** – это веб-приложение, которое позволяет создавать и делиться документами, содержащими код на различных языках программирования, математические выражения, графики и текстовые комментарии. Оно широко используется в научных исследованиях, анализе данных и образовании благодаря своей интуитивной и интерактивной природе.

Jupyter Notebook состоит из ячеек двух типов: ячейки кода и ячейки markdown. Ячейки кода содержат исходный код, который можно выполнять для получения результатов, в то время как ячейки markdown позволяют добавлять форматированный текст, изображения и математические формулы. Примеры работы с Jupyter Notebook:

Запуск Jupyter Notebook: Для запуска Jupyter Notebook необходимо открыть терминал (на MacOS и Linux) или командную строку (на Windows) и ввести команду: ```bash jupyter notebook ``` Это запустит сервер Jupyter Notebook, и браузер автоматически откроется на странице Jupyter.

Создание нового ноутбука: На главной странице Jupyter Notebook нажмите кнопку "New" (Новый) в верхнем правом углу и выберите язык программирования, например, "Python 3". Откроется новый ноутбук с именем Untitled (Безымянный).

Выполнение кода: В новоутбуке создайте ячейку кода и введите следующий код: ```python print("Hello, Jupyter Notebook!") ``` Чтобы выполнить код, нажмите кнопку "Run" (Выполнить) в верхнем меню или используйте сочетание клавиш Shift + Enter. Вы увидите результат выполнения кода под ячейкой.

Использование ячеек markdown: Чтобы добавить ячейку markdown, нажмите кнопку "Insert" (Вставить) в верхнем меню и выберите "Insert Cell Above" (Вставить ячейку над) или "Insert Cell Below" (Вставить ячейку под). Затем выберите тип ячейки "Markdown" в выпадающем меню. В ячейке markdown введите следующее: ```markdown # Заголовок 1 ## Заголовок 2 ### Заголовок 3 Текст с \*курсивом\* и \*\*жирным шрифтом\*\*. - Элемент списка 1 - Элемент списка 2 - Элемент списка 3 ``` Выполните ячейку markdown,в кнопку "Run" или используя сочетание клавиш Shift + Enter. Вы увидите, что markdown-код был отформатирован в соответствии с его синтаксисом. Это лишь несколько примеров работы с Jupyter Notebook. Вы можете использовать его для решения различных задач, таких как анализ данных, машинное обучение, визуализация данных и многое другое.

**Statistica** – это мощный программный пакет для анализа данных и статистического моделирования, разработанный компанией StatSoft. Он предоставляет широкий спектр статистических инструментов и методов, а также возможности для визуализации данных и их предварительной обработки. Statistica широко используется в научных исследованиях, промышленности, медицине и образовании. Основные возможности Statistica:

1. Предварительная обработка данных: очистка данных, изменение формата, объединение и разделение таблиц, выборка данных и т.д.

2. Статистический анализ: описательная статистика, корреляционный анализ, регрессионный анализ, дисперсионный анализ, кластерный анализ, факторный анализ и многие другие методы.

3. Визуализация данных: гистограммы, точечные диаграммы, круговые диаграммы, ящичковые диаграммы, матрицы диаграмм рассеяния и другие типы графиков.

4. Моделирование и прогнозирование: нейронные сети, деревья решений, случайные леса, временные ряды и другие методы.

5. Интеграция с другими программными продуктами: возможность импорта и экспорта данных в различных форматах, включая Excel, CSV, SPSS и другие.

Пример работы с Statistica: рассмотрим простой пример анализа данных с использованием Statistica. Предположим, у нас есть набор данных о сотрудниках компании, включающий информацию о возрасте, стаже работы, уровне образования и зарплате.

1. Импортируем данные в Statistica, выбрав пункт меню "Файл" > "Открыть" и выбрав файл с данными в формате CSV или Excel.

2. Проведем описательный статистический анализ, чтобы получить основные характеристики набора данных. Для этого выберем пункт меню "Анализ" > "Основные статистики/Таблицы" > "Описательные статистики и таблицы". В открывшемся диалоговом окне выберем интересующие нас переменные и нажмем "OK".

3. Визуализируем данные с помощью гистограммы зарплат. Для этого выберем пункт меню "Графики" > "Гистограммы" и выберем переменную "Зарплата" в открывшемся диалоговом окне.

4. Проведем корреляционный анализ для изучения взаимосвязи между переменными. Выберем пункт меню "Анализ" > "Основные статистики/Таблицы" > "Корреляции" и выберем переменные для анализа.

5. Сохраним результаты анализа, выбрав пункт меню "Файл" > "Сохранить как" и указав имя файла и формат (например, Statistica, Excel или CSV). Этот пример показывает, как можно использовать Statistica для анализа и визуализации данных. Программа предоставляет широкие возможности для проведения более сложных статистических исследований и моделирования.

# **Установка программ**

Jupyter Notebook - это веб-приложение, которое позволяет создавать и делиться документами, содержащими код, текст и графики. Оно широко используется в научных исследованиях и анализе данных. Вот как установить Jupyter Notebook на Windows:

1. Убедитесь, что у Вас установлен Python. Если нет, загрузите и установите последнюю версию с сайта python.org.

2. Откройте командную строку (cmd) или PowerShell.

3. Установите Jupyter Notebook с помощью pip, который является менеджером пакетов для Python. Введите следующую команду: ``` pip install jupyter ```

4. После установки запустите Jupyter Notebook, введя следующую команду: ``` jupyter notebook ```

5. Откроется браузер с интерфейсом Jupyter Notebook. Теперь вы можете создавать новые ноутбуки, открывать существующие и работать с ними.

Статистика (Statistica) - это программный пакет для статистического анализа данных и машинного обучения. Он не является бесплатным и открытым программным обеспечением, как Jupyter Notebook, и его установка немного отличается. Вот общий процесс установки Statistica на Windows:

1. Для данной работы использовалась базовая версия для студентов, которая распространяется бесплатно.

2. \*\*Скачивание установщика\*\*: После покупки лицензии Вы получите доступ к загрузке установщика программы. Обычно это делается через личный кабинет на сайте производителя или по электронной почте, которую Вы указали при покупке.

3. \*\*Запуск установщика\*\*: Найдите скачанный файл установщика (обычно имеет расширение `.exe`) и запустите его. Убедитесь, что у Вас есть права администратора на компьютере, так как установка может потребовать доступа к системным файлам.

4. \*\*Установка программы\*\*: Следуйте инструкциям мастера установки. Выберите язык, папку установки и другие параметры по своему усмотрению. Возможно, Вам потребуется ввести лицензионный ключ, который Вы получили при покупке.

5. \*\*Завершение установки\*\*: после того как установка завершена, вы можете запустить Statistica и начать работу с ней.

**Данные для работы**

Для данной работы были использованы данные скаченные с сайта <https://www.kaggle.com/> Прогноз цен на акции Nvidia с течением времени.

# **Заключение**

Jupyter Notebook и Statistica являются мощными инструментами для анализа данных и проведения исследований в различных областях. Jupyter Notebook предоставляет гибкую и интерактивную среду для работы с кодом, визуализации данных и документирования результатов. Он поддерживает множество языков программирования, что делает его универсальным инструментом для специалистов в области науки о данных, машинного обучения и анализа данных. С другой стороны, Statistica является специализированным программным обеспечением для статистического анализа и обработки данных. Он предлагает широкий спектр статистических методов и инструментов, а также возможности визуализации данных и создания отчетов. Statistica часто используется в академических исследованиях, а также в промышленности и бизнесе для принятия обоснованных решений на основе данных. Оба инструмента имеют свои сильные стороны и могут быть использованы в различных контекстах в зависимости от потребностей пользователя. Jupyter Notebook лучше подходит для гибкой и интерактивной работы с данными, а также для реализации собственных алгоритмов и моделей. Statistica, в свою очередь, предоставляет готовые решения для статистического анализа и может быть полезен для пользователей, не имеющих глубоких знаний в области программирования. В целом, выбор между Jupyter Notebook и Statistica зависит от конкретных задач и предпочтений пользователя.

# **Контрольные вопросы**

Термины:

1. Искусственный интеллект (ИИ): Это раздел компьютерных наук, который фокусируется на создании машин или программ, способных выполнять задачи, которые обычно требуют человеческого интеллекта, такие как обучение, обобщение, распознавание образов, решение проблем и т.д.

2. Технологии искусственного интеллекта: Это набор инструментов и методов, используемых для создания искусственного интеллекта, включая машинное обучение, нейронные сети, глубокое обучение, распознавание речи и т.д.

3. Перспективные методы искусственного интеллекта: Это новые или развивающиеся подходы в области искусственного интеллекта, которые могут предоставить новые возможности или улучшить существующие методы.

4. Набор данных: Это коллекция данных, которые используются для обучения, тестирования и оценки моделей искусственного интеллекта и машинного обучения.

5. Разметка данных: Это процесс добавления меток или аннотаций к данным для их лучшей интерпретации и использования в обучении мо.

6. Общедоступная платформа: Это платформа, доступная для широкой публики, которая может использоваться для разработки, тестирования и развертывания искусственного интеллекта и машинного обучения.

7. Открытая библиотека искусственного интеллекта: Это набор инструментов и библиотек, которые предоставляются бесплатно и с открытым исходным кодом для разработки искусственного интеллекта и машинного обучения.

8. Машинное обучение: Это раздел искусственного интеллекта, который фокусируется на разработке алгоритмов, которые позволяют компьютерам учиться на основе данных.

9. Автоматизированная система: Это система, которая может выполнять задачи без прямого вмешательства человека, используя искусственный интеллект и машинное обучение.

10. Агент: Это объект или сущность, которая может действовать самостоятельно в окружающей среде, используя искусственный интеллект и машинное обучение.

11. Библиотека знаний: Это коллекция информации и знаний, которые используются для обучения искусственного интеллекта и машинного обучения.

12. Большие данные: Это термин, используемый для описания очень больших наборов данных, которые могут быть использованы для обучения искусственного интеллекта и машинного обучения.

13. Знания (в искусственном интеллекте): Это информация и знания, которые используются для обучения искусственного интеллекта и машинного обучения.

14. Интернет вещей (IoT): Это сеть физическихустройств, которые связаны через Интернет и могут обмениваться данными и взаимодействовать друг с другом, используя искусственный интеллект и машинное обучение.

15. Интероперабельность: Это способность различных систем и устройств взаимодействовать друг с другом и обмениваться данными, используя искусственный интеллект и машинно.

16. Классификация: Это процесс группировки объектов или сущностей на основе их сходства или различий, используя искусственный интеллект и машинное обучение.

17. Компьютерное зрение: Это раздел искусственного интеллекта, который фокусируется на разработке алгоритмов, которые позволяют компьютерам видеть и интерпретировать изображения и видео.

18. Многоагентная система: Это система, состоящая из нескольких агентов, которые могут взаимодействовать друг с другом и с окружающей средой, используя искусственный интеллект и машинное обучение.

19. Модель знаний: Это представление знаний и информации в виде структуры или схемы, которая может быть использована для искусственного интеллекта и машинного обучения.

20. Распределенная система: Это система, состоящая из нескольких компонентов, которые распределены по разным узлам или устройствам и могут взаимодействовать друг с другом, используя искусственный интеллект и машинное обучение.

21. Обучение с учителем: Это метод машинного обучения, в котором модель обучается на основе данных с метками или аннотациями, предоставленными человеком.

22. Обучение без учителя: Это метод машинного обучения, в котором модель обучается на основе данных без меток или аннотаций, предоставленных человеком.

23. Обучение (модели): Это процесс, в котором модель искусственного интеллекта или машинного обучения обучается на основе данных и извлекает знания из них.

24. Система искусственного интеллекта: Это система, которая использует искусственный интеллект и машинное обучение для выполнения задач, которые обычно требуют человеческого интеллекта.

25. Адаптивность в контексте искусственного интеллекта (ИИ) относится к способности системы изменять свое поведение или алгоритмы в ответ на изменения в окружающей среде или данных. Это означает, что система может приспосабливаться к новым условиям, неизвестным данным или изменениям в требованиях задачи, не требуя ручного перепрограммирования.

26. Способность к самообучению: Способность к самообучению, также известная как самообучающиеся системы, относится к способности искусственных интеллектуальных систем улучшать свои модели и алгоритмы без прямого вмешательства человека. Это означает, что система может извлекать знания из новых данных, корректировать свои параметры и улучшать свои прогнозы или решения с течением времени. Самообучение является ключевым аспектом многих методов машинного обучения и глубокого обучения, где алгоритмы могут обучаться на больших объемах данных и автоматически настраивать свои внутренние параметры для достижения лучших результатов.

27. Oracle Business Intelligence Cloud Service: Достоинства: - Гибкость и масштабируемость благодаря облачной архитектуре. - Интеграция с другими продуктами Oracle. - Широкий набор инструментов для анализа данных и создания отчетов. - Поддержка мобильных устройств. Недостатки: - Высокая стоимость для небольших и средних предприятий. - Требуется определенный уровень знаний и опыта работы с продуктами Oracle. - Не всегда удобный и интуитивно понятный интерфейс.

28. Sisense: Достоинства: - Простота использования и быстрая интеграция с различными источниками данных. - Возможность обработки больших объемов данных. - Интуитивно понятный интерфейс и широкие возможности для создания визуализаций. Недостатки: - Высокая стоимость для небольших и средних предприятий. - Некоторые функции могут быть ограничены в бесплатной версии. - Требуется время для освоения всех возможностей платформы. Аналитическая платформа B3: Достоинства: - Гибкость и масштабируемость благодаря облачной архитектуре. - Интеграция с различными источниками данных. - Возможности для создания сложных моделей и прогнозов. Недостатки: - Отсутствие подробной документации и технической поддержки. - Не всегда удобный и интуитивно понятный интерфейс. - Высокая стоимость для небольших и средних предприятий.

29. Dataiku Data Science Studio: Достоинства: - Интеграция с различными источниками данных и инструментами машинного обучения. - Возможности для совместной работы и управления проектами. - Поддержка различных языков программирования и фреймворков. Недостатки: - Высокая стоимость для небольших и средних предприятий. - Требуется определенный уровень знаний и опыта работы с аналитическими инструментами. - Не всегда удобный и интуитивно понятный интерфейс.

30. KNIME Analytics Platform: Достоинства: - Открытый исходный код и бесплатная лицензия. - Широкий набор инструментов для анализа данных и машинного обучения. - Возможности для совместной работы и управления проектами. Недостатки: - Требуется время для освоения всех возможностей платформы. - Не всегда удобный и интуитивно понятный интерфейс. - Отсутствие технической поддержки для бесплатной версии.

31. Gephi: Достоинства: - Бесплатная лицензия и открытый исходный код. - Инструменты для визуализации и анализа графовых структур. - Возможности для совместной работы и управления проектами. Недостатки: - Узкая специализация на анализе графовых структур. - Требуется время для освоения всех возможностей платформы. - Отсутствие технической поддержки для бесплатной версии.

32. Python: Достоинства: - Широкий набор библиотек и фреймворков для анализа данных и машинного обучения. - Простота использования и легкость изучения. - Большое сообщество пользователей и разработчиков. Недостатки: - Немного меньшая скорость выполнения по сравнению с некоторыми другими языками. - Не всегда удобный и интуитивно понятный синтаксис. - Некоторые функции могут быть ограничены в бесплатных библиотеках. R: Достоинства: - Широкий набор библиотек и фреймворков для анализа данных и статистики. - Возможности для создания сложных моделей и прогнозов. - Большое сообщество пользователей и разработчиков. Недостатки: - Требуется время для освоения всех возможностей языка. - Не всегда удобный и интуитивно понятный синтаксис. - Немного меньшая скорость выполнения по сравнению с некоторыми другими языками.

33. Julia: Достоинства: - Высокая скорость выполнения и эффективность. - Интеграция с различными источниками данных и инструментами машинного обучения. - Возможности для создания сложных моделей и прогнозов. Недостатки: - Меньшее количество библиотек и фреймворков по сравнению с Python и R. - Требуется время для освоения всех возможностей языка. - Меньшее сообщество пользователей и разработчиков.

34. Scala: Достоинства: - Высокая скорость выполнения и эффективность. - Интеграция с различными источниками данных и инструментами машинного обучения. - Возможности для создания сложных моделей и прогнозов. Недостатки: - Требуется время для освоения всех возможностей языка. - Не всегда удобный и интуитивно понятный синтаксис. - Меньшее сообщество пользователей и разработчиков по сравнению с Python и R.

35. SQL: Достоинства: - Широкий набор инструментов для работы с реляционными базами данных. - Простота использования и легкость изучения. - Большое сообщество пользователей и разработчиков. Недостатки: - Не подходит для работы с неструктурированными данными. - Не всегда удобный и интуитивно понятный синтаксис.

36. PolyAnalyst: Достоинства: - Высокая степень автоматизации аналитических процессов. - Поддержка различных типов данных и форматов. - Интеграция с другими системами и инструментами. - Возможность работы с большими объемами данных. Недостатки: - Высокая стоимость лицензии. - Относительно сложное освоение для начинающих пользователей. - Недостаточно развитая сообщество пользователей и ресурсы для обучения.

37. Polymatica: Достоинства: - Интуитивно понятный интерфейс. - Возможность быстрой подготовки и анализа данных. - Поддержка больших объемов данных. - Интеграция с другими системами и инструментами. Недостатки: - Высокая стоимость лицензии. - Недостаточно развитая сообщество пользователей и ресурсы для обучения. - Отсутствие некоторых продвинутых функций, доступных в конкурирующих продуктах.

38. Stata: Достоинства: - Широкий набор статистических методов и моделей. - Высокая скорость обработки данных. - Возможность написания собственных скриптов и макросов. - Хорошо документированная и регулярно обновляемая. Недостатки: - Высокая стоимость лицензии. - Относительно сложное освоение для начинающих пользователей. - Недостаточно развитая визуализация данных по сравнению с конкурирующими продуктами.

39. IBM SPSS Statistics: Достоинства: - Широкий набор статистических методов и моделей. - Интуитивно понятный интерфейс. - Возможность интеграции с другими продуктами IBM. - Хорошо документированная и регулярно обновляемая. Недостатки: - Высокая стоимость лицензии. - Относительно низкая скорость обработки данных по сравнению с конкурирующими продукта - Недостаточно развитая визуализация данных.

40. Loginom: Достоинства: - Гибкая и модульная архитектура. - Возможность работы с большими объемами данных. - Интеграция с другими системами и инструментами. - Удобный и интуитивно понятный интерфейс. Недостатки: - Высокая стоимость лицензии. - Недостаточно развитая сообщество пользователей и ресурсы для обучения. - Отсутствие некоторых продвинутых функций, доступных в конкурирующих продуктах.

41. SAP Analytics Cloud: Достоинства: - Интеграция с другими продуктами SAP. - Возможность работы с большими объемами данных. - Интуитивно понятный интерфейс. - Возможность проведения сложных аналитических исследований. Недостатки: - Высокая стоимость лицензии. - Относительно сложное освоение для начинающих пользователей. - Недостаточно развитая сообщество пользователей и ресурсы для обучения.

42. Qlik: Достоинства: - Интуитивно понятный интерфейс. - Возможность быстрой подготовки и анализа данных. - Поддержка больших объемов данных. - Интеграция с другими системами и инструментами. Недостатки: - Высокая стоимость лицензии. - Недостаточно развитая сообщество пользователей и ресурсы для обучения. - Отсутствие некоторых продвинутых функций, доступных в конкурирующих продуктах. 43. Tableau: Достоинства: - Высококачественная визуализация данных. - Интуитивно понятный интерфейс. - Возможность быстрой подготовки и анализа данных. - Поддержка больших объемов данных. Недостатки: - Высокая стоимость лицензии. - Относительно низкая скорость обработки данных по сравнению с конкурирующими продуктами. - Недостаточно развитая сообщество пользователей и ресурсы для обучения.

44. Proceset: Достоинства: - Гибкая и модульная архитектура. - Возможность работы с большими объемами данных. - Интеграция с другими системами и инструментами. - Удобный и интуитивно понятный интерфейс. Недостатки: - Высокая стоимость лицензии. - Недостаточно развитая сообщество пользователей и ресурсы для обучения. - Отсутствие некоторых продвинутых функций, доступных в конкурирующих продуктах.

45. TIBCO Data Science: Достоинства: интегрированная платформа для аналитики и науки о данных, поддержка различных языков программирования, возможность совместной работы, широкий набор инструментов для анализа и визуализации данных. Недостатки: сложность в освоении для новичков, высокая стоимость, зависимость от других продуктов TIBCO.

46. Jupyter Notebook: Достоинства: интерактивная среда для работы с кодом, возможность визуализации данных, поддержка множества языков программирования, легкость в использовании, открытый исходный код. Недостатки: ограниченные возможности управления версиями, не всегда удобная работа с большими наборами данных, отсутствие встроенной поддержки совместной работы.

47. IDLE Python: Достоинства: простой и легкий в использовании редактор кода, встроен в Python, поддержка автодополнения кода, открытый исходный код. Недостатки: ограниченные возможности по сравнению с другими IDE, отсутствие встроенных инструментов для анализа и визуализации данных, не поддерживает многие современные функции и технологии.

48.Google Colab: Достоинства: бесплатная облачная платформа, возможность использования GPU и TPU, интеграция с Google Drive, поддержка Jupyter Notebook, возможность совместной работы. Недостатки: ограниченный объем бесплатного хранилища, зависимость от скорости интернет-соединения, возможные проблемы с конфиденциальностью данных.

50. Anaconda: Достоинства: удобная установка и управление пакетами, поддержка множества научных библиотек Python, возможность создания изолированных окружений, открытый исходный код. Недостатки: большой размер дистрибутива, возможные конфликты версий пакетов, некоторые функции доступны только в коммерческой версии.

51. MatLa: Достоинства: отсутствует информация о продукте с таким названием. Недостатки: невозможно оценить достоинства и недостатки, так как информация о продукте отсутствует.

52. Statistica: Достоинства: широкий набор статистических методов и инструментов, удобный интерфейс, возможность визуализации данных, поддержка совместной работы. Недостатки: высокая стоимость, сложность в освоении для новичков, зависимость от производителя для обновлений и поддержки.

53. Power BI: Достоинства: мощные инструменты для анализа и визуализации данных, интеграция с другими продуктами Microsoft, возможность создания сложных отчетов, относительно низкая стоимость. Недостатки: сложность в освоении для новичков, возможные проблемы с производительностью при работе с большими наборами данных, ограниченная поддержка сторонних источников данных.

54. Julia: Возможности языка: Julia - высокоуровневый динамический язык программирования, предназначенный для научных вычислений и анализа данных. Он обладает высокой производительностью, благодаря использованию JIT-компиляции и поддержке параллелизма. Julia также имеет мощную математическую библиотеку и поддержку внешних пакетов. Наиболее часто решаемые задачи: анализ данных, машинное обучение, оптимизация, обработка сигналов, моделирование и имитация.

55. Scala: Возможности языка: Scala - функциональный и объектно-ориентированный язык программирования, работающий на виртуальной машине Java. Он сочетает в себе возможности обоих парадигм и позволяет разрабатывать масштабируемые и высокопроизводительные приложения.Scala хорошо подходит для работы с большими данными и облачными вычислениями. Наиболее часто решаемые задачи: анализ больших данных, машинное обучение, разработка веб-приложений, системы рекомендаций.

56. SQL: Возможности языка: SQL (Structured Query Language) - язык запросов к реляционным базам данных. Он позволяет выполнять операции выборки, вставки, обновления и удаления данных, а также управлять структурой базы данных. Наиболее часто решаемые задачи: работа с реляционными базами данных, выборка и анализ данных, создание отчетов, управление данными.

57. PyCharm: Достоинства: интегрированная среда разработки (IDE) для Python с мощным функционалом, включая автодополнение кода, отладку, тестирование, поддержку виртуальных окружений и многие другие функции. Недостатки:может быть медленным на старых компьютерах, требует значительных ресурсов, есть ограничения в бесплатной версии.

54. Visual Studio Code: Достоинства: легковесная и быстрая IDE с открытым исходным кодом, поддерживает множество языков программирования, имеет обширный набор расширений, включая отладку, тестирование и многие другие функции. Недостатки: может быть сложным для новичков, требует настройки и установки дополнительных расширений для работы с некоторыми языками.

55. Microsoft Visual Studio: Достоинства: мощная и функциональная IDE, поддерживает множество языков программирования, включая C#, C++, Python и другие. Имеет широкий набор инструментов для разработки, отладки и тестирования приложений. Недостатки: требует значительных ресурсов, может быть медленным на старых компьютерах, имеет сложный интерфейс и высокую стоимость для коммерческого использования.

56. R-Studio: Достоинства: интегрированная среда разработки для языка R с функционалом для анализа данных, включая отладку, визуализацию данных, поддержку пакетов и многие другие функции. Недостатки: может быть сложным для новичков, требует настройки и установки дополнительных пакетов для работы с некоторыми задачами.

57. Microsoft Excel и NodeXL: Достоинства: Microsoft Excel - популярный инструмент для работы с данными, включая ввод, редактирование, анализ и визуализацию. NodeXL - дополнение для Excel, позволяющее анализировать и визуализировать социальные сети. Недостатки: ограниченная возможность работы с большими данными, сложность в использовании NodeXL для новичков, необходимость дополнительного обучения для эффективного использования функционала.

58. MatLab: Достоинства: мощный инструмент для математических вычислений, моделирования и анализа данных. Имеет обширную библиотеку функций и поддержку внешних пакетов. Недостатки: высокая стоимость, сложность в освоении для новичков, ограниченная возможность работы с большими данными.

59. Statistica: Достоинства: программный продукт для статистического анализа данных с широким набором функций, включая визуализацию данных, моделирование, прогнозирование и многие другие. Недостатки: высокая стоимость, сложность в освоении для новичков,

60. Oracle Business Intelligence Cloud Service (BICS) - это облачная платформа для аналитики и бизнес-интеллекта, предлагающая широкий спектр возможностей для анализа данных и принятия решений. Она позволяет создавать интерактивные отчеты, панели мониторинга и визуализации данных. Часто используется для финансового анализа, управления ресурсами, маркетинга и CRM.

61. Sisense - это программное обеспечение для аналитики и бизнес-интеллекта, которое позволяет создавать сложные визуализации и отчеты на основе больших объемов данных. Sisense поддерживает интеграцию с различными источниками данных и предлагает гибкие инструменты для анализа данных, что делает его популярным среди компаний, которые хотят улучшить свою аналитическую деятельность. 62. Аналитическая платформа B3 - это специализированное программное обеспечение для финансового анализа и мониторинга рынков. Она предоставляет широкий спектр инструментов для анализа данных, включая графики, индикаторы и прогнозы. Платформа часто используется трейдерами и аналитиками для принятия инвестиционных решений.

63. Dataiku Data Science Studio (DSS) - это платформа для коллаборативной работы над проектами машинного обучения и анализа данных. Она предлагает полный набор инструментов для подготовки данных, разработки моделей машинного обучения и их развертывания. DSS поддерживает множество языков программирования и алгоритмов машинного обучения.

64. KNIME Analytics Platform - это открытая платформа для анализа данных и машинного обучения, которая используется для интеграции, анализа и визуализации данных. KNIME предоставляет графический интерфейс для создания потоков данных, включая различные модули для обработки данных и алгоритмы машинного обучения.

65. Gephi - это программное обеспечение для визуализации и анализа графовых структур. Оно часто используется для исследования социальных сетей, семантических сетей и других типов связанных данных. Gephi предоставляет мощные инструменты для визуализации и анализа графов, включая алгоритмы для выделения сообществ и измерения центральности узлов.

66. Python - это универсальный язык программирования, который широко используется в области анализа данных и машинного обучения благодаря своей простоте, мощным библиотекам (таким как NumPy, Pandas, Scikit-learn, TensorFlow, PyTorch) и активной сообществу. Python поддерживает различные алгоритмы машинного обучения, обработку и анализ данных, а также визуализацию результатов.

67. R - это язык программирования, специализированный для статистических вычислений и графики. Он широко используется в академических кругах и промышленности для анализа данных и разработки статистических моделей. R обладает обширным набором пакетов, которые предоставляют функции для различных

68. Под искусственным интеллектом (ИИ) понимается способность компьютерных систем выполнять задачи, которые обычно требуют человеческого интеллекта, такие как понимание естественного языка, распознавание образов, обучение, планирование и решение проблем.

69. Понятие искусственного интеллекта возникло в 1956 году на конференции в Дартмутском колледже, когда группа исследователей предложила создать науку о "мыслящих" машинах.

70. Рост количества данных (больших данных) оказал значительное влияние на развитие ИИ, так как позволил создавать и обучать более сложные модели машинного обучения и глубокого обучения, которые способны извлекать знания и делать прогнозы на основе больших объемов информации.

71. Развитие техники, такого как вычислительные мощности, памяти, сетевые технологии и интерфейсы, оказало значительное влияние на ИИ, позволяя создавать более сложные и эффективные системы, а такжеширять возможности их применения.

72. Примеры применения ИИ в промышленности включают автоматизацию производственных процессов, контроль качества, прогнозирование поломок оборудования и оптимизацию логистики.

73. Примеры применения ИИ в интернет-сервисах включают персональных рекомендации, фильтрацию спама, распознавание речи и оптимизацию контента для пользователей.

74. В сфере услуг, торговли и логистики ИИ используется для автоматизации процессов, оптимизации распределения ресурсов, прогнозирования спроса и предложения, а также для улучшения обслуживания клиентов.

75. Искусственный интеллект может быть частью информационной системы, обеспечивая более глубокий анализ данных, прогнозирование, оптимизацию процессов и поддержку принятия решений.

76. Классификация систем искусственного интеллекта может быть основана на различных критериях, таких как тип используемого алгоритма (распознавание образов, машинное обучение, нейронные сети), область применения (промышленность, здравоохранение, финансы) или уровень автономности (уровни от полностью автономных систем до систем с человеком в контуре).

77. Математический аппарат искусственного интеллекта включает в себя теорию вероятностей, статистику, оптимизацию, теорию игр, теорию графов и другие области математики, которые используются для построения и анализа моделей, лежащих в основе алгоритмов ИИ.

78. Машинное обучение: основные термины Машинное обучение (Machine Learning, ML) - это раздел искусственного интеллекта, который фокусируется на разработке алгоритмов, способных обучаться на основе данных и улучшать свою производительность с течением времени. Основные термины в машинном обучении включают:- Алгоритмы обучения: набор правил или процедур, используемых для обучения модели на основе данных. - Обучающая выборка: набор данных, используемый для обучения модели. - Тестовая выборка: набор данных, используемый для оценки производительности обученной модели. - Метрики оценки: показатели, используемые для измерения качества работы модели (например, точность, полнота, F1-мера). - Оверфиттинг (переобучение): ситуация, когда модель слишком сильно подстраивается под обучающую выборку и плохо работает на новых данных. - Андерфиттинг (недообучение): ситуация, когда модель недостаточно обучена и плохо работает на обеих выборках.

79. Машинное обучение: решаемые задачи и инструментарий Машинное обучение используется для решения широкого круга задач, включая: - Классификация: отнесение объектов к определенным категориям. - Регрессия: предсказание непрерывных значений. - Кластеризация: группировка объектов на основе их сходства. - Рекомендации: предложение товаров или услуг на основе предпочтений пользователя. - Обработка естественного языка (NLP): анализ и синтез текста. - Распознавание изображений: идентификация объектов на изображениях. Инструментарий машинного обучения включает библиотеки и фреймворки, такие как Scikit-learn, TensorFlow, PyTorch, Keras, и другие, которые предоставляют API для реализации алгоритмов ML.

80. Наиболее популярные библиотеки для машинного обучения. Краткая характеристика. - Scikit-learn: простая и эффективная библиотека для машинного обу Python, которая поддерживает множество алгоритмов классификации, регрессии и кластеризации. - TensorFlow: открытая библиотека разработанная компанией Google, ориентированная на глубокое обучение и нейронные сети. - PyTorch: динамическая библиотека для глубокого обучения, разработанная компанией Facebook, которая предоставляет высоровневые инструменты для создания нейронных сетей. - Keras: высокоуровневая библиотека для глубокого обучения, которая может работать поверх TensorFlow, Theano или CNTK и предоставляет простой в использовании API.

81. Прикладное программное обеспечение, используемое для решения задач ИИ Прикладное программное обеспечение для решения задач ИИ включает: - Scikit-learn, TensorFlow, PyTorch и другие библиотеки для реализации алгоритмов машинного обучения. - Jupyter Notebook и другие интерактивные среды разработки для анализа данных и прототипирования моделей. - Веб-фреймворки и сервисы для развертывания и масштабирования моделей (например, Flask, Django, AWS SageMaker, Google AI Platform).

82. Специалисты в области Data Science используют широкий спектр инструментальных средств для организации своего рабочего места. К ним относятся: - \*\*Компьютерные языки программирования\*\*: Python, R, Julia, Scala, Java и другие. - \*\*Среды разработки и интегрированные среды разработки (IDE)\*\*: PyCharm, RStudio, Jupyter Notebook, Visual Studio Code, Spyder и другие. - \*\*Системы управления версиями\*\*: Git, SVN, Mercurial и другие. - \*\*Базы данных и системы управления базами данных (DBMS)\*\*: PostgreSQL, MySQL, MongoDB, Apache Cassandra и другие. - \*\*Инструменты для работы с большими данными\*\*: Apache Hadoop, Apache Spark, Dask, TensorFlow, PyTorch и другие. - \*\*Инструменты для визуализации данных\*\*: Tableau, PowerBI, Matplotlib, Seaborn, ggplot2 и другие. - \*\*Инструменты для анализа данных и машинного обучения\*\*: scikit-learn, XGBoost, LightGBM, Keras, H2O и другие. - \*\*Облачные сервисы\*\*: Amazon AWS, Google Cloud, Microsoft Azure, IBM Watson и другие.

83. \*\*Jupyter Notebook\*\* - это веб-приложение, которое позволяет создавать и делиться документами, содержащими код, текст, математику, графики и мультимедиа. Возможности Jupyter Notebook включают: - \*\*Интерактивное программирование\*\*: позволяет пользователям писать и запускать код в интерактивном режиме. - \*\*Многоязычность\*\*: поддержи множество языков программирования, включая Python, R, Julia и другие. - \*\*Визуализация данных\*\*: предоставляет возможности для создания графиков и диаграмм прямо в ноутбуке. - \*\*Документирование\*\*: позволяет добавлять markdown-текст, формулы LaTeX и комментарии к коду. - \*\*Работа с данными\*\*: поддерживает работу с большими наборами данных и предобработку данных. - \*\*Разработка моделей машинного обучения\*\*: может использоваться для построения, обучения и оценки моделей машинного обучения.

84. IDLE Python: IDLE (Integrated Development and Learning Environment) - это простой текстовый редактор и среда разработки, встроенная в стандартную установку Python. Она подходит для начинающих разработчиков и исследователей данных, которые хотят начать работать с Python. Возможности IDLE включают: - Интерактивный интерпретатор Python- Цветное форматирование синтаксиса для упрощения чтения кода. - Подсветка ошибок во время ввода кода. - Простой многооконный интерфейс. - Поддержка поиска и замены текста. - Отладка с пошаговым выполнением и просмотром переменных. IDLE часто используется для обучения основам Python, простых скриптов и прототипирования, но для более сложных проектов анализа данных и машинного обучения, как правило, выбирают более мощные инструменты.

85. PyCharm: PyCharm - это интегрированная среда разработки (IDE) для Python, разработанная JetBrains. Она предлагает широкий спектр функций, которые делают ее популярным выбором среди профессиональных разработчиков и специалистов по данным. Возможности PyCharm включают: - Автозавершение кода и проверка синтаксиса. - Встроенный отладчик для пошагового выполнения кода и просмотра переменных. - Поддержка виртуальных окружений и менеджер пакетов. - Интеграция с системами контроля версий (Git, SVN и др.). - Встроенный терминал. - Поддержка Jupyter Notebooks. - Интеграция с базами данных. - Плагины для расширения функциональности. PyCharm Professional Edition также включает поддержку веб-разработки на фреймворках Django и Flask, а также инструменты для анализа данных и машинного обучения.

86. Visual Studio Code: Visual Studio Code (VS Code) - это легковесная, но мощная среда разработки с открытым исходным кодом,ная Microsoft. Она поддерживает множество языков программирования, включая Python, и предлагает обширный набор функций через расширения. Возможности VS Code включают: - Автозавершение кода и проверка синтаксиса. - Встроенный отладчик с поддержкой нескольких языков. - Поддержка виртуальных окружений и менеджер пакетов. - Интеграция с системами контроля версий (Git и др.). - Встроенный терминал и интегрированный поиск. - Поддержка Jupyter Notebooks через расширение. - Широкий выбор расширений для различных задач, включая анализ данных и машинное обучение. VS Code является гибким инструментом, который может быть настроен для решения широкого круга задач в области анализа данных и машинного обучения, благодаря обширному набору расширений и плагинов.

87. Microsoft Visual Studio: Это интегрированная среда разработки (IDE), которая поддерживает множество языков программирования, включая C#, C++, Python и другие. В области анализа данных и машинного обучения, Visual Studio может использоваться для разработки и отладки сложных алгоритмов, а также для создания пользовательских приложений, которые могут обрабатывать и анализировать большие объемы данных.

88. R-Studio: Это мощная и гибкая среда для работы с языком R, который широко используется в статистическом анализе и машинном обучении. R-Studio предоставляет обширный набор инструментов для анализа данных, включая графические средства для визуализации данных, а также интегрированный отладчик и редактор кода.

89. Microsoft Excel и NodeXL: Excel является широко используемым инструментом для работы с данными, предоставляя возможности для сортировки, фильтрации, визуализации и базового анализа данных. NodeXL — это дополнение для Excel, которое позволяет анализировать и визуализировать социальные сети. Вместе они могут использоваться для предварительной обработки данных и проведения базового анализа.

90. MatLab: Это высокоуровневый язык и интерактивная среда для вычислений, визуализации и программирования. MatLab широко используется в научных и инженерных расчетах, включая обработку сигналов, анализ данных и моделирование систем. В области машинного обучения, MatLab предоставляет библиотеки для алгоритмов обучения, а также мощные инструменты для визуализации результатов.

91 Statistica - это программное обеспечение для статистического анализа и работы с данными. Оно предоставляет широкий спектр статистических методов, включая простые статистики, регрессионный анализ, кластерный анализ, факторный анализ и многомерные методы. Statistica также поддерживает визуализацию данных и может быть использовано для предварительной обработки данных, моделирования и проверки гипотез. В области анализа данных и машинного обучения Statistica часто используется для проведения сложных статистических исследований и для разработки моделей прогнозирования.

92. Google Colab - это бесплатный облачный сервис для работы с Jupyter Notebooks, который предоставляет доступ к вычислительным ресурсам, включая GPU и TPU, для обучения моделей машинного обучения. Google Colab позволяет пользователям писать и выполнять код на нескольких языках программирования, включая Python, а также предоставляет интерактивные визуализации и возможность делиться своими работами с другими пользователями. В области анализа данных и машинного обучения Google Colab часто используется для обучения и тестирования моделей машинного обучения, а также для проведения экспериментов с данными.

93. Anaconda - это дистрибутив Python и R, предназначенный для научных вычислений и работы с данными. Он включает в себя среду разработки Anaconda Navigator и множество популярных пакетов для анализа данных, таких как NumPy, Pandas, Matplotlib и Scikit-learn. Anaconda также предоставляет инструменты для управления зависимостями и версиями пакетов, а также для создания изолированных сред разработки (conda environments). В области анализа данных и машинного обучения Anaconda часто используется для установки и управления пакетами, необходимыми для работы с данными и машинного обучения, а также для обеспечения совместимости и воспроизводимости результатов.

94. MatLab - это высокоуровневый язык и интерактивная среда для численных вычислений, визуализации данныхи программирования. MatLab предоставляет широкий набор инструментов для анализа данных, моделирования и алгоритмизации. В области анализа данных и машинного обучения MatLab часто используется для разработки и реализации алгоритмов обработки сигналов, систем управления, статистического анализа и машинного обучения. Он также поддерживает работу с матрицами и векторами, что делает его удобным инструментом для математических вычислений.

95. Power BI. Возможности программногосредства и наиболее часто решаемые задачи в области анализа данных и машинного обучения. Power BI - это набор инструментов бизнес-аналитики от Microsoft, предназначенный для анализа данных и обмена сведениями. Он позволяет создавать интерактивные визуализации данных и отчеты, а также предоставляет возможности для их публикации и совместного использования. В области анализа данных и машинного обучения Power BI часто используется для: - Визуализации данных: создание диаграмм, графиков, карт и других визуальных представлений для наглядного отображения данных. - Анализ данных: проведение детального анализа данных с использованием различных методов и алгоритмов. - Создание отчетов: разработка интерактивных отчетов, которые могут быть легко обновлены и распространены среди коллег. - Интеграция с другими инструментами: возможность работы с данными из различных источников, включая базы данных, файлы Excel, облачные сервисы и т.д. - Совместная работа: обмен данными и отчетами с другими пользователями, а также возможность работы в режиме реального времени.

96. OWOX BI. Возможности программного средства и наиболее часто решаемые задачи в области анализа данных и машинного обучения. OWOX BI - это комплексный инструмент для бизнес-аналитики, который предоставляет широкие возможности для сбора, анализа и визуализации данных. В области анализа данных и машинного обучения OWOX BI часто используется для: - Сбор данных из различных источников: интеграция с системами веб-аналитики, CRM, ERP и другими источниками данных. - Обработка и трансформация данных: преобразование данных для подготовки их к анализу и визуализации. - Анализ данных: проведение детальногоанализа данных с использованием различных методов и алгоритмов. - Визуализация данных: создание диаграмм, графиков, карт и других визуальных представлений для наглядного отображения данных. - Создание отчетов: разработка интерактивных отчетов, которые могут быть легко обновлены и распространены среди коллег.

97. PolyAnalyst. Возможности программного средства и наиболее часто решаемые задачи в области анализа данных и машинного обучения. PolyAnalyst - это мощный инструмент для анализа данных, который предоставляет широкие возможности для обработки и анализа больших объемов данных. В области анализа данных и машинного обучения PolyAnalyst часто используется для: - Обработка и трансформация данных: преобразование данных для подготовки их к анализу и визуализации. - Анализ данных: проведение детального анализа данных с использованием различных методов и алгоритмов, включая статистические методы, машинное обучение и др. - Визуализация данных: создание диаграмм, графиков, карт и других визуальных представлений для наглядного отображения данных. - Создание отчетов: разработка интерактивных отчетов, которые могут быть легко обновлены и распространены среди коллег. - Интеграция с другими инструментами: возможность работы с данными из различных источников, включая базы данных, файлы Excel, облачные сервисы и т.д.

98. Polymatica. Возможности программного средства и наиболее часто решаемые задачи в области анализа данных и машинного обучения. Polymatica - это инструмент для анализа данных и бизнес-аналитики, который предоставляет широкие возможности для обработки и анализа больших объемов данных. В области анализа данных и машинного обучения Polymatica часто используется для: - Обработка и трансформация данных: преобразование данных для подготовки их к анализу и визуализации. - Анализ данных: проведение детального анализа данных с использованием различных методов и алгоритмов, включая статистические методы, машинное обучение и др. - Визуализация данных: создание диаграмм, графиков, карт и других визуальных представлений для наглядного отображения данных. - Создание отчетов: разработка интерактивных отчетов, которые могут быть легко обновлены и распространены среди коллег. - Интеграция с другими инструментами: возможность работы с данными из различных источников, включая базы данных, файлы Excel, облачные сервисы и т.д.

**Список литературы:**

1. Уэс, Маккинли. Python и анализ данных Электронный ресурс / Маккинли Уэс ; пер. А. А. Слинкин. - Python и анализ данных,2022-04-19. - Саратов : Профобразование, 2017. - 482 с. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4488-0046-7, экземпляров неограниченно.

2. Сузи, Р.А. Язык программирования Python Электронный ресурс : учебное пособие / Р.А. Сузи. - Язык программирования Python,2020-07-28. - Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. - 350 c. - Книга находится в базовой версии ЭБС IPRbooks. - ISBN 5-9556-0058-2, экземпляров неограниченно.

3. Стенли, Липпман. Язык программирования С++ Электронный ресурс : Полное руководство / Липпман Стенли, Лажойе Жози ; пер. А. Слинкин. - Язык программирования С++,2022-04-19. - Саратов : Профобразование, 2017. - 1104 с. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4488-0136-5, экземпляров неограниченно

4. https://github.com/enikolaev/MMO – Репозиторий с примерами кода из лабораторных работ.

5. https://archive.ics.uci.edu/ml/index.html – Репозиторий наборов данных для машинного обучения (Центр машинного обучения и интеллектуальных систем).

6. https://www.kaggle.com – Портал и система проведения соревнований по проблемам анализа данных.

7. https://www.mockaroo.com – Сайт для генерации наборов данных.

8. <https://habr.com/ru/articles/349204> – Статья на хабре: «Моделирование динамических систем: введение в GNU Octave»