4 слайд

Основными причинами выбора данной темы являются актуальность и потребности бизнеса.

То есть, социальные медиа становятся все более значимым инструментом для коммуникации, маркетинга и анализа данных. Разработка приложения для вычисления метрик в таких сетях поможет отслеживать эффективность деятельности пользователей в социальных медиа и принимать обоснованные решения на основе аналитики.

Также многие компании и маркетологи активно используют социальные медиа для продвижения продуктов и услуг. Понимание эффективности распространения информации важно для оптимизации стратегий.

На графике представлено количество пользователей социальных сетей, на котором видно, что с каждым годом это количество увеличивается.

5 слайд

Объектом исследования являются характеристики пользователей социальных медиа.

Предметом исследования являются алгоритмы определения метрик.

6 слайд

Основная цель работы заключается в создании программного инструмента, способного собирать информацию о пользователях социальной сети, вычислять различные метрики и на основе полученных данных определять наиболее влиятельных пользователей для лучшего распространения информации.

Необходимо: провести анализ предметной области, изучить существующие решения, создать модель виртуальной сети в СИМ, спроектировать приложения, реализовать приложения, провести тестирование, развернуть приложение на сервере.

7 слайды

Вся работа будет проходить в несколько этапов: необходимо реализовать сервис и приложение для вычисления метрик пользователей, модель виртуальной сети в системе информационного моделирования, организовать сбор информации в онтологии и отслеживание через журналы событий.

9 слайд

* На слайде представлены существующие решения для вычисления метрик и анализа социальных сетей
* Каждое решение предлагает свой набор функций и инструментов для анализа метрик, что позволяет выбрать наиболее подходящее в зависимости от потребностей и целей бизнеса.
* Некоторые решения, такие как Buffer и Sprout Social, известны своим простым и интуитивно понятным интерфейсом, что делает их привлекательными для небольших компаний или предпринимателей.
* Решения типа Google Analytics предлагают более широкий набор функций и гибкость в анализе данных, что делает их предпочтительными для крупных компаний с сложными потребностями в аналитике.
* Важным фактором при выборе решения является его стоимость. Некоторые решения, такие как Facebook Analytics и Google Analytics, предоставляют базовый функционал бесплатно, в то время как другие, например, Sprout Social и Talkwalker, требуют платных подписок.
* При выборе решения необходимо учитывать его возможность интеграции с другими инструментами и платформами, такими как социальные сети, CRM-системы и т.д.

10 слайд

Метрики позволяют отслеживать эффективность работы в социальных сетях и планировать выпуск контента или проведение рекламных кампаний.

Взаимная направленность — это свойство, которое указывает на то, является ли отношение между вершинами бинарным.

Гомогенность — это свойство, которое указывает на степень появления связей между акторами, которые имеют схожие характеристики, такие как пол, возраст или интересы.

Транзитивность связей — это свойство, которое указывает на увеличение вероятности появления связей между акторами, у которых есть связи с одними и теми же вершинами.

Разница в распределении — это свойство, которое указывает на то, что у одних акторов есть большое количество связей, а у других - минимальное.

Центральность — это метрика, которая позволяет определить значительность или влияние определенного узла или группы в сети.

Ассортативность — это свойство, которое указывает на склонность к образованию связей между вершинами большой степени.

PageRank — это алгоритм, разработанный компанией Google, который используется для оценки важности веб-страниц на основе структуры ссылок между ними.

11 слайд

В мире социальных медиа взаимодействие пользователей и формирование связей играют ключевую роль в распространении информации и влиянии. Для анализа этих связей используются концепции виртуальных и реальных сетей.

Виртуальные сети представляют собой абстрактные модели социальных связей и взаимодействий между участниками. Они часто используются в системах информационного моделирования для визуализации и анализа социальных структур.

Реальные сети представляют собой фактические социальные связи между реальными пользователями социальных медиа. Они формируются на основе действительных взаимодействий и коммуникаций между участниками.

12 слайд

Для исследования социальной сети требуется анализ информационных каналов, то есть, исследование того, как пользователи взаимодействуют друг с другом и откуда они получают информацию. Эффективным методом представления социальной сети является использование графа, где узлы представляют собой пользователей или сообщества, а рёбра отображают каналы связи между ними.

Возникает вопрос: в каком формате следует хранить информацию о реальной социальной сети?

Для решения этой задачи было предложено применять онтологии.

Онтология — это детальное определение концепции или предметной области. Под детальностью понимается способ представления понятий и их определений таким образом, чтобы они могли быть однозначно поняты как человеком, так и компьютером.

Онтологии служат своего рода словарем для представления и обмена знаниями в определенной предметной области, а также определяют связи между терминами в этом словаре

13 слайд

Для успешной реализации приложения необходимо проанализировать основные бизнес-процессы, которые будут выполняться в системе.

Процесс сбора данных о пользователях социальной сети. **Прецеденты данного процесса**: **автоматизированное сканирование профилей, Сбор социальных связей, Обработка и фильтрация данных.**

Процесс анализа метрик и оценки влияния пользователей. Прецеденты: Вычисление метрик активности, Оценка влияния.

Процесс мониторинга и анализа результатов. **Прецеденты:** **Постоянный мониторинг активности, Анализ результатов.**

На слайде представлены диаграмма бизнес-процесса и диаграмма последовательности прецедента.

15 слайд

С целью изучения возможности распространения информации в социальных сетях, с учетом активности пользователей и степени реализации возможностей влияния на социальные сети, в данной работе применяется совместная разработка агентского подхода (Agent-based modelling) и математического моделирования. Данный графовый подход, который можно назвать моделью Барабаши-Альберта или же моделью агентного моделирования, применяется в совокупности с концепцией графового подхода, которая позволяет наглядно представить взаимодействие и взаимосвязи между агентами социальной сети, а также визуализировать динамику распространения информации.

В данной исследовательской работе в качестве основы для моделирования процесса распространения контента используется модель SEIR, которая базируется на принципах эпидемиологии. Такой выбор обоснован тем, что эти модели считаются наиболее соответствующими для анализа социальных сетей, поскольку механизм передачи информации в таких сетях аналогичен механизму распространения инфекций.

В модели выделяются следующие категории: Восприимчивые - пользователи, которые еще не видели информацию; Воздействованные - пользователи, которые видели информацию, но еще не начали ее распространять; Инфекционные - пользователи, которые начали активно распространять информацию среди своих подписчиков и друзей; Выздоровевшие - пользователи, которые уже видели информацию и перестали ее активно распространять.

16 слайд

В процессе реализации были созданы необходимые агенты, такие как User и Main, которая содержит популяцию User.

Реализована диаграмма состояний с переходами между ними представленная на слайде.

Также на слайде представлена работа модели, где виден процесс распространения, состояние пользователей и гистограмма, оценивающая центральность, то есть количество связей между users.

18 слайд

* В качестве СУБД для хранения информации о пользователя была выбрана PostgreSQL.
* В качестве языка программирования был выбран Python, так имеет множество полезных возможностей.
* Для вывода данных и удобного развертывания системы был выбрал фреймворк Flask.
* Для взаимодействия и отображения онтологий будет использоваться Protégé.

19 слайд

Для реализации сбора и отслеживания информации изначально необходимо определить сущности, которые будут отслеживаться, то есть пользователь, сообщество и пост, то есть класс Person, Community и Post.

Следующее, что необходимо сделать сформировать методы для взаимодействия с VK API, для этого создан класс VK, который обрабатывает взаимодействие с VK API, включая получение данных о пользователях, сообществах и записях, проверку статуса онлайн пользователя и получение понравившихся записей.

Далее реализуем класс Ontology, который управляет созданием и сохранением онтологии, а также предоставляет методы для сохранения информации о пользователях, сообществах и записях в онтологию.

Класс EventLog управляет ведением журнала событий, включая добавление трассировок, запись журналов в файлы и определение различных типов событий, таких как статус онлайн, добавление, просмотр, лайк и копирование записей.

На слайде представлены полученная диаграмма классов, структура получившейся онтологии и журнал событий.

20 слайд

Была реализована База данных, которая будет содержать информацию о пользователях.

Также были реализованы методы для взаимодействия с VK API.

Следующим шагом является создание графа друзей определенного пользователя, то есть графа реальной сети. С помощью этого графа будет возможно вычислить метрики.

На слайде представлены полученные диаграммы классов, реализованный граф друзей пользователя, и посчитанные метрики у пользователей, которые сохранены в базу данных.

21 слайд

Веб приложение отвечает за просмотр собранной информации.

Первый элемент – это таблица, собранной информации о пользователях.

Следующим элементом является слайдер графиков для сравнения различных метрик у пользователя. Для реализации слайдера использовался Bootstrap.

Третий элемент информационный, он содержит описания основных метрик социального графа

22 слайд

Было проведено:

* Тестирование сбора информации в онтологии и отслеживания через журналы событий.
* Проверка точности вычисления метрик.
* Для тестирования функциональности веб-приложения был использован метод, исключающий применение автоматизированных средств и включающий активное участие группы пользователей.

Для успешного развертывания системы используется Docker. Docker позволяет упаковать приложение и его зависимости в контейнер для обеспечения единообразной среды выполнения на различных платформах, необходимо создать Dockerfile в корневой директории приложения и настроить Docker Compose для контейнеризации приложения.

24 слайд

Подведение итогов:

* Создана модель виртуальной сети в системе информационного моделирования, включающая в себя визуализацию виртуальной сети
* Проведен анализ реальной сети с целью реализации сервиса для вычисления метрик на основе фактических данных
* Проведена работа по сбору информации в онтологии и отслеживанию через журналы событий

Потенциал для дальнейшего развития:

* улучшение методов анализа
* интеграция новых источников данных
* разработка персонализированных рекомендаций
* анализ влияния внешних факторов
* более глубокое применение машинного обучения и искусственного интеллекта