**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ**

**РОССИЙСКОЙ ФЕДЕРАЦИИ**

Федеральное государственное автономное образовательное учреждение высшего образования

«Южный федеральный университет»

Институт математики, механики и компьютерных наук им. И.И.Воровича

Кафедра прикладной математики и программирования

Направление 01.03.02 Прикладная математика и информатика

**ОТЧЕТ**

о производственной практике

Бакалавра 4 курса

Боруса Александра Игоревича

Руководитель практики   
от структурного подразделения ЮФУ:

доц. Махно В.В.

г. Ростов-на-Дону 2020

**Требования**

**по оформлению отчета о прохождении практики**

Объем отчета не должен превышать 10-15 страниц печатного текста, формат А4, шрифт 14, Times New Roman, интервал полуторный.

Структура отчета:

- титульный лист;

- содержание;

- введение;

- пояснительная записка (цели и задачи работы, методика проведения исследований, анализ полученных результатов);

-список использованных источников;

-приложение (материалы и документы, предоставленные организацией, методические материалы, т.п.).

Отчет – это самостоятельный документ, который студент представляет на зачет по практике.

Отчет по практике оформляется по мере изучения материала в соответствии с программой практики.Отчет по практике должен содержать анализ изучаемых материалов, конкретные расчеты, лично проведенные исследования. По материалам проведенных исследований должны быть сделаны выводы и предложения. Анализ материалов и представленные выводы должны отличаться самостоятельностью суждений.

Отчет составляется студентом в конце практики, к которому прилагаются материалы, собранные в период практики. Отчет представляется руководителю от профильной организации, который подписывает его.

**Оглавление**

Введение 4

1. Цели и задачи работы 4

2. Возможности социальных сетей и мессенджеров 5

3. Основной параметр выбора платформы. Cредства взаимодействия с API 7

4. Анализ возможностей, выбор платформы и средств разработки 7

5. Aнализ результатов опроса. Определение методов анализа 9

Заключение 11

Список литературы 12

Введение

Развитие цифровых технологий и сети Интернет за последние десятилетия оказывает существенное влияние на стиль и образ жизни общества на современном этапе. Как молодежь, так и старшее поколение активно применяют в своей повседневной жизни социальные сети и мессенджеры. Удобство, комфорт, доступность позволяет таким элементам интернета значительно улучшить качество жизни населения. Пользователи за доли секунды связываются друг с другом в любой точке земного шара, получают ответы на любой вопрос, удовлетворяют любые свои желания и потребности при помощи пары кликов. Создание приложений на базе таких коммуникационных площадок позволяет получить очень важные для разработчиков возможности. Например: прямой диалог и обратная связь с пользователями, сбор разнообразных обучающих материалов, более легкое отслеживание необходимой статистики, относительно низкая стоимость содержания приложений и т. п.

1. Цели и задачи работы

Цель работы: изучение возможностей социальных сетей и мессенджеров для определения наиболее функциональной платформы реализации системы, позволяющей накапливать и интерпретировать высказывания пользователей, выбранной социальной платформы, по некоторой фиксированной теме в режиме интерактивного взаимодействия. Система выступает инициатором общения с пользователями, периодически отправляя сообщения с просьбой выразить своё мнение по некоторому вопросу, и анализирует полученные ответы. Считаем, что для интерпретации высказываний пользователей реализован механизм анализа фраз, например с помощью других пользователей, а также учёт недобросовестных (заведомо ложных) комментариев. Изучить подходы к анализу результатов опроса.

Темой общения будем считать фильмы, просмотренные пользователями. Мнениями - оценки и комментарии, которые пользователи оставляют к предложенным (ими или программой) фильмам. Анализ комментария – определение соответствует ли оценка и комментарий друг другу.

Задачи работы:

– изучение возможностей социальных сетей и мессенджеров для определения наиболее функциональной платформы реализации системы;

– сравнительный анализ систем разработки приложений для социальных сетей и мессенджеров;

– изучение подходов и методов статистического анализа результатов опросов;

2. Возможности социальных сетей и мессенджеров

На данные момент не существует общепринятой классификации социальных средств коммуникации. Ряд существующих социальных сетей очень широк – от микроблогинга, до бизнес-сетей и виртуальных миров, и для лучшего понимания каждого типа, представляется удобным разделить их по функциональным возможностям:

1. Социальные сети для общения (Relationnetworks) – самый популярный и востребованный тип. К таким социальным сетям относятся: «Facebook», «Вконтакте», «Одноклассники» и др. Взаимоотношения в данной группе сетей могут быть как личными, так и профессиональными. Организации могут вести профессиональные сообщества, группы, корпоративные страницы.
2. Сети, предназначенные для обмена медиа-контентом, таким как фото, видео, аудио (Mediasharingnetworks). В данной группе представлены: «Instagram», «YouTube», «Vine», «Snapchat» и др.
3. Социальные платформы для публикации отзывов и обзоров (Onlinereviews). Представляют собой базы знаний для пользователей и разработчиков, желающих собрать необходимую информацию о чем-либо. Например, для зарубежных сетей – это «Yelp», «Forsquare», а российский аналог «Flamp», где используется геолокация и возможность оставлять отзывы и рекомендации о локальном бизнесе. Также существуют ресурсы, на котором пользователи оставляют отзывы о покупках и другие услуги, в России – это «Яндекс Маркет», «Отзыв.ру», «Irecommend» и подобные.
4. Для коллективных обсуждений (Discussion forums/Social news). Это Различные форумы сообщества и Q&A-сервисы (вопрос-ответ) были одними из первых социальных медиа, к современным платформам можно отнести «Mail.ru», «Reddit», и т. п.
5. Сети для авторских записей (Social publishing platforms) представляют собой блоггинг, микро-блоггинг, с помощью которого люди создают и обмениваются текстово-медийный контентом. В данную группу входят: «Twitter», «Tumblr», и т. п.
6. Сервисы социальных закладок (Bookmarkingsites). Представители данного вида – «Pinterest», «Flipboard», «Scoop.it!», «Diigo» и другие. По своей сути –это сервисы, на которых пользователи собирают свою личную библиотеку, на которую впрочем могут подписываться и другие пользователи. Обычно, такие социальные сети занимаются изучением интересов пользователей в целях более релевантного контента для каждого.
7. Социальные сети по интересам (Interest-based networks) – это сети, для нахождения единомышленников. «Last.fm» – для меломанов, «Goodreads» – для любителей литературы. Данный вид социальных медиа отлично подходит для создания отраслевых сообществ.
8. Профессиональные социальные сети (Professionnetworks). Ресурсы, главной целью которых является профессиональное общение. Данные сети отлично подходят для рекрутинга, создания бизнес-проектов, продвижения товаров и услуг в бизнес-среде.

В данный момент времени виды социальных сетей очень разнообразны, что позволяет удовлетворять различные потребности пользователей и организаций, кроме того отличительной чертой социальных сетей является возможность точно взаимодействовать с целевой аудиторией, выбирать площадки с ее наибольшей концентрацией и более эффективные способы коммуникации.

Далее поговорим о мессенджерах. Многие не видят разницы между социальными сетями и мессенджерами, однако определенная разница все же существует. Социальные сети дают больше возможностей для удовлетворения потребностей и интересов пользователей. Мессенджеры же — это программы, которые работают только с целью обмена текстовыми сообщениями, голосовыми или видео вызовами, а также обмена медиа файлами. В каком-то роде их можно назвать упрощенными версиями соцсетей 1-го типа, и, хотя возможности таких социальных платформ постоянно пополняются, все же их функционал в разы меньше.

3. Основной параметр выбора платформы. Cредства  
взаимодействия с API

Т. к. для корректного функционирования моей программы необходимы возможности отправки, приема и обработки сообщений, то, следовательно, мне подойдут социальные сети 1-го типа и мессенджеры. Для взаимодействия с пользователями важны средства работы с интерфейсом прикладного программирования (API).

API — это платформа, предоставляемая социальными сетями, позволяющая другим приложениям и веб-сайтам получать данные социальных сетей и интегрироваться с их сайтом или приложением. В случае моей задачи легкий и удобный для взаимодействия API — основной параметр выбора площадки реализации программы. Существует множество соцсетей и мессенджеров с разнообразными возможностями работы с их API. Самые популярные из них предоставляют разработчикам библиотеки для работы с их интерфейсами, написанными на различных языках программирования.

Сейчас популярны, так называемые “Чатботы” — программы, работающие внутри мессенджера или соцсети, которые по заданному плану совершают какие-либо действия через их API. Они способны отвечать на вопросы, а также самостоятельно задавать их. Зачастую в таких программах используют экспертные системы, применимые в разных сферах, решая типовые задачи. Самыми популярными платформами, где используют ботов являются “ВКонтакте” и “Telegram”. Обе предоставляют отличные библиотеки, с подробной документацией, для работы с API, а также, что не мало важно, совершенно бесплатные, в отличии от того же “Whatsapp”.

Для поставленной задачи написание такого “Чатбота” с использование экспертной системы оценки комментариев наиболее эффективный выбор.

4. Анализ возможностей, выбор платформы и средств разработки

Разница во взаимодействии ботов с API мессенджера или соцсети, к которой они относится, заключается в строении самой социальной платформы, в возможностях взаимодействия с пользователем, количестве параметров у объектов (сообщений, CallBack ответов и т. д.) и многом другом. Так, например, при создании чатбота в “ВКонтакте” нужно обязательно соединить его с сообществом, от имени которого он будет общаться-взаимодействовать с пользователями “ВКонтакте”, соответственно ему доступны не только работа с чатом, но и с постами. При создании же чатбота в “Telegram” мы не обязаны прикреплять его ни к каким каналам. Однако у обоих платформ получение событий и обновлений от бота реализовано одинаково. И там, и там существует два диаметрально противоположных по логике метода — Callback(Webhook) и Long Poll.

CallBack(Webhook) - сервер присылает уведомления на сервер клиента, как только происходит определенное событие. Событие может быть каким угодно: комментарий к фотографии, новая запись на стене, вступление в чат, отправка сообщения, и многое другое.

LongPoll - клиент запрашивает информацию с сервера точно так же, как при обычном опросе, но с ожиданием, что сервер может не ответить немедленно. Если на сервере нет новой информации для клиента при получении опроса, вместо отправки пустого ответа сервер удерживает запрос открытым и ожидает, когда информация ответа станет доступной. Как только он получит новую информацию, сервер немедленно отправит клиенту ответ. После получения ответа от сервера клиент часто сразу же отправляет другой запрос к серверу.

Без необходимости привязки к сообществу и работы с постами, мной при разработке программы был выбран мессенджер Telegram, 2ой вариант получения обновлений и библиотека TelegramBotAPI, доступная на разных языках. Наиболее обширные варианты библиотеки написаны на Python, Java, PHP. Мной был выбран Python в связи с удобством реализации экспертной системы и организации элементарной работы с базой данных. Средой разработки выбран PyCharm. Он имеет возможности интегрированного модульного тестирования, проверки кода, интегрированного контроля версий, инструменты рефакторинга кода, набор инструментов для навигации проекта, выделения и автоматического завершения. Поддержка ряда сторонних фреймворков для веб-разработки, что делает его универсальной IDE (Integrated Development Environment) для быстрой разработки приложений.

5. Aнализ результатов опроса. Определение методов анализа

Начнем с определения понятия опрос:

Опрос — это метод сбора первичной информации, основанный на непосредственном или опосредованном взаимодействии исследователя и опрашиваемого. Источником информации служит словесное или письменное суждение человека. Данный метод широко распространен из-за его универсальности, сравнительной легкости применения и обработки данных. Однако не полностью раскрывает исследуемую тему. Одна из трудностей, с которой сталкивается исследователь, применяющий методы опроса – это обеспечение достоверности и надежности полученных данных. Информация, которую он получает, носит субъективный характер, так как зависит от степени искренности отвечающего и ряда других факторов. Поэтому данные, полученные в результате опроса, следует сопоставлять с данными, полученными другими методами (эксперимент, наблюдение, анализ документации и т. п.).

Методология обработки анализа результатов опроса необычайно обширна, и варьируется в зависимости от задачи. Для выбора подходящих нам методов, определим какую информацию мы получим в результате опроса:

1. Название фильма
2. Оценка фильма от опрашиваемого
3. Комментарий от опрашиваемого

Мы имеем объект изучения, оцениваемые и не оцениваемые (комментарий) параметры. Нас интересует соотносится ли информация, содержащаяся в комментарии, с оставленной оценкой. Следовательно, выбор метода анализа результатов опроса сводится к выбору метода определения тональности текста. Тональность текста подразумевает эмоциональное отношение автора высказывания к некоторому объекту (объекту реального мира, событию, процессу и т. п.), выраженному в тексте.

На данный момент существуют три основных метода определения тональности текста:

1. Анализ методами векторного анализа и машинного обучения, сравнение с ранее размеченным эталонным корпусом по выбранной мере близости и отнесение (классификация) текста к негативу или позитиву на основании полученного результата сравнения. Данный метод работает достаточно быстро, но требует наличия предварительно размеченного эталонного корпуса, на основе которого происходит обучение алгоритма сравнения. Существенными недостатками такого подхода оказываются увеличение трудоемкости и ограничение разнородности корпуса (т. е. неполнота лексического покрытия), что приводит к потере точности. К тому же данный метод не позволяет провести глубокий анализ текста, то есть выявить и показать эмотивность на уровне предложения.   
   Примером может служить Кросс-валидация – методика, в основе которой лежит разделение исходного множества данных на k примерно равных блоков. Затем на k-1-ом блоке, производится обучение модели, а k-ый блок используется для тестирования. Процедура повторяется k раз, при этом на каждом проходе для проверки выбирается новый блок, а обучение производится на оставшихся. При k=5, Схема-1

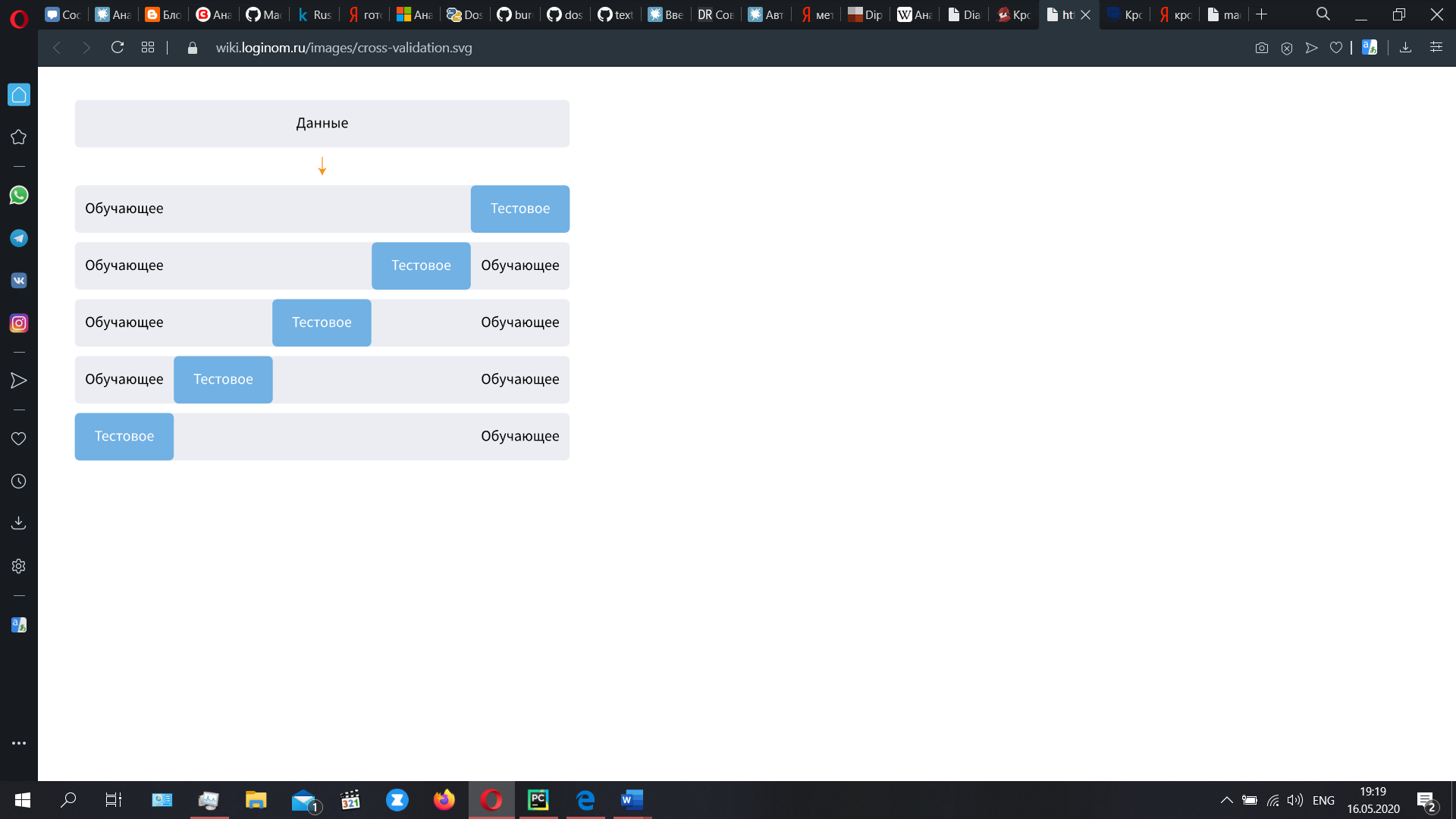


Схема-1.

1. Поиск эмотивной лексики (лексической тональности) в тексте по заранее составленным тональным словарям (спискам паттернов) с применением лингвистического анализа. По совокупности найденной эмотивной лексики текст может быть оценен по шкале, отражающей количество негативной и позитивной лексики. Этот метод может использовать как списки паттернов, подставляемые в регулярные выражения, так и правила соединения тональной лексики внутри предложения. Является не менее трудоемким из-за составления тональных словарей, но в сочетании с синтаксическим и морфологическим анализом более гибок: он позволяет не только показать цепочки тональной лексики, но и получить синтаксически корректные эмоциональные выражения. При хорошем наполнении тональных словарных списков этот метод позволяет достичь хорошего покрытия эмотивной лексики. Недостаток этого метода в том, что с помощью него сложно дать количественную оценку негативности-позитивности текста.

Пример словаря:

|  |  |
| --- | --- |
| Слово | Оценка |
| Хороший | 1 |
| Великолепный | 2 |
| Никакой | 0 |
| Отвратительный | -2 |

1. Смешанный метод (комбинация первого и второго подходов).

Заключение

Изучил функциональные возможности социальных сетей и мессенджеров, определил наиболее функциональную платформу реализации системы. Провел сравнительный анализ систем разработки приложений для социальных сетей и мессенджеров. Изучил подходы и методы анализа результатов опроса системы.

Список литературы

1. Олейник Р. В., Исследовательская работа “Интернет-мессенджеры”
2. Шиманская Д. В., “Социальные сети и мессенджеры, как каналы коммуникации в связях с общественностью”
3. Официальная документация VK для работы с API: <https://vk.com/dev/manuals>
4. Официальная документация Telegram для работы с API: <https://tlgrm.ru/docs>
5. Материалы сайта <http://wikipedia.org>
6. Пазельская А. Г., Соловьев А. Н., “ Метод определения эмоций в текстах на русском языке”