

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования

«ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ» (ТУСУР)

Кафедра комплексной информационной безопасности электронно-
вычислительных систем (КИБЭВС)

УТВЕРЖДАЮ

Заведующий каф. КИБЭВС

_____ Шелупанов А.А.

_____._____._____

дата

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ «АНАЛИЗАТОР
ТОНАЛЬНОСТИ»

Курсовая работа по дисциплине «Технологии и методы программирования»

Пояснительная записка

СОГЛАСОВАНО

Преподаватель каф. КИБЭВС

_____ Никифоров Д.С.

_____._____._____

дата

РАЗРАБОТЧИКИ

Студенты группы 726:

_____ Кравцова А.В.

_____ Грохотова Е.А.

_____ Васильева М.И.

_____._____._____

дата

Томск 2018

Реферат

Курсовая работа содержит 45 страниц, 10 рис., 1 таблица, 4 источника, 4 прил.

АНАЛИЗАТОР, ТОНАЛЬНОСТЬ, АВТОМАТИЗИРОВАННАЯ СИСТЕМА, АНАЛИЗ, ТЕКСТ.

Объектом разработки является автоматизированная система (АС) «Анализатор тональности».

Цель работы: разработка АС, анализирующей эмоциональную окраску вводимого текста.

Разработка АС проводилась на языке программирования Python с помощью программного обеспечения Python 3.6, Django 2.0, PyCharm 2018.2.

В результате работы была создана АС, анализирующая тональность осмысленного русского текста длиной не более 250 символов.

Достигнутые технико-эксплуатационные показатели: высокая скорость анализа текста.

Степень внедрения: АС внедрена в сеть Интернет и может использоваться любым человеком, подключённым к сети.

Область применения АС – анализ большого количества текстов.

В дальнейшем планируется развитие АС путём увеличения максимальной длины анализируемого текста.

Пояснительная записка выполнена в текстовом редакторе Microsoft Word 2016.

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования

«ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ
УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ» (ТУСУР)

Кафедра комплексной информационной безопасности электронно-
вычислительных систем (КИБЭВС)

УТВЕРЖДАЮ

Заведующий каф. КИБЭВС

_____ Шелупанов А.А.

_____._____._____

дата

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ «АНАЛИЗАТОР
ТОНАЛЬНОСТИ»

Курсовая работа по дисциплине «Технологии и методы программирования»
Техническое задание

СОГЛАСОВАНО

Преподаватель каф. КИБЭВС

_____ Никифоров Д.С.

_____._____._____

дата

РАЗРАБОТЧИКИ

Студенты группы 726:

_____ Кравцова А.В.

_____ Грохотова Е.А.

_____ Васильева М.И.

_____._____._____

дата

1 Общие сведения

Полное название системы: информационно-аналитическая система для определения тональности текста. Далее автоматизированная система (АС) будет именоваться «Анализатор тональности».

Заказчик: преподаватель кафедры комплексной информационной безопасности электронно-вычислительных систем (КИБЭВС) Томского Государственного Университета Систем Управления и Радиоэлектроники (ТУСУР) Никифоров Дмитрий Сергеевич.

Разработчики: студенты группы 726 факультета безопасности кафедры КИБЭВС ТУСУР Кравцова Анастасия Владимировна, Васильева Мария Игоревна и Грохотова Екатерина Андреевна.

Разработка системы ведётся на основании следующих документов:

- ГОСТ 34.602-89. Утверждён Государственным комитетом СССР по стандартам Министерством приборостроения, средств автоматизации и систем управления СССР 01.01.1990;
- Методические указания к курсовой работе по дисциплине «Методы программирования» для специальности 090105 Комплексное обеспечение информационной безопасности автоматизированных систем. Утверждено Е.М. Давыдовой и Р.В. Мещеряковым в 2012 году;
- Требования к курсовой работе. Утверждено Никифоровым Д.С. 13.09.2018.

Плановый срок начала работы по созданию системы 01.10.2018.

Плановый срок окончания работы по созданию системы 03.12.2018.

Работы не финансируются.

Результаты работы предъявляются в сроки, указанные в разделе 6, в виде документов, указанных в разделе 8.

2 Назначение и цели создания системы

2.1 Назначение системы

Система предназначена для управления автоматизированным анализом текстового сообщения. Использовать систему предполагается на коротких сообщениях, содержащих неформальную лексику.

2.2 Цель создания системы

Для данной автоматизированной системы показателем производительности является точность оценки тональности текстового сообщения.

Точность этой оценки есть критерии оценки достижения целей создания системы.

3 Характеристики объекта автоматизации

Объект автоматизации представляет собой анализ текстовых сообщений.

Климатические условия эксплуатации должны удовлетворять требованиям, предъявляемым к техническим средствам в части условий их эксплуатации.

4 Требования к системе

4.1 Требования к системе в целом

4.1.1 Требования к структуре и иерархии системы заключаются в создании клиент-серверной автоматизированной системы. Клиенты должны активно взаимодействовать с сервером, формируя и отправляя запросы. Сервер должен обрабатывать полученные запросы с помощью нейронной сети, обученной с помощью базы данных. Обмен информацией с сервером должен происходить автоматически. Для АС определен нормальный режим функционирования:

- серверное программное обеспечение и технические средства серверов обеспечивают возможность круглосуточного функционирования, с перерывами на обслуживание;
- исправно работает оборудование, составляющее комплекс технических средств;
- исправно функционирует системное, базовое и прикладное программное обеспечение системы.

Для обеспечения нормального режима функционирования системы необходимо выполнять требования и выдерживать условия эксплуатации программного обеспечения и комплекса технических средств системы, указанные в соответствующих технических документах (техническая документация, инструкции по эксплуатации и т.д.).

При отказе серверной части АС клиентам должно выводиться сообщение о недоступности АС.

В перспективе развития системы рассматривается расширение размера анализируемого текста.

4.1.2 Для работы с автоматизированной системой достаточно одного человека. Для работы с системой требуются базовые знания работы с компьютером и рекомендуется ознакомление с документацией по

использованию сервера. Требования к режиму работы персонала АС не предъявляются.

4.1.3 В системе должен быть обеспечен уровень надежности, включающий в себя сохранение работоспособности АС и всей накопленной информации при отказе или выходе из строя по любым причинам одного из технических средств, с последующим восстановлением функционирования системы после проведения ремонтных и восстановительных работ. Уровень надёжности также должен обеспечивать нормальное функционирование системы при количестве параллельно обрабатываемых запросов клиентов не более 5.

Перечень аварийных ситуаций, по которым должны быть регламентированы требования к надежности:

- при возникновении локальных отказов компонентов системы (отказ функционального рабочего места пользователя) должна обеспечиваться сохранность работоспособности системы;
- сохранность информации в системе должна обеспечиваться при нарушении электропитания или нарушении работы провайдера.

Время на восстановление работоспособности отдельных элементов сетевого и серверного оборудования не должно превышать два часа, в прочих случаях – определяется временем заказа и поставки необходимого оборудования. Для обеспечения электропитания серверного и другого оборудования, обеспечивающего функционирование АС, должны использоваться источники бесперебойного питания. Надежность рабочих мест должна быть обеспечена унификацией используемого системного ПО, централизованным хранением и копированием данных, а также системных настроек средствами резервного копирования. Выход из строя рабочего места пользователя не должен влиять на работоспособность системы в целом. Оценка надежности осуществляется на стадии проектирования за счет анализа полноты архитектуры и технических решений по построению системы и их соответствия техническим требованиям данного ТЗ.

4.1.4 Все внешние элементы технических средств системы, находящиеся под напряжением, должны иметь защиту от случайного прикосновения, а сами технические средства иметь зануление или защитное заземление в соответствии с ГОСТ 12.1.030-81 и ПУЭ. Система электропитания должна обеспечивать защитное отключение при перегрузках и коротких замыканиях в цепях нагрузки, а также аварийное ручное отключение. Общие требования пожарной безопасности должны соответствовать нормам на бытовое электрооборудование. В случае возгорания не должно выделяться ядовитых газов и дымов. После снятия электропитания должно быть допустимо применение любых средств пожаротушения. Факторы, оказывающие вредные воздействия на здоровье со стороны всех элементов системы (в том числе инфракрасное, ультрафиолетовое, рентгеновское и электромагнитное излучения, вибрация, шум, электростатические поля, ультразвук строчной частоты и т.д.), не должны превышать действующих норм (СанПиН 2.2.2/2.4.1340-03 от 03.06.2003 г.).

4.1.5 Взаимодействие пользователей с прикладным программным обеспечением, входящим в состав системы должно осуществляться посредством визуального графического интерфейса (GUI). Интерфейс системы должен быть понятным и удобным, не должен быть перегружен графическими элементами и должен обеспечивать быстрое отображение экранных форм. Навигационные элементы должны быть выполнены в удобной для пользователя форме. Средства редактирования информации должны удовлетворять принятым соглашениям в части использования функциональных клавиш, режимов работы, поиска, использования оконной системы. Ввод-вывод данных системы, прием управляющих команд и отображение результатов их исполнения должны выполняться в интерактивном режиме. Интерфейс должен соответствовать современным эргономическим требованиям и обеспечивать удобный доступ к основным функциям и операциям системы. Интерфейс должен быть рассчитан на преимущественное использование манипулятора типа «мышь», то есть управление системой должно осуществляться с помощью набора экранных меню, кнопок, значков и т. п. элементов. Клавиатурный режим ввода должен

использоваться главным образом при заполнении и/или редактировании текстовых экранных форм. Все надписи экранных форм, а также сообщения, выдаваемые пользователю (кроме системных сообщений) должны быть на русском языке.

Система должна обеспечивать корректную обработку аварийных ситуаций, вызванных неверными действиями пользователей, неверным форматом или недопустимыми значениями входных данных. В указанных случаях система должна выдавать пользователю соответствующие сообщения, после чего возвращаться в рабочее состояние, предшествовавшее неверной (недопустимой) команде или некорректному вводу данных.

Экранные формы должны проектироваться с учетом требований унификации:

- все экранные формы пользовательского интерфейса должны быть выполнены в едином графическом дизайне, с одинаковым расположением основных элементов управления и навигации;
- для обозначения сходных операций должны использоваться сходные графические значки, кнопки и другие управляющие (навигационные) элементы. Термины, используемые для обозначения типовых операций (добавление информационной сущности, редактирование поля данных), а также последовательности действий пользователя при их выполнении, должны быть унифицированы;
- внешнее поведение сходных элементов интерфейса (реакция на наведение указателя «мыши», переключение фокуса, нажатие кнопки) должны реализовываться одинаково для однотипных элементов. Система должна соответствовать требованиям эргономики и профессиональной медицины при условии комплектования высококачественным оборудованием (ПЭВМ, монитор и прочее оборудование), имеющим необходимые сертификаты соответствия и безопасности Росстандарта.

4.1.6 Требования к транспортабельности для АС не предъявляются.

4.1.7 АС должна бесперебойно работать в течении рабочего дня (с 9:00 до 20:00). АС должна обеспечивать одновременную работу максимум пяти пользователей. Для работы с автоматизированной системой достаточно одного человека. Для работы с системой требуются базовые знания работы с компьютером и рекомендуется ознакомление с документацией по использованию сервера. Требования к режиму работы персонала АС не предъявляются.

4.1.8 Требования к защите информации от несанкционированного доступа заключаются в обеспечении целостности и конфиденциальности информации, отправляемой клиентами на сервер.

4.1.9 Требования к защите информации от влияния внешних воздействий не предъявляются. Сохранность информации, являющейся объектом хранения системы, должна обеспечиваться при следующих аварийных ситуациях:

- сбой или выход из строя технических средств, на которых осуществляется эксплуатация системы;
- сбой системного программного обеспечения АС;
- сбой или отказ прикладного программного обеспечения АС из-за ошибок в настройках.

4.1.10 Требования к патентной чистоте не предъявляются.

4.1.11 Защита от влияния внешних воздействий должна обеспечиваться средствами владельца хостинга.

4.1.12 Система должна использовать современные протоколы TCP/IP, HTTP, HTTPS. В системе должны использоваться API готовых фреймворков.

4.2 Требования к функциям, выполняемым системой

Автоматизированная система должна выполнять одну функцию – анализ тональности текста. Требования к функции:

- выходная информация должна быть представлена в виде текста на русском языке;
- точность анализа текста не должна быть меньше 60%;
- функция должна корректно работать при одновременном подключении к системе не более пяти пользователей.

Корректность работы данной функции гарантируется при выполнении пользователем следующих требований:

- входная информация должна представлять собой осмысленный текст на русском языке от 200 до 300 символов без орфографических ошибок;
- пользователь подключён к сети Интернет;
- пользователь использует браузер.

При несоответствии входных данных заданным требованиям или одновременном подключении более пяти пользователей, функция может работать некорректно.

4.3 Требования к видам обеспечения

4.3.1 Требования к математическому обеспечению системы заключаются в создании рекуррентной нейронной сети с архитектурой долгой краткосрочной памятью (LSTM), выполняющей анализ тональности текста.

4.3.2 Состав, структура и способы организации данных в системе должны быть определены на этапе технического проектирования. Клиенты системы должны активно взаимодействовать с сервером, отправляя запросы. Обмен информацией с сервером должен происходить автоматически. При проектировании и развертывании системы необходимо рассмотреть возможность использования накопленной информации из уже функционирующих информационных систем.

4.3.3 Требования к лингвистическому обеспечению системы обуславливаются требованиями к программному обеспечению Python 3.6, Django 2.0, PyCharm 2018.2. АС должна быть разработана на языке программирования Python. Язык ввода/вывода данных должен быть русским.

4.3.4 Требования к программному обеспечению системы:

- поддержка архитектуры клиент-сервер.
- совместимость программных продуктов в части используемых технических средств, системного программного обеспечения в пределах требований к техническому обеспечению, а также их информационная совместимость в пределах требований к информационному обмену.
- русификация и наличие эксплуатационной документации на русском языке.

4.3.5 Требования к техническому обеспечению системы заключаются в наличии у клиента браузера и выхода в сеть Интернет. Для развёртывания сервера техническое обеспечение должно удовлетворять следующим требованиям:

- ОС Windows 7 и выше;
- минимум 2 ГБ ОЗУ, рекомендуется 4 ГБ ОЗУ;

- минимальное разрешение экрана 1024x786.

4.3.6 Требования к метрологическому обеспечению системы не предъявляются.

4.3.7 Для организационного обеспечения предлагается выполнение следующих пунктов:

- формирование регламентов взаимодействия участников проекта, разграничение их функций;
- разработка процедуры управления проектом на всех фазах жизненного цикла системы.

Для защиты от ошибочных действий персонала АС рекомендуется ознакомление персонала с инструкцией по использованию АС.

4.3.8 Нормативно техническая документация системы должна содержать в себе пояснительную записку, включающую в себя следующие разделы:

- описание разрабатываемой системы;
- описание архитектуры системы;
- описание протокола взаимодействия клиента и сервера;
- описание функций, предоставляемых сервером;
- описание инструментов, использованных при разработке;
- описание деталей реализации системы;
- описание системы тестирования;
- инструкция по настройке сервера;
- инструкция по использованию клиента.

К пояснительной записке должен прикладываться диск с:

- электронной версией пояснительной записки в формате PDF;
- исходные файлы пояснительной записки (docx, TeX и т.п.);
- ссылкой на репозиторий;
- файлом с описанием процесса установки сервера и получения клиента.

5 Состав и содержание работ по созданию системы

5.1 Подготовительный этап

Он включает в себя постановку задачи, сбор и анализ требований к разработке, тестирование, формирование целевой функции. Первыми проводятся системный анализ и анализ требований заказчика.

На этом этапе:

- выясняется, что необходимо заказчику;
- проводится оценка возможности выполнимости заказа (системы);
- определяются необходимые затраты (временные, организационные и технические);
- проводится распределение функций по элементам технической системы (людям, БД, аппаратуре, программам и т.д.);
- определяются ограничения на систему.

Результаты анализа сводятся в спецификацию требований к программному обеспечению, на основе чего формируется техническое задание на разрабатываемую автоматизированную систему.

5.2 Проектирование

Проектирование проводится в три этапа: концептуальное, логическое и физическое.

На этапе концептуального проектирования необходимо учесть требования заказчика и пользователей. Это сбор, документирование, проверка требований пользователей и выработка способов их реализации. Результатом является описание задачи и ее решения.

На этапе логического проектирования на основе концептуальной модели, построенной на предыдущем этапе, формулируется абстрактная модель решения:

- разрабатывается структура приложения, описание частей системы и их взаимодействие;
- описываются интерфейсы, обеспечивающие организационную структуру взаимодействия между компонентами и механизмы передачи параметров;
- выявляются ошибки концептуального проектирования.

Логический проект описывает, как должна работать система. Результатом является описание разрабатываемой системы, включающее в себя её архитектуру и описание всех её подсистем.

На этапе физического проектирования в проект включаются требования разработчиков. Результатом является описание сервисов и технологий, необходимых для реализации решения.

5.3 Реализация

Данный этап включает в себя:

- реализацию спроектированной АС;
- полнофункциональное тестирование.

Результаты тестирования должны соответствовать предъявляемым в техническом задании критериям качества. Отчёт о тестировании должен быть представлен в пояснительной записке.

6 Порядок контроля и приёмки системы

Сроки разработки и сдачи проекта приведены в таблице 6.1.

Таблица 6.1 – Сроки сдачи стадий проекта

Содержание этапа или стадии	Срок		Форма отчетности
Составление технического задания	14.09.18	11.10.18	ТЗ
Проектирование системы	11.10.18	20.10.18	Алгоритм
Составление системы и программ для тестирования	20.10.18	1.12.18	Код программы
Тестирование системы	1.11.18	20.11.18	Отчёт о тестировании
Составление пояснительной записки	21.11.18	3.12.18	ПЗ
Сдача проекта	4.12.18	29.12.8	Оценка

В состав приемной комиссии войдут сотрудники Томского Государственного Университета Систем Управления и Радиоэлектроники (ТУСУР) Факультета Безопасности (ФБ) с кафедры комплексной информационной безопасности электронно-вычислительных систем (КИБЭВС).

Прием данной системы будет осуществляться в Томском Государственном Университете Систем Управления и Радиоэлектроники (ТУСУР) на кафедре Комплексной Информационной Безопасности Электронно-Вычислительных Систем (КИБЭВС) в установленные планом сроки.

7 Требования к составу и содержанию работ по подготовке объекта автоматизации к вводу системы в действие

Инструкция по развёртыванию сервера и получению клиента должна быть в README файле в репозитории на <https://github.com>. Ссылка на репозиторий должна быть на диске, прилагаемом к пояснительной записке.

8 Требования к документированию

Состав программной документации:

- техническое задание (ТЗ);
- пояснительная записка (ПЗ);
- документация к серверу в электронном виде.

Документация должна быть оформлена с использованием:

- ГОСТ 34.602-89;
- ОС ТУСУР 01-2013 для технического задания;
- ГОСТ 19.701-90 для схем алгоритмов, программ, данных и систем;
- ОС ТУСУР 01-2013 для пояснительной записки.

9 Источники разработки

Техническое задание разрабатывалось на основе:

- ОС ТУСУР 01-2013;
- ГОСТ 34.602-89;
- задания на сайте <http://edu.fb.tusur.ru/course/view.php?id=1296>.

СОСТАВИЛИ

Наименование организации	Должность исполнителя	Фамилия, имя, отчество	Подпись	Дата
ТУСУР	студент гр. 726	Кравцова Анастасия Владимировна		
ТУСУР	студент гр. 726	Васильева Мария Игоревна		
ТУСУР	студент гр. 726	Грохотова Екатерина Андреевна		

СОГЛАСОВАНО

Наименование организации	Должность	Фамилия, имя, отчество	Подпись	Дата

Оглавление

1 Введение.....	24
2 Описание разрабатываемой системы.....	25
3 Описание архитектуры системы.....	26
4 Описание протокола взаимодействия клиента и сервера.....	27
5 Описание функций, предоставляемых сервером.....	28
6 Описание инструментов, используемых при разработке.....	29
7 Описание деталей реализации системы.....	30
8 Описание системы тестирования.....	31
9 Инструкция по настройке сервера.....	37
10 Инструкция по использованию клиента.....	38
11 Заключение.....	40
Список использованных источников.....	41
Приложение А (справочное) Диаграмма создания и обучения нейронной сети.....	42
Приложение Б (справочное) Диаграмма активности.....	43
Приложение В (справочное) Анкетирование.....	44
Приложение Г (обязательное) Диск.....	45

1 Введение

Цель работы: разработка АС, анализирующей эмоциональную окраску вводимого текста.

Область применения АС – работа над созданием искусственного интеллекта в целях обеспечения безопасности социальных платформ и предотвращения террористических актов.

2 Описание разрабатываемой системы

Разрабатываемая система представляет собой клиент-серверное приложение [1], осуществляющее анализ тональности текста, введённого пользователем. Анализ основан на использовании нейронной сети бинарной классификации [2]. Результат анализа выводится на экран в виде сообщения «Ваш отзыв является позитивным» или «Ваш отзыв является негативным».

3 Описание архитектуры системы

Архитектура системы представляет собой один сервер и множество клиентов. Клиенты отправляют запросы серверу и принимают ответы от него, сервер проводит анализ тональности сообщения и отправляет результат анализа клиентам. Обмен информацией с сервером происходит автоматически сразу после нажатия соответствующей кнопки в интерфейсе клиентской подсистемы. Благодаря тому, что серверная часть системы развёрнута с помощью хостинга [3], клиентом является браузер, и пользователем может стать любой желающий. Количество клиентов может быть неограниченно, но корректная работа гарантируется при одновременном подключении не более 5 пользователей.

4 Описание протокола взаимодействия клиента и сервера

Основополагающим протоколом сети Internet является протокол TCP/IP. Каждый информационный пакет содержит IP-адреса (IP – Internet Protocol) компьютера-отправителя и компьютера-получателя. Специальные компьютеры, называемые маршрутизаторами, используя IP-адреса, направляют информационные пакеты в нужную сторону, то есть к указанному в них получателю.

5 Описание функций, предоставляемых сервером

Функции, предоставляемые сервером:

- `home(request);`
- `about(request);`
- `document(request);`
- `predict_sentiment(fulltext);`
- `analysis(request).`

Функция `home(request)` – это функция отображения для домашней страницы сайта.

Функция `about(request)` – это функция отображения страницы сайта с информацией о команде разработчика.

Функция `document(request)` – это функция отображения страницы сайта с файлами документации для скачивания.

Функция `predict_sentiment(fulltext)` осуществляет предсказание, полученного текста. Пошаговая работа функции:

- а) загрузка архитектуры сети;
- б) создание модели;
- в) загрузка сохранённых весов в модель;
- г) задание максимальной длины отзыва;
- д) добавление в переменную `text` переменной `fulltext` в виде массива;
- е) векторизация текстового корпуса, превращение каждого текста в виде последовательности целых чисел;
- ж) генерация выходных предсказаний для входных выборок.

Функция `analysis(request)` – это главная функция анализа тональности текста, через которую происходит переход к функции предсказания и получение от нее оценки тональности текста.

Описание функций так же представлено в репозитории проекта.

6 Описание инструментов, использованных при разработке

Автоматизированная система реализована на языке Python 3.6. Сайт был создан с использованием фреймворка Django 2.0 и развернут на хостинге Pythonanywhere, который поддерживает актуальные версии всех библиотек Python. При разработке использовались библиотеки Keras 2.2.X и TensorFlow.

Для обучения нейронной сети использовалась база данных Юлии Рубцовой [4]. Работа с базой данных проводилась в Sqlite Browser.

Для развёртывания сервера использовались программы PyCharm Community Edition 2018.2.5, Python 3.6, pip 18.1 и дополнительно скачанные модули, список которых хранится в репозитории в файле «requirements.txt».

7 Описание деталей реализации системы

Для реализации данного проекта было выполнено 3 больших этапа разработки:

- написание нейронной сети с её последующим обучением и сохранением архитектуры и весов (приложение А);
- создание сайта django;
- загрузка готовой модели с сохранёнными весами и архитектурой на сайт.

Для написания нейронной сети понадобились программы PyCharm 2018.2.5, интерпретатор питона 3.6 и дополнительно скачанные модули, список которых хранится в репозитории в файле «requirements.txt».

Создание сайта django состояло из следующих пунктов:

- подгрузка библиотеки Django 2.1;
- создание макета сайта;
- рисование дизайна с использованием Photoshop;
- вёрстка html страниц.

На последнем этапе создания системы в главный файл views были добавлены функции предсказания тонального анализа, в которые входит загрузка сохранённой модели и её архитектуры.

8 Описание системы тестирования

8.1 Этап перед тестированием

При реализации данного этапа была написана программа-парсер для сайта «<http://vseotzyvy.ru>» и собрана начальная база на 16 356 отзывов. Диаграмма активности данной программы прикреплена в приложении Б.

Данная база содержала всего 377 негативных отзывов, поэтому на ее основе была сформирована сбалансированная тестовая выборка (test_base.db) на 751 отзыв, в которой 377 негативных и 374 позитивных отзывов.

Еще одна тестовая база (test_base11.db) была создана на основе данных для обучения нейросети, и состоит из 200 отрицательных и 200 положительных отзывов.

Также был реализован тест для проверки отсутствия одинаковых отзывов, результат работы теста представлен на рисунке 8.1.

```
D:\grohotova\Рабочий стол\тестирование\Site\test_base>python -m unittest test_for_testbase
..
-----
Ran 2 tests in 0.353s

OK
```

Рисунок 8.1 – Тестирование тестовых баз

Пример тестовой базы «test_base.db» представлен на рисунке 8.2.

feedback	score
Пылесос Thomas Vestfalia XT мне купил муж, когда я забеременела. Дело в том, что у нас е...	1
Пылесос очень надёжный и практичный, если деньги есть советую брать, а денег на такую т...	1
После возни со всякими пылевыми мешками и прочей гадостью пылесос с аквафильтром ст...	1
Хорошее качество изготовления, прекрасное соотношение цены и качества, очень удобно ег...	1
Действительно удивительно, как такой небольшой пылесос умудряется настолько эффектив...	1
Долго решали, что выбрать эту модель и тThomas Twin XT. Всё такие остановились на нём и...	1
Давно хотела Томас домой и вот решились с мужем на покупку. Правда сначала хотели выб...	1
Пылесос Thomas Vestfalia XT очень удобный в использовании. Он не тяжелый и манёврнн...	1

Рисунок 8.2 – Тестовая база «test_base.db»

Пример тестовой базы «test_base11.db» представлен на рисунке 8.3.

feedback	score
USER хоть я и школота но поверь у нас то же самое d общество профилирующий предмет...	1
да все таки он немного похож на него но мой мальчик все равно лучше d	1
rt USER ну ты идиотка я испугалась за тебя	1
rt USER кто то в углу сидит и погибает от голода а мы еще 2 порции взяли хотя уже и так ...	1
USER вот что значит страшилка d но блин посмотрев все части у тебя создается ощущени...	1
ну любишь или нет я не знаю кто ты бля d URL	1
rt USER ох 9 d ну это конечно же USER чтобы у нее было много друзей ведь она такая ми...	1
rt USER у тебя есть ухажер нет мои уши не кто не жрет d	1
поприветствуем моего нового читателя USER	1

Рисунок 8.3 – Тестовая база «test_base11.db»

8.2 Первый этап тестирования

При тестировании точности используются тестовые базы с предыдущего этапа. Тесты сопоставляют метки размеченной выборки (тестовой базы) с результатом работы программы и подсчитывают точность, при точности выше 60% тест считается пройденным. Данный тест проводится с помощью Continuous Integration, из сервисов был использован Travis-CI.

Результат последнего тестирования представлен на рисунке 8.4

```
=====
FAIL: test_rnn_base (test_RNN.Test_RNN)
-----
Traceback (most recent call last):
  File "/home/travis/build/courswork2018/Site/sentifilm/test_RNN.py", line 43, in test_rnn_base
    self.assertGreaterEqual(round(k / len(list) * 100), 60)
AssertionError: 51 not greater than or equal to 60
=====
FAIL: test_rnn_base11 (test_RNN.Test_RNN)
-----
Traceback (most recent call last):
  File "/home/travis/build/courswork2018/Site/sentifilm/test_RNN.py", line 25, in test_rnn_base11
    self.assertGreaterEqual(round(k / len(list) * 100), 60)
AssertionError: 52 not greater than or equal to 60
-----
Ran 2 tests in 300.096s

FAILED (errors=2)
The command "python -m unittest test_RNN" exited with 1.
```

Рисунок 8.4 – Результат тестирования точности нейросети

Также вручную была проверена валидация пустого сообщения, результат работы сайта представлен на рисунке 8.5.

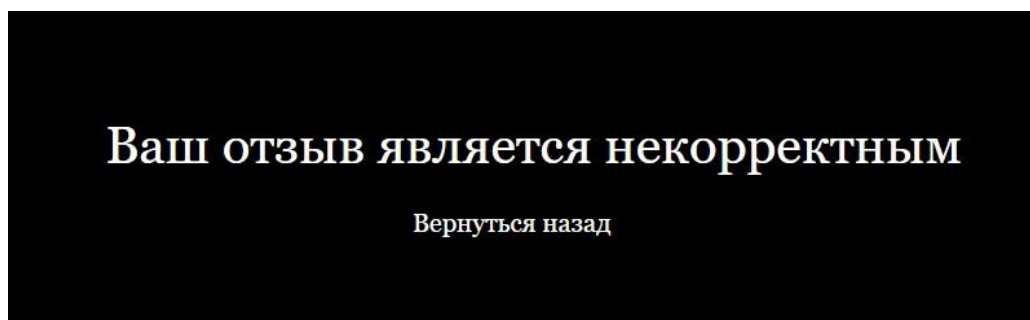


Рисунок 8.5 – Валидация пустого сообщения

При обучении нейросети была использована кросс-валидация, которая длилась 1 час 40 минут и точность которой составила 77%.

Результаты представлены на рисунке 8.6.

```

111104/111923 [=====>.] - ETA: 0s
111168/111923 [=====>.] - ETA: 0s
111232/111923 [=====>.] - ETA: 0s
111296/111923 [=====>.] - ETA: 0s
111360/111923 [=====>.] - ETA: 0s
111424/111923 [=====>.] - ETA: 0s
111488/111923 [=====>.] - ETA: 0s
111552/111923 [=====>.] - ETA: 0s
111616/111923 [=====>.] - ETA: 0s
111680/111923 [=====>.] - ETA: 0s
111744/111923 [=====>.] - ETA: 0s
111808/111923 [=====>.] - ETA: 0s
111872/111923 [=====>.] - ETA: 0s
111923/111923 [=====] - 113s 1ms/step
Accuracy: 77.11%
Saved model to disk

Process finished with exit code 0
|

```

Рисунок 8.6 – Результат кросс-валидации

8.3 Второй этап тестирования

На втором этапе было проведено тестирование удобства и привлекательности разработанного сайта. Тестирование проводилось с помощью предоставления бета-версии сайта некоторому количеству пользователей и последующему их анкетированию. Анализ результатов показал, что всем опрошенным было приятно работать с разработанным сайтом, также был предложен ряд небольших улучшений дизайна. В последствии разработчик реализовал некоторые предложенные идеи, но повторное анкетирование не проводилось.

Подписи анкетированных представлены в приложении В.

Также было проведено тестирование длительности отклика. Подсчитывалось среднее время работы нейросети при вводе отзывов из тестовых баз данных. Тест был пройден. Среднее время отклика составило менее 60 секунд.

8.4 Третий этап тестирования

На третьем этапе было проведено конфигурационное тестирование. При изменении масштаба дизайн сайта также меняет масштаб. Тест не пройден.

9 Инструкция по настройке сервера

Подробная инструкция по настройке сервера представлена в файле README.md в репозитории на GitHub (<https://github.com/courswork2018/Site>).

Алгоритм развёртывания сервера:

- а) скачать PyCharm;
- б) скачать Python 3.6.X;
- в) скачать с репозитория на GitHub все файлы;
- г) установить менеджер пакетов pip 18.1 или обновить его до этой версии;
- д) установить Django 2.1;
- е) скачать библиотеки Keras и TensorFlow;

ж) открыть проект с помощью Windows PowerShell. Чтобы это сделать, необходимо зажать Shift и кликнуть правой кнопкой мыши на папку с проектом. В открывшемся меню следует выбрать "Открыть окно PowerShell здесь". Если такой записи нет, то PowerShell можно скачать [здесь](#);

- з) в открывшейся консоли следует ввести `python manag.py runserver`.

10 Инструкция по использованию клиента

Для начала работы с автоматизированной системой пользователю необходимо перейти на сайт <https://tusurgpo726.pythonanywhere.com>. В верхней части сайта расположено меню разделов сайта (рис. 10.1).

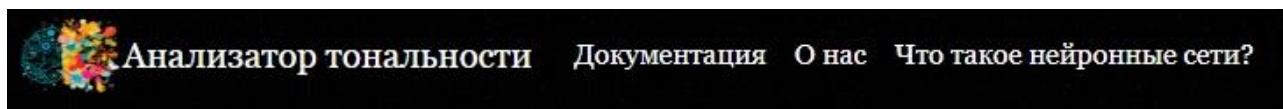


Рисунок 10.1 – Меню

В разделе «Документация» располагаются ссылки на техническое задание, пояснительную записку, инструкцию по развёртыванию сервера и инструкцию по использованию клиента.

В разделе «О нас» содержится информация о разработчике, тестировщике и документаторе АС «Анализатор тональности».

При нажатии на раздел «Что такое нейронные сети» сайт переходит на страницу Википедии с информацией о нейронных сетях.

Раздел «Анализатор тональности» является главным и содержит информацию об АС, окно для ввода данных и кнопку «Проанализировать!» (рис. 10.2).

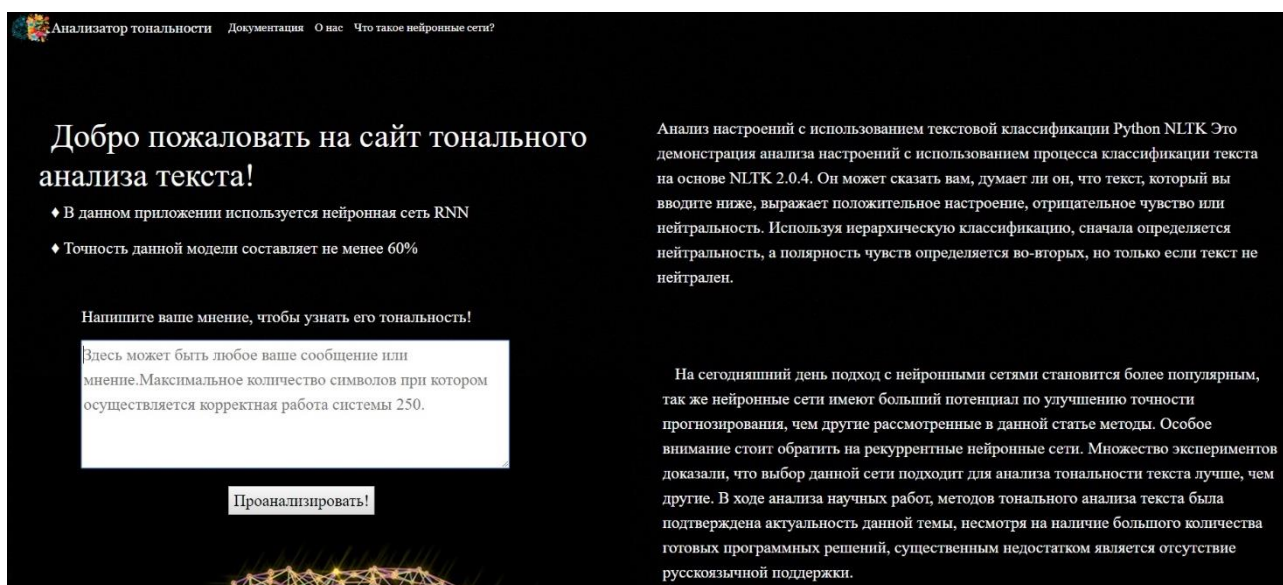


Рисунок 10.2 – Главное окно

Для анализа сообщения требуется ввести текст сообщения в окно для ввода данных и нажать кнопку «Проанализировать!» (рис. 10.3).

Напишите ваше мнение, чтобы узнать его тональность!

Сегодня хорошая погода, и у меня отличное настроение! Я надеюсь, что этот день пройдет на "ура" и закончится приятным сюрпризом!

Проанализировать!

Рисунок 10.3 – Ввод текста

Для корректной работы системы текст сообщения должен представлять собой осмысленное предложение на русском языке без орфографических ошибок размером от 200 до 300 символов.

После нажатия кнопки «Проанализировать!» Выведется одно из сообщений об оценке тональности, представленных на рисунке 10.4.

Ваш отзыв является позитивным

Вернуться назад

Ваш отзыв является негативным

Вернуться назад

Рисунок 10.4 – Оценка тональности

11 Заключение

В результате данной работы была спроектирована и реализована АС «Анализатор тональности». АС была протестирована в соответствии с тест планом. По выполненной работе была написана вся необходимая документация.

Список использованных источников

1 Архитектура «клиент-сервер» с примерами [Электронный ресурс]. – Режим доступа: <https://zametkinapolyah.ru/servera-i-protokoly/o-modeli-vzaimodejstviya-klient-server-prostymi-slovami-arxitektura-klient-server-s-primerami.html> (дата обращения: 10.10.2018)

2 Применение нейронных сетей для задач классификации [Электронный ресурс]. – Режим доступа: <https://basegroup.ru/community/articles/classification> (дата обращения: 15.10.2018)

3 Хостинг [Электронный ресурс]. – Режим доступа: <https://ru.wikipedia.org/wiki/Хостинг> (дата обращения: 3.11.2018)

4 Рубцова, Ю.В. База размеченных и неразмеченных данных [Электронный ресурс]. – Режим доступа: <http://study.mokoron.com> (дата обращения: 15.11.2018)

Приложение Б

(справочное)

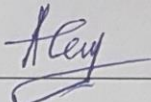
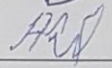
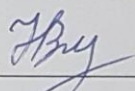
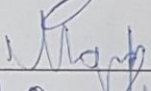
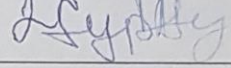
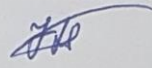
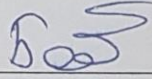
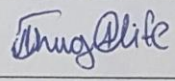
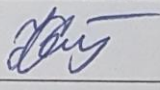

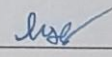
Диаграмма активности



Приложение В

(справочное)

Анкетирование

	ФИО	Подпись
1	Саченков А. А.	
2	Авдеевич К. В.	
3	Жуанчестов В. А.	
4	Морозов В. М.	
5	Дугаров Д. Б.	
6	Кузнецов Р. Ю.	
7	Бобков А. Д.	
8	Касанов З. А.	
9	Суханов Р. А.	
10	Тахтеев И. А.	
11	Лобкин В. В.	
12		

Приложение Г
(обязательное)

Диск