

Разработка алгоритма фильтрации спама gmail

Выполнил:

Левкович Дмитрий Валерьевич, гр. 7383

Руководитель:
доцент

Кирияничков Владимир Андреевич, к.т.н.,

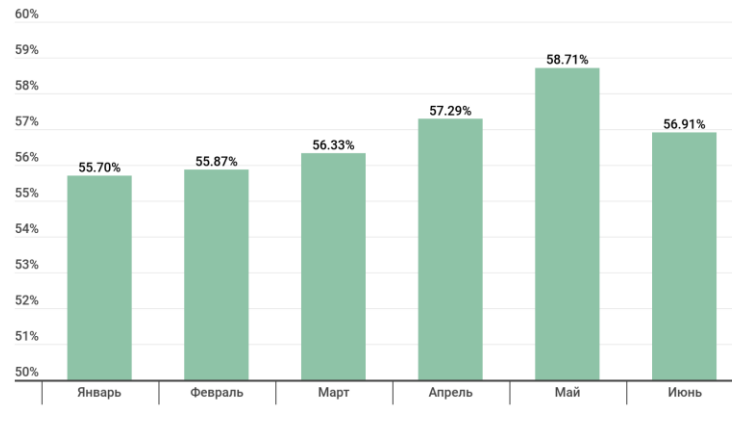
Консультант:

Шевская Наталья Владимировна

Актуальность

Актуальность:

1. Существующие методы фильтрации могут ошибочно вносить в базу спамеров IP и доменные адреса.
2. Доля почтового спама в мире согласно исследованию лаборатории Касперского высокая.



kaspersky

Рисунок 1 – Доля спама в почтовом трафике за второй квартал 2019 года

Цели и задачи

Цель: разработка алгоритма для фильтрации спама с электронной почты Gmail.

Задачи:

1. Обзор предметной области;
2. Разработать алгоритм классификации спама;
3. Исследование характеристик разработанного алгоритма;
4. Сравнение разработанного алгоритма с аналогами;
5. Интеграция алгоритма с Gmail почтой.

Обзор предметной области

Алгоритм	Средняя точность	Стандартное отклонение
Нейронная сеть	0.910	0.011
Метод опорных векторов	0.904	0.010
Наивный байесовский классификатор	0.880	0.017
Ближайших k соседей	0.826	0.018
Случайный лес	0.769	0.073

Описание исходных данных

Описание первого набора:

Описание первого набора:

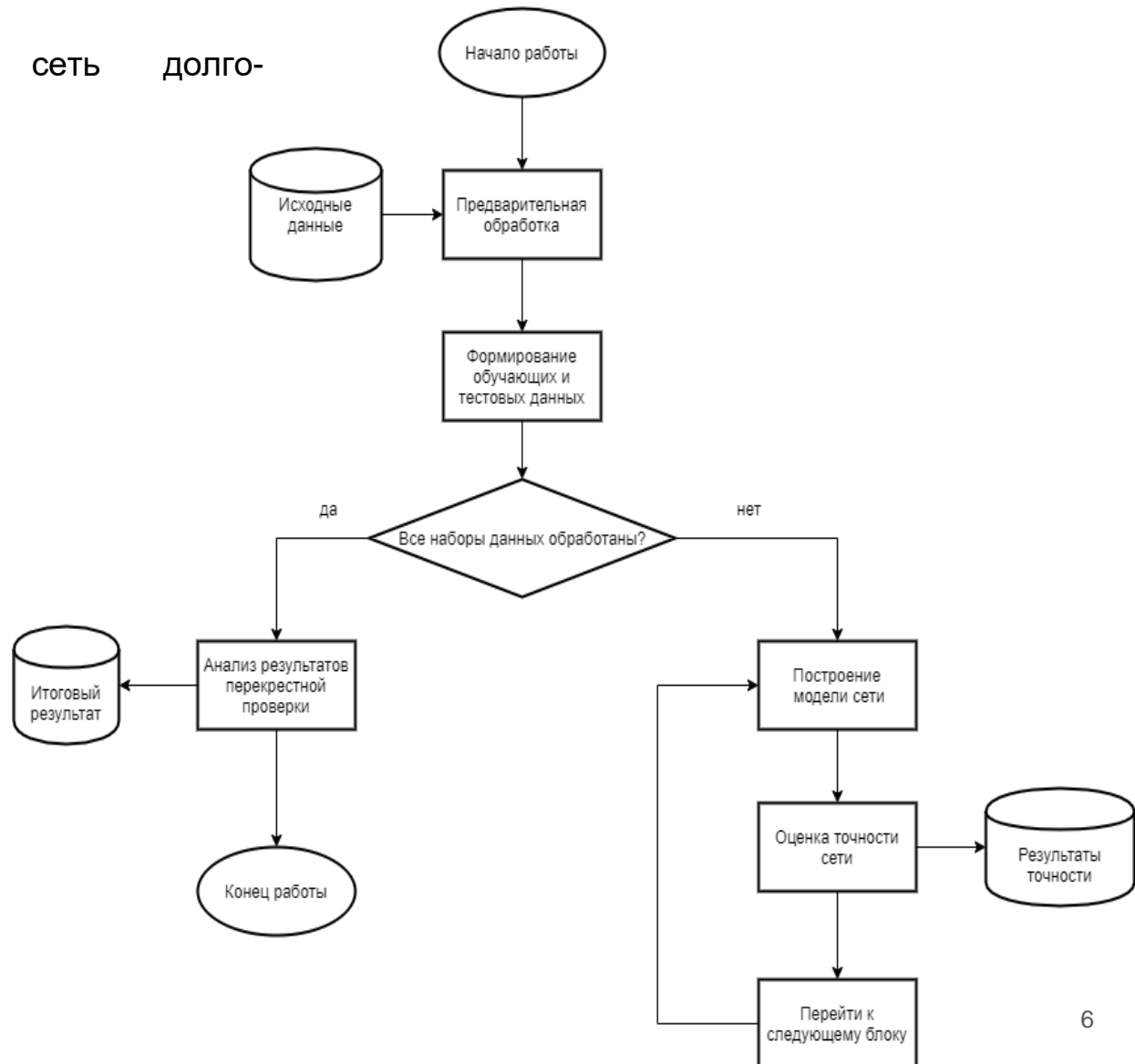
- 48 признаков – частота встречаемости слов
- 6 признаков – частота встречаемости символов
- средняя длина непрерывных последовательностей заглавных букв
- длина самой длинной непрерывной последовательности заглавных букв
- суммарное количество заглавных букв
- Спам/ не спам

Описание второго набора:

- Спам/ не спам
- Текстовое представление письма

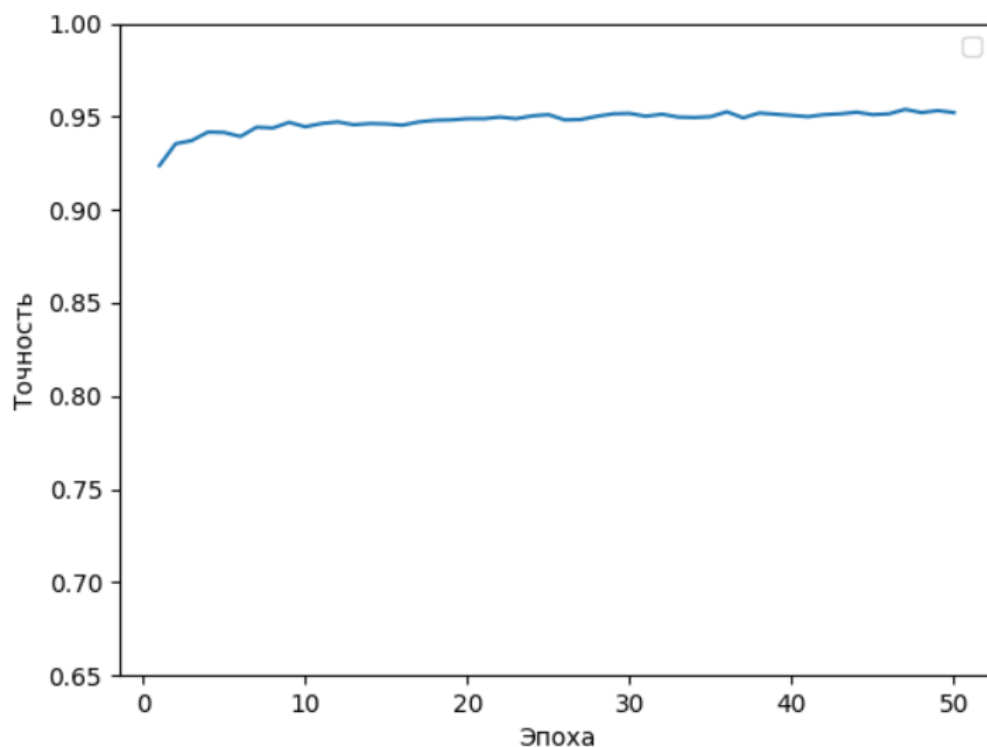
Алгоритм классификации спама

Алгоритм — нейронная сеть долго-краткосрочной памяти.



Исследование характеристик разработанного алгоритма

Метод оценки модели по заданным критериям — перекрестная проверка по 10 блокам.

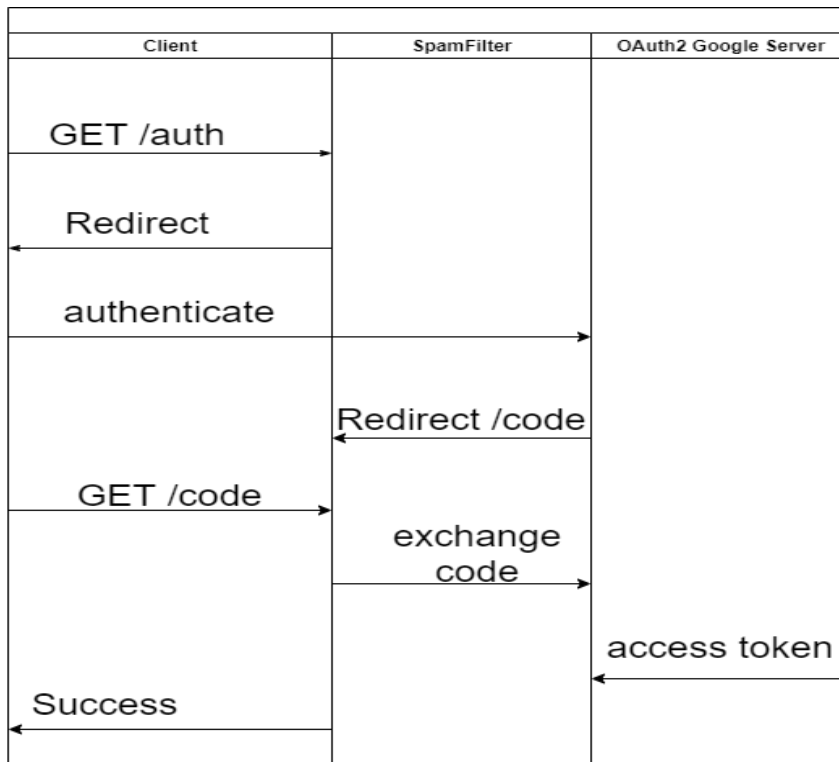


Сравнение алгоритма с аналогами

Алгоритм	Средняя точность	Стандартное отклонение, 10^{-2}
Рекуррентная нейронная сеть	0.9552	0.746
Наивный байесовский классификатор	0.8850	1.700
Метод опорных векторов	0.9040	1.000
Метод k-ближайших соседей	0.8260	1.800
Случайный лес	0.7690	7.300

Интеграция алгоритма с Gmail почтой

- Взаимодействия с почтой происходит с помощью Gmail API.
- Авторизация в приложение с помощью OAuth 2.0



Заключение

- Обзор аналогов показал, что нейронные сети лучше остальных алгоритмов справляются с задачей классификации спама;
- Была разработана нейронная сеть для классификации спама
- Экспериментальное исследование показало, что разработанный алгоритм ведет себя стабильно для обоих наборов данных, стандартное отклонение небольшое;
- Сравнение с аналогами показало, что разработанный алгоритм позволяет повысить точность классификации, а также ведет себя стабильнее остальных.
- Разработанный алгоритм был интегрирован с почтой Gmail.

Апробация работы

- Левкович Д.В. / IX Научно-практической конференции с международным участием «Наука настоящего и будущего»
- Репозиторий проекта
https://github.com/DmLvkvch/gmail_filter

Точность алгоритма на втором наборе данных

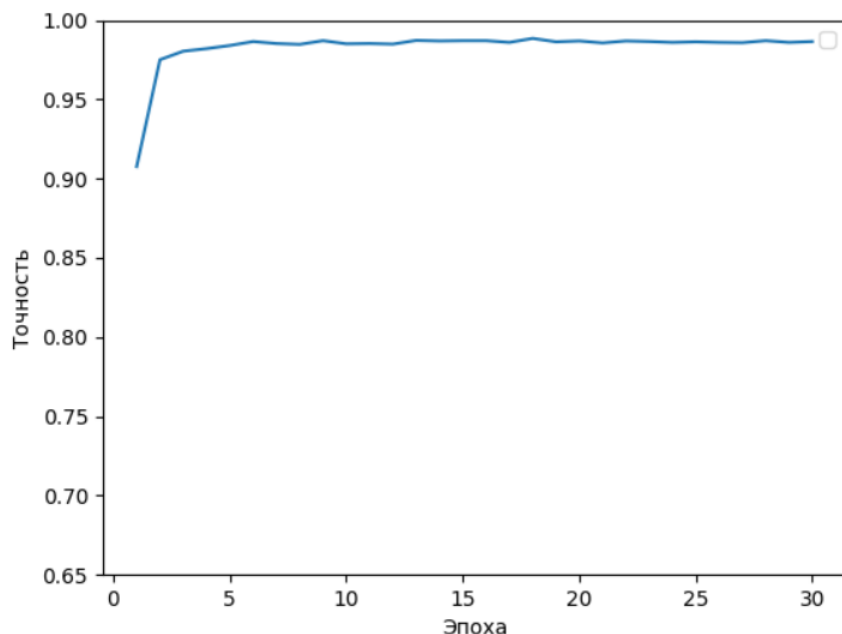


Рисунок 4 – Точность для второго набора данных, используемый для оценки качества классификации текстового вида данных

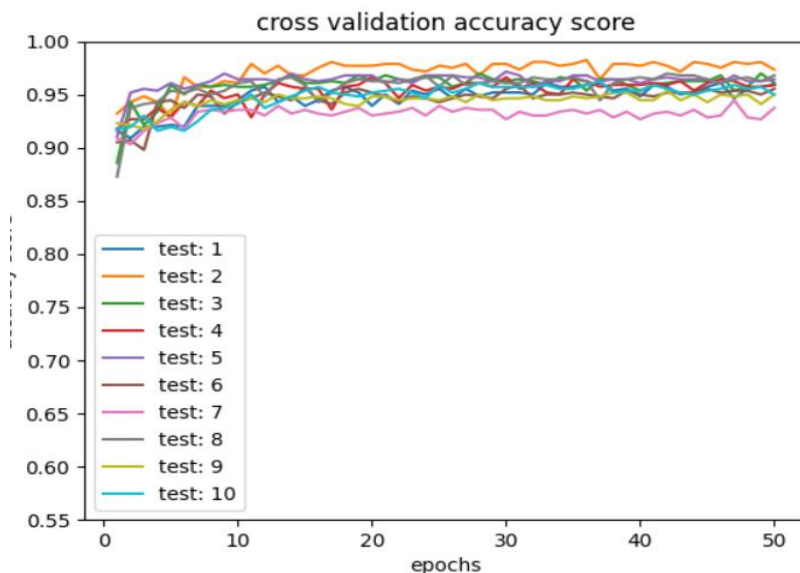


Рисунок 5 – Точность для второго набора данных, используемый для оценки качества классификации текстового вида данных

Авторизация в приложение

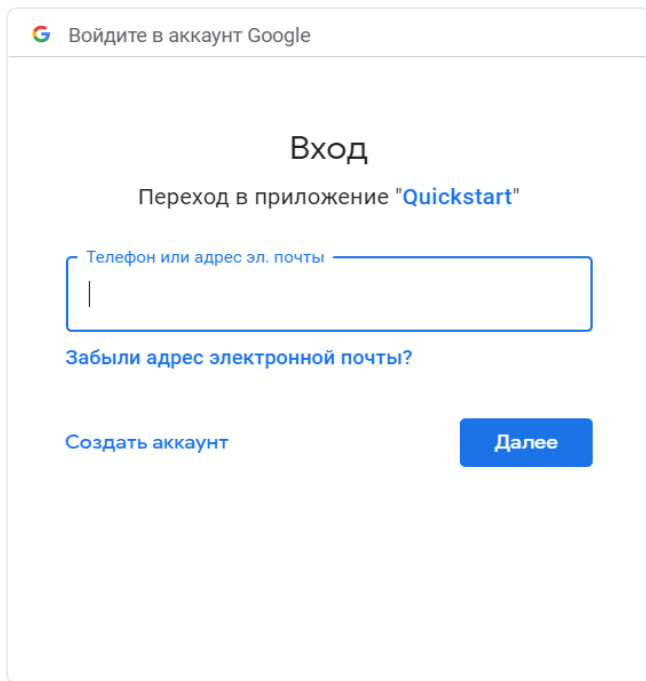


Рисунок 5 - Страница аутентификации пользователя в системе Google

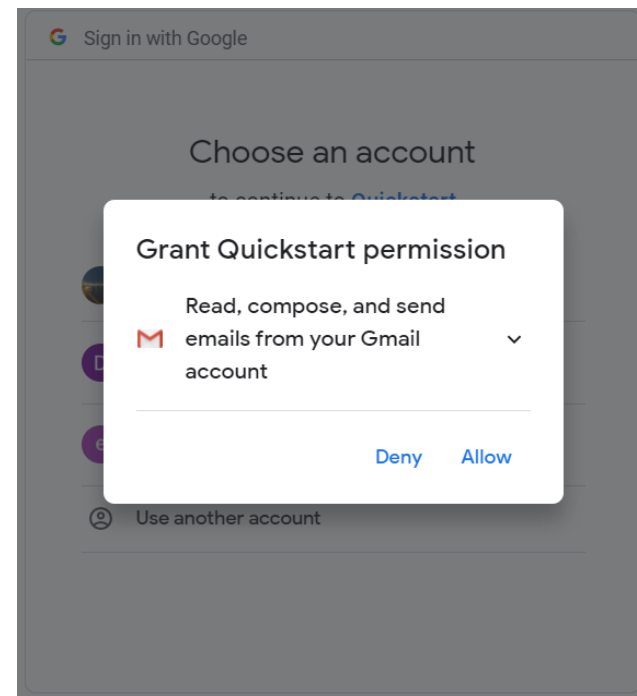


Рисунок 6 - Страница с подтверждением предоставления приложению прав доступа к ресурсам аккаунта

UML-диаграмма разработанного приложения

