

# Разработка фильтра электронных писем на основе методов машинного обучения как способ борьбы со спамом

С.А. Федорова

Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, Россия

**Обоснование.** В настоящее время электронная почта является не только эффективным способом отправки и получения данных, но и еще она стала использоваться пользователями для регистрации на различных интернет-платформах. В связи с этим количество самых разных писем на почте сильно увеличилось, и найти нужное письмо в этом огромном потоке информации становится нелегкой задачей. Кроме того, некоторые письма «превращаются» в спам, и мы никогда не узнаем, почему. Вопрос в том, как работают спам-фильтры [1]. Можно ли адаптировать современные системы под свои потребности? Эти вопросы достаточно актуальны в современном информационном мире, поэтому именно они легли в основу нашего исследования.

**Цель** — определить наилучший алгоритм и создать на его основе собственную модель.

**Методы.** В работе применяли следующие теоретические методы исследования: анализ и синтез, классификация. Кроме того, для решения поставленных задач был использован такой математический метод, как визуализация данных, а также основой для всей исследовательской работы являлся метод динамического программирования.

**Результаты.** В первой главе были рассмотрены и проанализированы следующие методы автоматической классификации [2]: наивный байесовский классификатор, логистическая регрессия и векторизация слов, на которых базируются большое количество спам-фильтров. Кроме того, было проведено сравнение метрик, которые в дальнейшем будут применены для оценки алгоритмов, используемых в этой работе. В состав метрик вошли: точность (precision), полнота (recall),  $f$ -мера (f-measure), коэффициент корреляции Мэтьюса (MCC). Также в ходе исследования были проанализированы алгоритмы, используемые в популярных компаниях, таких как Mail.ru, Gmail.com и Yandex.ru [3].

В практической главе были выполнены поиск, сбор и обработка баз спам-сообщений. Все материалы, использованные в нашем исследовании, были получены путем поиска данных на сайтах Kaggle.com, archive.org, а также с помощью поисковой системы Google. В результате мощность полученной обучающей выборки составила 14 299 сообщений.

Далее, было проведено сравнение наивного байесовского классификатора и логистической регрессии на полученной выборке с помощью такой статической модели, как  $n$ -граммы ( $n$ -граммы — это последовательность из  $n$  подряд идущих слов в тексте). В качестве оценки была взята  $f$ -мера. По результатам сравнения мы стали использовать логистическую регрессию, т. к. она справляется с поставленной задачей лучше при любом наборе  $n$ -грамм. После этого был создан спам-фильтр на основе полученного алгоритма [4].

На следующем этапе была произведена классификация ряда писем из личной почты на основе созданной модели и сравнение качества работы спам-фильтра, основанного на предложенном алгоритме, со спам-фильтром Gmail и с экспертным мнением.

После разработки, применения и оценки алгоритмов фильтрации спама мы рассмотрели еще вопрос касательно применимости построенной модели на других выборках, например на письмах на другом языке. Нами была определена область применимости построенной модели.

**Выводы.** В результате работы был создан математический алгоритм, основанный на векторизации слов и логистической регрессии. Исследование доказало, что данный спам-фильтр является эффективным инструментом, который может быть использован в различных областях.

**Ключевые слова:** спам-фильтр; метод фильтрации; методы машинного обучения; логистическая регрессия; векторизация слов.

## Список литературы

1. Скляренко Н.С. Обзор алгоритмов машинного обучения, решающих задачу обнаружения спама // Новые информационные технологии в автоматизированных системах. 2017. № 20. С. 251–257.
2. [ru.coursera.org](https://ru.coursera.org) [Электронный ресурс]. Рябенко Е., Слесарев А., Кантор В., и др. Спецкурс «Машинное обучение и анализ данных». Лекция «Прикладные задачи анализа данных» / под ред. Д.П. Ветрова, Д.А. Кропотова. [дата обращения 18.01.2023]. Доступ по: <https://ru.coursera.org/learn/data-analysis-applications>
3. [habr.com](https://habr.com/ru/company/mailru/blog/476714/) [Электронный ресурс]. Эксплуатация машинного обучения в Почте Mail.ru [дата обращения: 29.01.2022]. Доступ по: <https://habr.com/ru/company/mailru/blog/476714/>
4. Гмурман В.Е. Теория вероятностей и математическая статистика: Учебное пособие для вузов. 9-е изд., стер. Москва: Высшая школа, 2003. С. 52–53.

*Сведения об авторе:*

**Софья Александровна Федорова** — студентка, группа 6101-010302D, институт информатики и кибернетики, Самарский университет, Самара, Россия. E-mail: [feodorowa.sof@yandex.ru](mailto:feodorowa.sof@yandex.ru)