

Проблема фильтрации спама

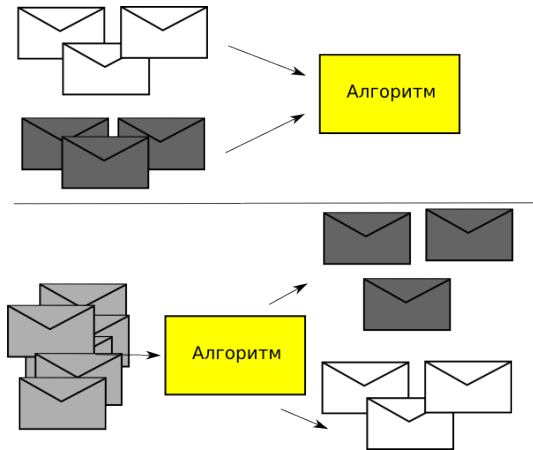
Определение

Спам - нежелательная почта. Та почта, которую пользователь не хотел бы получить даже зная о факте ее отправки.

Определение

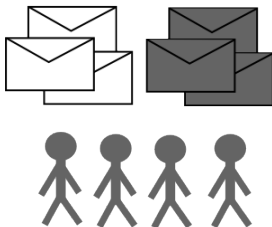
Задача фильтрации спама - задача обнаружения спам-сообщений для их последующего отсеивания из потока входящей почты.

Статистический подход к фильтрации спама



Персонифицированный и персонифицированный подходы

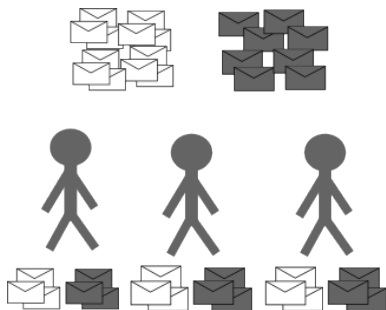
Неперсонифицированный



персонифицированный



Многопрофильный подход



Для классификации используются как собственные письма, так и письма от других пользователей

Задача классификации

Определение

Имеется множество **объектов**, разделённых некоторым образом на **классы**. Задано конечное подмножество объектов, для которых известно, к каким классам они относятся. Это множество называется **обучающей выборкой**. Классовая принадлежность остальных объектов не известна. Требуется построить алгоритм, способный классифицировать произвольный объект из исходного множества.

Определение

Классифицировать объект — значит, указать номер (или наименование класса), к которому относится данный объект.

Постановка задачи

Цели работы

- Реализовать свободную систему фильтрации спама, основанную на методе опорных векторов
- Разработать и реализовать на основе метода опорных векторов многопрофильный метод фильтрации спама.

Постановка задачи

- Произвести обзор открытых систем фильтрации спама и выбрать средство для расширения.
- На основе метода опорных векторов разработать классификатор сообщений, поддерживающий многопрофильность.
- Реализовать разработанный классификатор в рамках выбранной системы.
- Произвести экспериментальное исследование и апробацию на реальных данных.

Представление письма в виде вектора

Текст 1:	Соответствие слов позициям
<i>Привет мир</i>	
Текст 2:	
<i>Привет медведь</i>	Привет: 1
Текст 3:	мир: 2
<i>Привет Привет</i>	медведь: 3

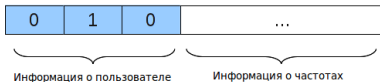
Векторное представление

Текст 1:	1	1	0
Текст 2:	1	0	1
Текст 3:	2	0	0

Представление письма в виде вектора признаков.

Многопрофильность

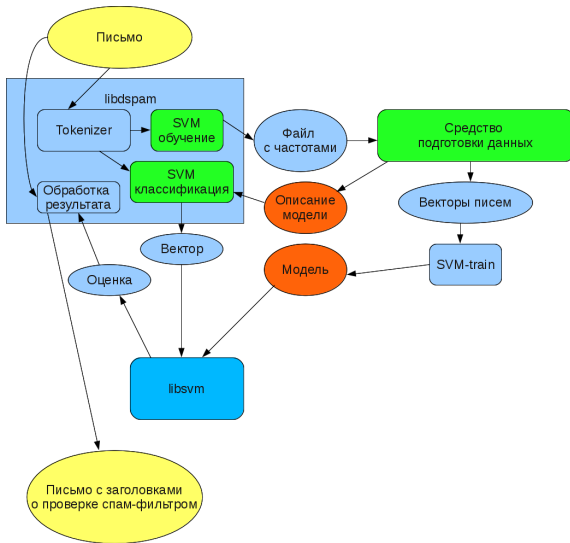
id	email
0	user1@example.com
1	user2@example.com
2	user3@example.com



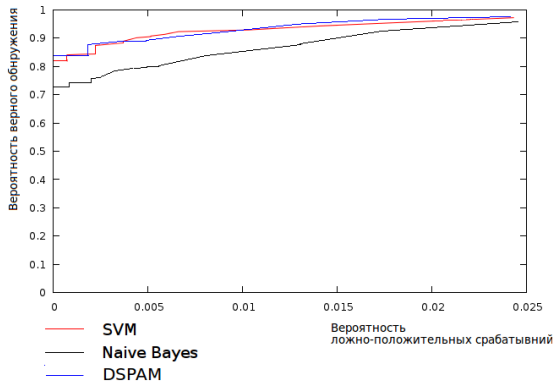
Добавление информации о пользователе

- Свободный
- Быстрый
- Многопользовательский

Схема работы модифицированного dspam



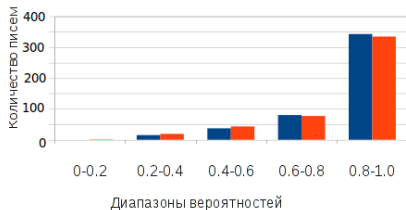
Результаты тестирования



Соотношение коэффициента верных обнаружений и
коэффициента ложных срабатываний

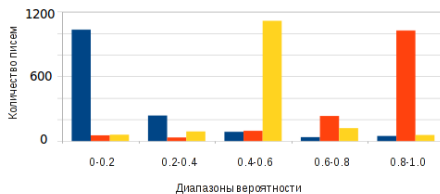
Тестирование многопрофильного режима

Распространение обучения



- Пользователь 1. Обучен как на спаме
- Пользователь 2. Не обучался на данном классе писем

Учет персональной информации



- Пользователь 1. Обучен как на легитимной почте
- Пользователь 2. Обучен как на спаме
- Пользователь 3. Не обучался на данном классе писем

Результаты

- Произведен обзор существующих средств фильтрации спама, выбрано средство для доработки.
- В рамках средства реализован алгоритм фильтрации спама на основе метода опорных векторов.
- Разработанный метод доработан для работы с несколькими профилями.
- Доработанный метод реализован в рамках системы dspam.
- Произведено экспериментальное исследование и апробация на реальных данных.