

AngryBlock

Фильтрация новостей

SPb AU project

Беляев Станислав
Степанов Владимир
Пластинин Виталий

Клиентская часть

- Расширение для Google Chrome:
 - 1) блокирует негативные новости
 - 2) собирает заголовки и посылает на сервер для обучения

Освоено:

- Основы работы с git
- Javascript + HTML
 - HTTP запросы
 - манипуляции с DOM
- Написание плагинов для Google Chrome
 - content script
 - context menu
 - local storage

Сервер

- Обучение:
 1. Пересчёт вероятностей для наивного байеса
 2. Вероятности хранятся в sqlite
- Классификация
 3. Разбиваем текст на лексемы (nltk.RegexpTokenizer)
 4. Наивный байесовский классификатор для оценки

Познакомился:

- библиотека flask, requests
- библиотека logging
- nltk Tokenizers и NaiveBayesClassifier
- git, GitHub
- Heroku, запуск удалённых приложений на сервере
- SQL: Расширения SQLAlchemy и управление SQL базой из python приложения

Проблемы

- Невозможность хранить локальные файлы на сервере heroku
- Особенности cross-domain запросов

<https://github.com/vitalik239/AngryBlock>