

Техническое задание

Формулировка задачи: На основании заявляемых провайдером характеристик качества соединения с сетью интернет и написанных пользователями отзывами, составить сравнительную таблицу и выявить лучших поставщиков услуги соединения с глобальной сетью Интернет.

Данные представляют собой отзывы пользователей интернет-услуг. Данные берутся с сайтов агрегаторов:

- <https://2ip.ru/isp-rating/?pageId=2>
- <https://ru.otzyv.com/it-internet/internet-provaydery/>
- <https://allcitynet.ru/reviews/>

Задачи:

1. Провести сбор данных с сайтов.
2. Разметить отзывы по признаку хороший/плохой для заданного провайдера .
3. Провести токенизацию, удалить стоп-слова, выделить основные слова, применить бинарной классификатор к подготовленным данным
4. Ранжировать полученные результаты по интернет-провайдерам.
5. Выявить лучших представителей услуг.

Используемые методы: нейронные сети, логистическая регрессия.

Оценка точности: Оценку точности классификации делать с помощью тестового и тренировочного набора, а также кросс-валидации. Строить confusion matrix и считать по ней разные значения, характеризующие качество созданных моделей классификации.

Новизна: Из-за быстрого развития информационных технологий, в частности сетевых, каждому человеку необходимо успевать следовать этой тенденции. Для этого необходимы услуги по предоставлению доступа к глобальной сети, предоставляемые интернет провайдерами. Выбор оптимального варианта из множества представителей поможет пользователям пользоваться услугами интернет соединения