

Senteval

У вас есть большая база отзывов о ресторанах, но вам совершенно некогда читать 100500 сообщений? Вам нужно сделать статистику по отзывам и отранжировать рестораны?



На помощь приходит Сентимент анализ (Sentiment Analysis)! Теперь можно просто проанализировать текст отзывов специальной программой, которая выдаст Вам оценку по всем аспектам:

- Насколько адекватны цены в ресторане?
- Вкусно ли там кормят?
- Приятная ли там обстановка?
- Каков уровень обслуживания?
- Ну и в целом, ресторан хороший?

Всем этим мы занимаемся в нашем проекте, и уже совсем скоро (через годик) станет доступен веб-сервис на основе нашего проекта! Ура!

А теперь серьезнее:

1) Название проекта - Senteval.

2) Программы, определяющие тональность текста, позволяют выявить отношение его автора к объекту высказывания. Такие программы регулярно используются для анализа отзывов в интернете (twitter, сайты с отзывами о ресторанах, фильмах, отелях), чтобы понять, довольны ли пользователи каким-либо продуктом.

В рамках проекта по компьютерной лингвистике SentiRuEval регулярно проводятся тестирования по обработке естественного языка в рамках автоматического sentiment analysis. В 2015 году в качестве одной из задач тестирования выступал аспектно-ориентированный анализ русскоязычных отзывов посетителей ресторанов. Участники нашего НИСа также приняли участие в этом соревновании, а теперь занимаются доработкой своего анализатора, который позволит эффективно давать оценку тональности впечатлений посетителей о ресторанах.

3) Наша задача - получить систему анализа тональности отзывов по отношению к конкретному объекту по его основным характеристикам (аспектам). В этом нам помогают словари оценочной лексики, которые мы получили из размеченного корпуса отзывов и впоследствии доработали. Также мы используем различные методы машинного обучения, чтобы добиться наиболее интересных результатов.

Аспектные категории включают в себя еду, сервис, цену, ресторан в целом, общее впечатление. Каждый из них делится еще на три типа, позволяющие сделать вывод об оценке этого аспекта ресторана посетителем. Отношение может быть позитивным, негативным и нейтральным. В итоге эти оценки суммируются и создают общую оценку ресторана.

Под чутким руководством мы извлекаем аспекты из размеченных отзывов, работаем со словарями оценочной лексики, пишем и дорабатываем программу на Python, придумываем разные возможности для оценки отзывов и осваиваем методы машинного обучения.

У нас интересно! =)

Руководитель проекта:

* Толдова Светлана Юрьевна, доцент

Участники:

* Андрей Кутузов (аспирант и сотрудник Mail.ru)

* Надя Мотина

* Дина Писаревская

* Дима Шубенок