

Лабораторная работа 4. Определение тональности текста

Под определением тональности текста подразумевается выделение фрагментов текста, выражающих позитивную или негативную эмоциональность по отношению к объекту этой эмоциональной оценки (объекту тональности). Объектом может выступать как имя собственное, так название продукта (например, наименование бренда), организации (например, Газпром), услуги, профессии и т.п., по отношению к которому анализируется текст. Таким образом, тональность текста определяется такими тремя факторами:

1. Субъект тональности – это автор текста, рецензии, комментария и т.п.

2. Тональная оценка (позитив/ нейтрально/ негатив) – это эмоциональное отношение автора к объекту тональности, выраженное в использовании специфической лексики при написании текста.

3. Объект тональности — продукт, событие, персона или явление, о которых высказывается автор.

Оценка тональности текстов предназначена для получения обратной связи между потребителями услуги и производителями услуги.

Например:

1. Оценка деятельности госучреждения в СМИ и в социальных сетях.

2. Возможность предсказания котировок (в качестве дополнительного индекса) на фондовых рынках.

3. Мониторинг экстремистских высказываний в СМИ, интернете.

Для морфологического анализа можно использовать MyStem:

<https://tech.yandex.ru/mystem/> и Sentistrength: <http://sentistrength.wlv.ac.uk/>. Инструкции по работе с данными программами размещены по приведенным выше ссылкам.

В качестве датасетов использовать корпус русскоязычных твитов:

<https://study.mokoron.com/>

Провести морфологический анализ твитов и представить результаты в виде текстового отчета со скриншотами.