
Text mining with Rapidminer - Group projects

You are an executive in a hotel/at attraction X. As competition is harsh, you would like to gain some additional insight, so you turn to Tripadvisor for reviews analysis.



Group project

Getting data:

- You found a great webcrawler tool at:
<https://github.com/urosgodnov/SimpleTripAdvisorScraping>. Use it and download data for you hotel/attraction.

Analyze the available data to gain insights into the following areas:

- To gain some basic insights we would like to provide some basic stats regarding terms and phrases in reviews and reviews overall.

After exploring data, your team has to analyze reviews deeply:

- Discover associations between terms and phrases;
- Try to identify few topics;
- Calculate sentiment and show the average sentiment score in time (by months);
- Show 3 best and 3 worst reviews;
- Extract entities and show basic stats;
- Investigate similarities in reviews;