Sentiment analysis and urban quality inference using machine learning algorithms

Description

Urban planning needs to face and integrate ecological, social, and economic aspects of city living. Previous pilot projects have investigated an automated pipeline to spatially integrate and visualize interview-based qualitative data on the perception of urban places.

The goal of this thesis project is to analyze the data from the interviews using statistical and machine learning algorithms. In particular, we are interested in the following aspects:

- Semantics analysis - what has been said about the place in the interviews?
- Sentiment analysis - polarity (positive, negative, neutral)
- Urban quality inference - identify the reasoning behind the sentiments using, e.g. feature-based sentiment analysis.

There are two main categories of urban qualities: people-based and place-based qualities. Loosely speaking, people-based quality is the cultural aspect and place-based mainly refers to the infrastructure. We are interested in both categories during this analysis.

The interview data has a semi-free form to make it expressive. However, this flexibility brings certain challenges, such as varying length, broken sentences, containing irrelevant and distracting descriptions, etc. Preprocessing techniques need to be applied when dealing with such data sets.

Research questions

- For the given data set and the goals, which state-of-the-art algorithms can be adopted and what are the pros and cons?
- Are there discrepancies in the data from different sources? How to identify and handle them?
- From the perspective of algorithm development, what is the limiting factor in the existing data set? How to improve the quality of the data?
- Using state-of-the-art techniques such as deep learning for interview data, what can be achieved in terms of semantic and sentiment analysis?
- How can we develop a system to visualize the interview data?
Prerequisites

- The applicants should have some knowledge of statistical analysis and machine learning with some experience in natural language processing. A background in urban planning is a merit. The applicants should have a strong interest in learning interdisciplinary subjects.
- The interviews are mainly in Swedish. Basic understanding of the Swedish language can be beneficial, but not required.
- One of the tasks is to visualize the interview data. A background in building visualization tools is a merit.

Contact

Yinan Yu, yinan@chalmers.se, Computer Science and Engineering
Sanjay Somanath, sanjay.somanath@chalmers.se, Architecture and Civil Engineering
Liane Thuvander, liane.thuvander@chalmers.se, Architecture and Civil Engineering

Number of projects: 1 - 2
Number of students: 1 - 2 per project

Background