

Xingming Qu

EMIS/CSE 8331: Data Mining, Spring 2019

Abstract - Intro to sentiment analysis

Sentiment analysis is ubiquitous, and it is a classification technique based on natural language processing. The main problem to be solved is to judge whether a sentence (or an article) is positive or negative. For example, on Amazon or Twitter, people post comments about a product, event, or person. Merchants can review these comments to understand the user experience and evaluation of their products. When large-scale sentiment analysis is required, the processing power of the naked eye becomes very limited. Thus, developing good sentiment analysis tools is very important. The essence of sentiment analysis is to speculate whether the text is positive or negative based on known text and emotional symbols. Dealing with sentiment analysis can greatly improve people's understanding of things, and can also use the conclusions of sentiment analysis to serve other people or things. For example, government can understand citizens' sentiment towards popular events. Therefore, they are able to grasp the public opinion orientation, and conduct public opinion monitoring in a timely and effective manner. At the same time, they may also support the formulation of relevant policies.

In this tutorial, we will talk about:

1. Text data preprocessing: One-hot encoding of words and Word Embedding.
2. Semantic analysis based on dictionary (old way).
3. Semantic analysis based on deep learning: IMDB movie-review sentiment-prediction task