






Programming Assignment

Opinion Mining and Sentiment Analysis

Consider the reviews of the Yelp Dataset Challenge for the city that you have been assigned. Build a Sentiment Analyzer that classifies a tweet as positive, negative or neutral. A review is positive, negative or neutral based upon the following definitions:

Sentiment	Star Rating (R)
Negative	$R \leq 2$
Neutral	$2.5 \leq R \leq 3.5$
Positive	$4 \leq R$

For your Sentiment Analyzer document the steps that you have performed mainly:

1. Pre-processing 
2. Features used for the classifier 
3. Classifier selected 
4. Evaluation Criteria selected 
5. Evaluation Strategy 
6. Conclusions

In this programming assignment, we do not consider a stream of reviews but consider that all the reviews are available. Suggest improvements to your Sentiment Analyzers so that it can be extended to handle streams. What are the difficulties that you might face while building the Sentiment Analyzer for streams?