Brands twitter customer service solution recommendation – MVP Proposal

Songlin Qing

Description: Based on twitter conversations between customers and customer service agents from different brands, build a solution recommendation system for customer tickets.

Data source: https://www.kaggle.com/thoughtvector/customer-support-on-twitter

Data description: 2.8 Million tweets with the following information:

- Tweet id
- Author id
- Timestamp
- Text
- Respond_to_tweet_id

Challenges:

- Customers usually do not provide all information in the initial tweet, most of the tweets are back and forth, asking questions and making clarifications. It is not clear which part is problem, which part is solution;
- Cases are not tagged explicitly as resolved or not resolved;
- No access to customer service knowledge base. We can only try to recover/build knowledge base from customer service agent responses in tweets;
- A single customer may initiate multiple cases on different days. Scope of a single case is unclear.

Approach:

1. Filter out non-English tweets and tweets without a response from customer service;
2. Group the tweets into conversations, by customer id and timestamp;
3. Extract the last tweet from customer, perform sentiment analysis and predict if the issue is resolved, not resolved, or work in progress;
4. Concatenate all tweets in a conversation. Use all tweets combined from customer as problem description, and all tweets combined from customer service as solution:
    a. Preprocess problem and solution into bag of words;
    b. Extract topics from problem and solution using TF-IDF and LSA;
    c. Perform clustering to categorize problem and solutions into clusters;
    d. For the problems that are resolved, label problem and solution pair as 1; For the problems that are not resolved, label problem and solution pair as 0;
    e. Build a binary classification to score if the problem solution pair is a match or not. Based on prediction scores, pick top 3 solutions for each problem.