Real-Time Summarization Track

Summary

The track began at TREC 2016 and is a combination of the **Microblog** track(2010 to 2015), and the **Temporal Summarization** track(2013 to 2015). RTS explores and focuses on information needs and systems that address prospective information needs against continuous document streams, exemplified by social media services such as Twitter. Users who have a number of "interest profiles" representing prospective information needs are considered. The task is to monitor the stream of documents to keep the user up to date on topics of interest. Updates are disseminated using two methods-

1. Push notifications can be sent to the user's phone as soon as a relevant post is identified. The notifications should be relevant, timely and novel.
2. Daily email digests summarizing activities in the day with respect to the interest profiles. Results should be relevant and novel though timeliness is not particularly important, provided that the tweets were all posted on the previous day.

Task(2018)

Participating systems listened to the Twitter sample stream using the Twitter streaming API. The API offers an approximately 1% sample of all tweets.

Data

Twitter stream.

Evaluation of Runs

A common pool will be constructed based on scenario A and scenario B submissions. Scenario A represents results submission via the REST API and scenario B represents results submission via Batch Upload to NIST. The evaluation methodology is based on pooling. A common pool is constructed based on scenario A and scenario B submissions.

1. Scenario A: All metrics are computed for each day for each interest profile and averaged. The metrics are Expected Gain, Normalized Cumulative Gain, Gain Minus Pain and Latency.
2. Scenario B: nDCG@10 score is computed for each day for each interest profile, and then averaged across them.

Approach to the Task

Text matching and NLP's Named Entity Recognition(NER) methods can be used for the task.