Group 11 Reviews

Gathering data from social network can be potentially useful and valuable. The group 11 states clearly in the problem motivation and points out four possible applications in different aspects. Such as Natural disasters, political uprising, crowdsourcing, and Geo-targeted advertising. Geotagged data from social media can also be implemented into local news. Another potential application using Geotagged data can be social intercourse, to make friend with people around you.

Group 11 want use Twitter data with spatial component to show the results about U.S presidential election. They fetch the tweets contains the key words like pro-Obama, pro-Romney, anti-Obama, anti-Romney, or neutral.

The hardness comes from both computational and spatial aspects. Tweets are high dimensional data in terms of word counts and also it is difficult to visualize tweet data in a map.

Group 11’s solutions are fetch useful tweet data from important milestone in the election, which avoid a lot of meaningless data and redundant data. Classify the data according to the support, and then aggregate the data by county. They also make a validation about their result according to the ground truth data from the election. I think fetch the data only from important milestone in the election is a good strategy to get ride of meaningless data and redundant data. The reason why it’s good is because the data from important milestone are most representative. One suggestion is to take use of the timestamp of the tweet data and aggregate the results by county and time. When visualize the results, we can also know the result according to the time. Another suggestion is about the acquire tweet data. I think only fetch tweet data from important milestone in election will lose many votes. If sampling tweet data can be combined with their current method, the accuracy would be improved in certain degree. The last suggestion could be the visualization. Group 11 shows very good results about the election. If they can implement intensity map on their results would be better to give people an idea about the number of votes.

Related work are presented in how they obtaining data, storing data, classifying tweets, and aggregating results. This is a common way to implement this kind of application. However, I think they miss the visualization part. After aggregating, the visualization is also an important part of the project. Group 11 did not talk about how they visualize their results. Although the visualize results in their presentation looks very good.

The ground truth data from the election are being used to validate their results. The first result tweet winner by state is reasonable because the way their acquire data. The second result tweet winner by county looks very close to the validation data.

Group 11 did a great job on their presentation. They clearly state their problem and provide a reasonable solution to solve their problem. The content of the presentation is clear and well organized. Their project has great potential value in many areas. The speakers were well engaged the audience and answered the short questions raised by audiences. I understand most of their talk and would like to learn more about their project. I would rate the presentation check+.