

1.125 Arch & Engineering Software Systems Term Project Assignment

Deliverable #1

Due Date:

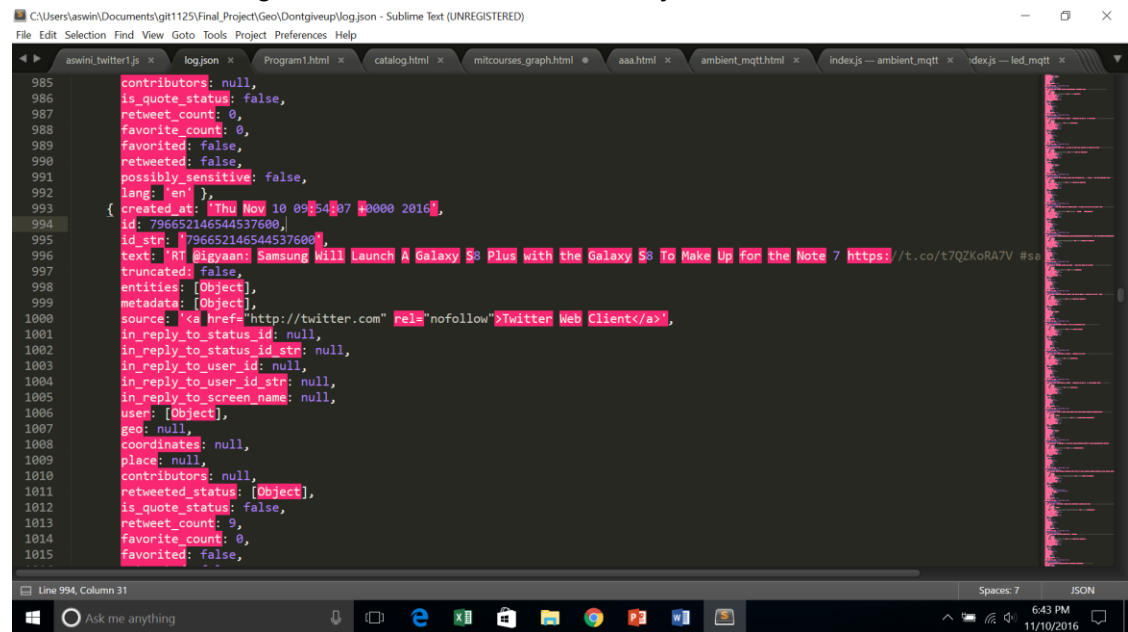
Thursday, Nov 10th, 2016

Team:

Aswini Narayana Prasad and Shweta Jindal

Technology Demonstrations

- In general these reviews are available in different platform but these data are large in volume and often contradictory and subjective too. Hence respective API packages are used to extract data.
 - API created and data extracted successfully using npm package 'Tweet'. All tweets related to 'Samsung' was extracted and stored as json file.



```

985 contributors: null,
986 is_quote_status: false,
987 retweet_count: 0,
988 favorite_count: 0,
989 favorited: false,
990 retweeted: false,
991 possibly_sensitive: false,
992 lang: "en",
993 { created_at: "Thu Nov 10 09:54:07 +0000 2016",
994   id: 796652146544537600,
995   id_str: "796652146544537600",
996   text: "RT @sanyaan: Samsung Will Launch A Galaxy S6 Plus with the Galaxy S6 To Make Up for the Note 7 https://t.co/t7QZKoRA7V #sa",
997   truncated: false,
998   entities: [Object],
999   metadata: [Object],
1000   source: "<a href='\"http://twitter.com\"' rel='\"nofollow\"'>Twitter Web Client</a>",
1001   in_reply_to_status_id: null,
1002   in_reply_to_status_id_str: null,
1003   in_reply_to_user_id: null,
1004   in_reply_to_user_id_str: null,
1005   in_reply_to_screen_name: null,
1006   user: [Object],
1007   geo: null,
1008   coordinates: null,
1009   place: null,
1010   contributors: null,
1011   retweeted_status: [Object],
1012   is_quote_status: false,
1013   retweet_count: 9,
1014   favorite_count: 0,
1015   favorited: false,
  
```

Later this dataset is cleaned and relevant reviews are extracted using Javascript. These codes are added to the filename 'Codeartifacts.pdf'.

- Machine learning is used to make meaningful conclusion from these review data. Accordingly, based on the data collected from the social media, sentiment analysis is done



Project

1.125 Final Term
Shweta Jindal, Aswini Prasad

on these reviews based on the semantic structures of these sentences and then tagged as positive/negative/neutral using classification methods followed by linear regression or similar techniques. Current plan is to use the statistical packages in R

- Data visualization using Javascript is done to show the overall trend.