Crystal Advanced Analytics
(Future-Oriented Benchmarking through Social Media Analysis)

Team Members: Claudino Alves, alvesc2014@my.fit.edu
Harshil Haumeer, hhaumeer2014@my.fit.edu

Faculty Sponsor: Dr Philip Chan, pkc@fit.edu

Client: Dr. Richard Griffith, griffith@fit.edu - *Department of Psychology*

Sponsor meetings: November 13th, 2017
November 27th, 2017

Client meetings: November 1st, 2017
November 8th, 2017

Task	Completion	Claudino	Harshil	To do
Implement, test and demo sentiment measurement of reviews (Sentiment analysis)	100%	0%	100%	None
Implement, test and demo visualization of sentiment in different categories of hierarchy	100%	0%	100%	None
Set up Website	100%	0%	100%	None
Implement web-scrapper on website	100%	0%	100%	None
Implement text classifier on website	100%	0%	100%	None
Implement, test and demo benchmarking analysis tool	40%	0%	40%	Research on benchmarking tool, and implementation

Discussion of each accomplished task for current milestone:

Task 1: A sentiment analysis analyzer was successfully implemented. The sentiment analysis tool analyzes the data(reviews) from the different categories and generates a sentiment score for each of the individual categories. The sentiment scores vary from -1 to +1, where -1 indicates the most negative score and +1 indicated the most positive score. The sentiment analysis uses a python library, textblob.

Task 2: A graphical illustration of the sentiment scores of the different categories was successfully implemented. The graph displays the average score for every category as well as an overall sentiment score of the organization combining the different categories. An attempt to the graph was made in Milestone 2, but improvements have been made which provides a better visualization of the data, such as the bar colors, labelling, accuracy and displaying of the data.

Task 3: A Website was successfully setup which runs locally for now. The website allows a user to sign up with a valid email, username, organization name and an appropriate password. The user is then able to log in to the system using the credentials and the user is automatically redirected to the home page of the website, where the user has the option to scrape (get reviews) data from different social media platforms, namely Facebook, Google and Yelp. The website was set up in 2 stages, the front-end which was done on Angular 2, Html, CSS and JavaScript. The back-end of the system was implemented in PyCharm, using Django and Django REST Framework. The back-end was then connected to the front-end through API endpoints.

Task 4: The web-scrapper was successfully implemented on the project website which returns the appropriate data (reviews) on the home page. The user has the option to view reviews from Facebook, Google and Yelp. The data is taken from the array and represented in a table.

Task 5: The text classifier was successfully implemented on the website. The classifier assigns the appropriate category for each review. The category assigned for the reviews is displayed in the respective column for each review.

Task 6: Benchmarking analysis tool in progress. Research was made on how to benchmark sentiment score and performance of one services of a particular organization to that of a competitor.

Discussion of Contribution for Each Task:

Claudino: None

Harshil: Harshil completed Task 1 through Task 4 mentioned above. He created the sentiment analysis tool using the TextBlob library in python. He implemented the sentiment analysis tool to use the data that are classified into the different categories to generate the sentiment scores of the individual categories. Moreover, he used the sentiment analyzer to generate an overall sentiment score for the organization including the different categories.

Harshil also implemented a visualization tool which displays the sentiment scores for each individual category as well as displaying the overall sentiment scores calculated as described above.

Moreover, Harshil started implementing a website for the system. He set up the front-end of the project using Angular 2, HTML and CSS. The proper fonts, colors and layout of the different pages were designed and implemented. The Website allows a user to sign up, log in and access the features of the website, which is at this point, scrapping from different social media platforms by specifying a time frame. Harshil then, implement the back-end of the system which comprises of the scrapper, classifier, sentiment analyzer and graphical visualization in the PyCharm IDE using Django and Django-REST Framework. He then connected the front-end and back-end using API end points. He then used Django Corse to make cross-domain requests.

Plan for the Next Milestone:

Task	Claudino	Harshil
Implement, test and demo benchmarking tool	60%	40%
Improve web interface	0%	100%
Implement sentiment analysis on website and implement drill-down feature	50%	50%

Discussion of each planned task for the next Milestone

Task 1: We plan on implementing a benchmarking tool that will be used to compare MCO's performance in the different services/factors to respective competitors. The comparison in performance will be displayed in graphical form.

Task 2: The interface on the website will be improved to allow users to view data/reviews from more social media platforms as well as allowing the users to get reviews within specific date range. The different categories labels will be represented by different color fonts for an easier overview of the data.

Task 3: The sentiment score analysis will be implemented on the website which will be used to generate scores for the different categories. The scores will be represented on interactive graphs where the user will be able to use the drill down option to see more specific data such as original reviews for the different categories.

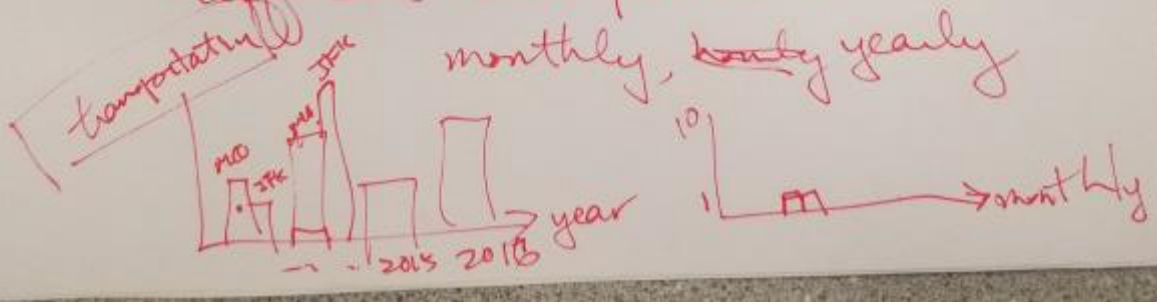
Sponsor Feedback:

1. takes too long, more than 10-15 minutes to generate web pages with the reviews
   - store reviews
   - Pre-compute web pages
   - update reviews every hour(?)

2. you're using API, not scraping

3. Text classification
   80/20 ~~fore~~ train/test
   → 7 files, 7 models, 7 predictions for each

4. Sentiment
   different time ~~periods~~ intervals
   monthly, ~~yearly~~ yearly



Sponsor Evaluation

- Sponsor: detach and return this page to Dr. Chan (HC 322)
- Score (0-10) for each member: circle a score (or circle two adjacent scores for .25 or write down a real number between 0 and 10)

Claudino	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10
Harshil	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10

Sponsor Signature: _____ Date: _____