Crystal Advanced Analytics

(Future-Oriented Benchmarking through Social Media Analysis)

Progress Evaluation for Milestone 5

Team Members: 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: February 12th, 2018
February 26th, 2018

Client meetings: February 14th, 2018
March 2nd, 2018
March 16th, 2018

Task	Completion	Harshil	To do
Complete implementation, test and demo of Benchmarking tool	100%	100%	None
Implement, Test and Demo Root-Cause Analysis	100%	100%	None
Implement graphs on Dashboard	80%	80%	Implement sentiment analysis graph and visualization for root cause analysis
Improve website	90%	90%	Implement more graphs and all necessary remaining components.
Showcase Poster	100%	100%	None

Discussion of each accomplished task for current milestone:

Task 1: A Benchmarking tool was successfully implemented and tested. It allows the benchmarking of MCO to 3 different airports as of now, but the system allows benchmarking of 5 different airports at once. Using API endpoints, the data from selected competitors are imported to the system. Then data returned is a json object which is then parsed into an array for classification. This array is stored in the database. After classifying the reviews, the data runs through the sentiment analyzer and each categorized review are assigned a polarity representing the sentiment score. The sentiment score is then aggregated and averaged for each category which is then used to benchmark against the respective categories. The benchmarking results is represented through bar charts.

Task 2: A root-cause analysis tool was successfully implemented and tested. The root-cause analysis tool takes categorized data with assigned sentiment score and retrieves the reviews with the most negative sentiment score or most positive sentiment score depending if the user wants to analyze the root cause for an underperforming service or the root cause for an over performing service. The root cause analysis tool tokenizes these reviews. It then filters those tokenized reviews to remove stop words. Stop words are repeated words of high frequency that are meaningless for the analysis such as “they”, “a”, “airport”, “#” etc. The filtered reviews are then iterated over to retrieve the most frequently occurring strings/words. Those strings/words are then parsed through an array which returns a sorted list of most frequent words with their respective frequency.

The results are represented on bar graphs with the x-axis representing most frequent words and the y-axis representing their occurrences. On the dashboard, besides displaying the bar graphs, there will be a results section which will take those most frequently used words and add a string to it to display a full sentence. Such as, for instance ‘TSA’ is a high frequently used word, the string ‘is negatively impacting your Speed service’ to display ‘TSA is negatively impacting your Speed service’ for an underperforming service or String ‘Friendly Staff’ will be concatenated to ‘is boosting an over performance in your Customer Centric Service’ displaying ‘Friendly Staff is boosting an over performance in your Customer Centric Service’.

Root-cause analysis will be performed in services over performing or underperforming, where the parameters for over performance or under performance will be set by the user available on the GUI.

Task 3: Different kinds of graphs were successfully implemented to represent different kinds of analysis and results. For Benchmarking analysis, bar graphs are being used to represent the data. There are 2 different benchmarking graphs, one to represent benchmarking based on customers ratings (1-5) and one to represent benchmarking based on sentiment score (-1 to 1). The bar graphs will allow benchmarking for up to 4 different competitors. The user will have the options to select which competitors they would like to compare their service to.

For Sentiment Analysis, gauges are being used where the gauge's gradient goes from Red to Orange/Yellow to Green. A negative sentiment will fall in the red zone of the gauge, while a neutral sentiment score will fall in the orange/yellow zone of the gauge and a positive sentiment score will fall in the green zone of the gauge.

To represent the client's service overview, circle graphs are being used. The service overview graph represents the overall service combining the 7 services based on customers' ratings (1-5). To represent service overview based on sentiment score, bar graphs with x-axis representing services and y-axis representing sentiment score from -1 to 1 are being used.

Root-cause analysis are being represented using bar graphs as well, where the x-axis represents the most frequent words and the y-axis represents the respective frequencies.

Task 4: The overall website was improved, where the main landing page includes more information about the project/services that Crystal Advanced Analytics provide. The main dashboard was also improved which includes sections for Service overview, Benchmarking, Root-Cause analysis. Different components were added for graphs implementations and results section. There are still more improvements left to be done, such as options to export the results locally as a pdf or forward it as an email attachment etc.

Task 5: The showcase poster was successfully designed and submitted to the HSDC. The poster includes an Abstract which provides an overview of the project, Features which outlines the features of the project, a labelled system architecture, Approach which provides an overview of how the system was implemented and how it works, the tools used to implement the project, a results section provided snapshots of classified data, graphs representing results from benchmarking and sentiment analysis, as well as an acknowledgement section.

Plan for the Next Milestone:

Task	Harshil
1. Complete and test website	100%
2. Complete E-book page	100%
3. Create Demo Video	100%
4. Create User Manual	100%

Discussion of each planned task for the next Milestone

Task 1: I plan on completing the website, by connecting the front-end to the back-end to represent real-time data. I will implement all the remaining necessary graphs, such as root-cause analysis and sentiment analysis. I also plan on improving existing graphs by allowing toggle options to change the range of date for which data is being displayed. I will fully test all the components of the website, including registration, log in, the different interactive graphs of the dashboards among others.

Task 2: I also plan on completing the E-book page which will include the necessary information about the project such as abstract and relevant graphics.

Task 3: I plan on creating a demo video of the project which will be used to demonstrate the tools and features that Crystal Advanced Analytics offer. The demo video will have sample real-time data of how the user will select the social media platforms reviews would be scrapped from, and the results of classified data, sentiment analysis, root cause analysis and benchmarking. The demo video will be made live on a YouTube channel.

Task 4: I plan on creating a user manual which will provide an overview of the project, the intended users and the different features of the projects. It will have the system/browser requirements to use the system. The user manual will as well include relevant graphics and all other relevant information for the user.

Sponsor feedback on each task for current Milestone (M5):

Task 1:

Task 2:

Task 3:

Task 4:

Task 5:

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)

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: _____