Crystal Advanced Analytics

(Future-Oriented Benchmarking through Social Media Analysis)

Team Member: Harshil Haumeer, hhaumeer2014@my.fit.edu

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

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

Client meetings: December 20th, 2017

Goals and motivation: The scope of this project is to create a system that will automatically retrieve data from social media platforms and analyze it to filter for customers feedback. The result will be further analyzed to identify the customers emotion and root cause of dissatisfaction, so airports can improve the areas where they are under-performing. The system will also allow businesses to compare their services to their competitors and predict future trends to identify any possible issues before they become a problem. These functions and tools will be available to businesses under one dynamic website.

Approach: The website will have an integrated web scrapping tool which gets data from the reporting platform; BrightLocal, which scrapes data from different social media platforms.

The data gathered from the web scrapping will then be further analyzed to determine the content type that will be assigned to different categories, namely the 7 World Class Factors. This will be done using a text classification algorithm.

After the text classification, the data from the different classes will then be sorted based on their ratings (1-5). For reviews without a rating, the sentiment analysis tool will provide a sentiment score.

The system will also allow a particular organization to compare its performance with its competitors. This will help organizations understand where they are underperforming and hence they can work on their weaknesses to provide a better service. Moreover, organizations will be able to identify trends and issues before they are an actual problem and take proactive actions to prevent them from occurring. Thus, working towards a future-oriented objective. This will be done using benchmarking analysis.

Novel feature: The novel feature of the system is the ability to provide businesses with the automated feature of analyzing customers' feedback through social media and understand the root cause of underperformance using text classification, sentiment analysis and root-cause analysis. On top of these, the system will auto-generate future-oriented trends where they could proactively work on identified issues before they become a problem. All these features will be made available to users under one website without having to rely on multiple systems.

Technical Challenges: I do not have prior experience to machine learning. Understand the core of it and learning how to use and implement the different algorithms I will be using along with machine learning will be challenging.

Despite having some knowledge of Python, using this language on a larger scale to implement a whole system using text classification, content analysis, and sentiment analysis among other tools on a framework that we never used before will be very challenging.

Implementing the system to provide the desired functions while making sure the system provides fast and efficient result will be another big challenge. We will need to pick the proper tools, algorithms, database/repository setup and SQL to ensure a fast performing system.

Design: System Architecture Diagram
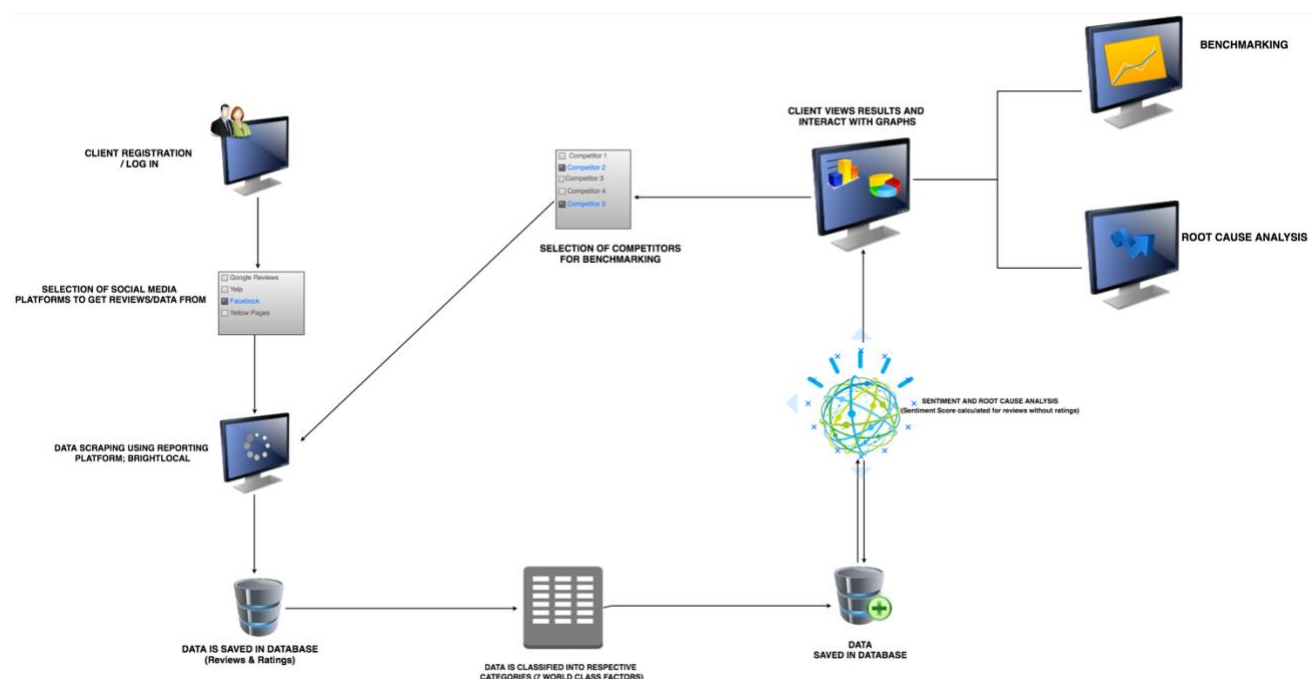
Progress Summary

Task	Completion	To do
Implement web scrapper (Use of BrightLocal)	100%	None
Implement Text Classifier	80%	Train with more data. Accuracy at 57% (initially 25%)
Implement sentiment score analyzer	100%	None
Setup up website	60%	Improve Graphs, Add more features
Implement Benchmarking tool	20%	Add option to select from multiple competitors
Implement Root Cause Analysis	0%	Identify repeating words/key words in negative reviews.

Milestones

Milestone 4 (February 12): Itemized Task

- Implement, test and Demo Benchmarking tool
- Add option to select from multiple competitors
- Improve website GUI to include Benchmarking tool
- Improve text classification accuracy

Milestone 5 (March 19): Itemized Task

- Implement, test and demo root cause analysis
- Improve GUI to include root cause analysis tool.

Milestone 6 (April 16): Itemized Task

- Add specific features such as review original reviews sorted from sentiment score polarity (e.g More negative to more positive)
- Complete Website
- Test/demo entire system
- Create User Manual
- Create demo video

Task Matrix for Milestone 4

Task	Harshil
Implement Benchmarking tool	Research, test and demo Benchmarking tool
Add option to select multiple competitors to benchmark	Add checkbox option to select competitors
Improve website	Add Dashboard, Improve GUI

Description of each planned task for Milestone 4:

I have researched methods to do an accurate benchmarking. In milestone 4, I plan on implementing a fully functional benchmarking tool which will allow users to select multiple competitors to compare their services to.

I also plan on improving the website in general, by improving graphs to be interactive and adding relevant features such as Benchmarking tool.

Approval from Faculty Sponsor

"I have discussed with the team and approve this project plan. I will evaluate the progress and assign a grade for each of the three milestones."

Signature: _____ Date: _____