

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*

Client meetings: August 30th 2017
September 6th 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 businesses 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 will be used to retrieve 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 of a predefined hierarchy, 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 analyzed and determined whether they are positive or negative. This will be done using sentiment analysis.

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 our 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 latent semantic analysis and sentiment analysis. On top of these, our 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: We do not have prior knowledge of data mining. Understand the core of it and learning how to use and implement data mining algorithms will be a big challenge in this project.

Despite having some knowledge of Python, using this language on a larger scale to implement a whole system using machine learning algorithms, content analysis, latent semantic 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.

Milestones

Milestone 1 (Oct 2): Itemized Task

- Investigate most appropriate programming language
- Select required tools and frameworks
- Research and learn about Algorithms, Sentiment analysis and Text classification
- Select collaboration tools, Task organizer and planner, communication tools
- Create requirement document
- Create Design document
- Create Test plan

Milestone 2 (Oct 30): Itemized Task

- Implement, test and demo desired features
- Set up website
- Implement, test and demo Web Scrapping feature
- Implement, test and demo Graphical Visualizations

Milestone 3 (Nov 27): Itemized Task

- Improve website
- Implement, test and demo text classification
- Implement, test and demo sentiment analysis
- Implement, test and demo benchmarking analysis tool
- Revise and edit website layout, design and features.

Task Matrix for Milestone 1

Task	Claudino	Harshil
Investigate/Select Tools	Select language, Web Scrapping tool, Front end and Back end development	Select language, Select framework, Front end and Back end development
Investigate and select algorithms	Research on Algorithms, Sentiment analysis	Research on Algorithms, Text classification
Select collaboration tools	Presentation, documents, communication	Task Calendar, communication, Task organizer
Requirement Document	60%	40%
Design Document	40%	60%
Test Plan	50%	50%

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: _____