

HEALTH MONITORING ANALYTICS

ABSTRACT

Team Profile

Zhan Chen (Team leader) johnalwaysyoung@gmail.com.

Data analysis database and server set up.

Yue Song ys394@scarletmail.rutgers.edu

Webpage design, Google maps invoking.

Xiaoyu Yu xiaoyu.yu@rutger.edu

TwitterAPIData mining.

Goal

We wish to implement community tracking, develop personalized regimens for individuals and the community, perform small scale social networking between people, generate real time health statistics for a community, generate community status awareness bulletins, and privacy awareness.

Target Audience

1. Fitness Buffs
2. Average Consumers
3. Business Owners
4. Academics/Researchers
5. Government Officials

Source of Domain Knowledge

The source of domain knowledge of this schema would be entirely through programming and the aggregation of data. And the data would be mainly from Twitter.

Scope and Impact

It draws health related information from Twitter. With the big data analysis, we would be hoping to target some of the issues relating to exercise habits of people.

Processing Steps

1. Crawling Twitter Data (Using the Twitter API)

Twitter uses Open Authentication (OAuth) to keep users' information safe from third parties. Information we could retrieve from Twitter includes users' profile, network (followers, friends), and most importantly their tweets with relevant hashtag metadata.

2. Storing Twitter Data

We will use MongoDB as one example of a NoSQL implementation to deal with the large volume of Tweets, Tweeters, and network information.

3. Analyzing Twitter Data:

When analyzing Twitter data, we will use proper network measures to find the important topics and information that are relative to our project. First, we will explore a model that exploits the links between the entities to find key players in the data and understand the text to find relative information. Then, we will explore topic model that find natural topics in the text. (By using LDA algorithm, we will discover (select) topics in the text and build a (hashtag) automatically.

Technical Challenges

1. How to ensure whether this information is real monitoring information or not.
2. How to update and maintain the website information various over time.
3. How to display our data virtually, word description is needed how to display our data virtually, word description is needed.

References

- [1] Software Engineering Project: Health Monitoring Analytics
<http://www.ece.rutgers.edu/~marsic/books/SE/projects/HealthMonitor/analytics.html>
- [2] Kumar, Shamanth, Morstatter, Fred, and Huan Liu. Twitter Data Analytics. Springer, 2013 <http://tweettracker.fulton.asu.edu/tda/>
- [3] Kenneth. M. Anderson, Aaron Schram, "Design and Implementation of a Data Analytics Infrastructure in Support of Crisis Information Research (NIER Track)"
<http://epic.cs.colorado.edu/wp-content/uploads/icse2011.pdf>
- [4] Twitter Developers, <https://dev.twitter.com/>
- [5] Engrammi, <http://engrammi.blogspot.com/2011/10/nanni-balestrini-original-tag-cloud.html>