Sporty Twitters

Virgile Landeiro Dos Reis

Machine Learning & Social Media – 10/08/2013

Problem Statement

- Regular aerobic exercise can have benefits on health.

- Fewer studies focused on the psychological effects of exercising.

- Is it possible to correlate the physical activity of a person with his psychological well-being ?

Hypothesis

- The main hypothesis is that we can monitor the physical activity and the well-being of Twitter users by analyzing their tweets.

Method

Set of tweets

- Filtering with physical activity thanks to exercise tracker applications (RunKeeper, Runtastic, NikePlus, etc)

Filtered set of Twitter data

- Classification of users according to the frequency of exercising

Twitter users divided in 3 classes corresponding to the physical activity (frequent, occasionnal, none)

- Well-being analysis based on sentiment analysis (OpinionFinder or GPOMS)

Users classified by physical activity Users classified by well-being

- Correlation between exercising and well-being.

Data

How to collect data from Twitter ?

- Use an existing data set (1.6M tweets from Sentiment140).

- Use the Streaming API from Twitter.

- The objective is to obtain a set of several millions of tweets for the experiment to be legitimate.

Related Works

- Go, A., Huang, L., & Bhayani, R. (2009). Twitter sentiment analysis. *Entropy, 17.*

- This paper is focused on the sentiment analysis of tweets by classifying them in two categories : positive and negative.

Related Works

- Roberts, K., Roach, M. A., Johnson, J., Guthrie, J., & Harabagiu, S. M. (2012). EmpaTweet: Annotating and Detecting Emotions on Twitter. In *LREC* (pp. 3806-3813).

- Classification of the tweets in seven emotions : anger, disgust, fear, joy, love, sadness and surprise.

Related Works

- Deslandes, A., Moraes, H., Ferreira, C., Veiga, H., Silveira, H., Mouta, R., ... & Laks, J. (2009). Exercise and mental health: many reasons to move. *Neuropsychobiology*, 59(4), 191-198.

- This paper shows that exercise ensure successful brain functioning using non-machine learning techniques.

Timeline

10/15/2013 — • Complete Twitter data set

10/22/2013 — • Classification according to the physical activity completed.

11/5/2013 — • Well-being analysis completed

11/10/2013 — • Report completed

Questions ?