Apply Machine Learning in Real World Applications: Showcases from Student’s Work from Social Data Mining Seminar WS2014/2015

Institute of Cognitive Science Osnabrück

Machine Learning Practice SS2015
Background Data:

- Twitter Data on 12-14 of July 2014
- Number of Tweets on 12th of July: 11,441,826
- Number of Tweets on 13th of July: 13,439,779
- Number of Tweets on 14th of July: 11,091,365
Assignment 5: Sentiment Analysis of Worldcup 2014

Objective:

- Investigate the tweets **before** the start of the match for third place and final

Tweet Examples from 13th of July before the final started:

Germany is gonna win though :o but I don’t mind is Argentina wins #ARG will hoist the cup!
VAMOOOOS #GER
VAMO ARGENTINA
World Cup Finals! Gonna have to ride with Germany 2-1!
#GERvsARG #WorldCupFinal
VAAAAAAAAAMOS ALEMANHAA
Vaaaaaamos alemania
Argentina about to get the dub
#ARG #ARG #ArgentinaCampeonBrasil2014
Made it to Houston In time to watch the #worldcupfinal at @BWWings :) buzzing for this game ! Come on Ze Germans #worldcup #VamosMessi #VamosMessi #VamosMessi
Messi is si adorable
VAAAAAI ALEMANHA
Ready for Germany vs Argentina! Good luck to both these great teams! #WorldCupFinal
Great opportunity for Kramer here. Is no Khedira, but has had a brilliant season for M’gladbach
All the Argentinean names look like types of good coffee. #WorldCup
AUF GEHTS GERMANY
The sadness of Brazil can be explained by their humiliation by Germany the days before and they stayed sad and were angry about the world cup.

Spain was not happy. It could be explained that they were the last world champions but failed bitterly during the group stage.

Indonesia tweeted a lot about the world cup.

Red dot is for sadness, Blue dot is for happiness.
Assignment 5: Sentiment Analysis of Worldcup 2014 on 13th of July. Argentina vs Germany

Gray is for undecided

Red is for Germany  Blue is for Argentina
Project Ideas

... the extraction of airport locations merely from geo-location and timestamps of tweets... It shows paths of Twitter users that span over a large area, meaning that they users have traveled a long distance between the tweets, which is indicative of airplanes since you are usually not allowed to use mobile devices on-board, but it is likely that people will want to inform friends or followers that they are departing or arriving.

Objective:

- Trace twitter users in different regions of the world
- Analyze the sentiment of these tweets
- Localize airports based on timestamps and GPS tags

Authors: Jonas Görlich, Constantin Lehmann and Pascal Nieters
Subjective Sentiment Analysis, Movement Tracking and Airport Localization 2/4
Subjective Sentiment Analysis, Movement Tracking and Airport Localization 3/4

Plot

Red is airport calculated from latitude and longitude given in Twitter data
Black is imported data from real airports
Computer Vision Group:

http://ikw.uni-osnabrueck.de/~cv/