Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences

BTI7535 Data Science

Project: Sentiment Analysis

Prof. Dr. Jürgen Vogel (juergen.vogel@bfh.ch)

Sentiment Analysis Project (1)

- Implement your own Sentiment Analysis approach
  - as a Java / Maven project
  - details see following slides

- Teams
  - 3 students per team
  - 1 GitHub-Repository `ch.bfh.bti7535.w2017.{group name}`
    - add vgj1
  - send mail to vgj1 with names & git login names

- Evaluation Criteria
  - 60% solution & achieved results
  - 30% code & documentation & teamwork
  - 10% final presentation

Sentiment Analysis Project (2)

- Deadline for implementation and presentation slides: 14.01.18, 23:59

- Presentation of results: 15.01.18, 18:10 – 19:45
  - slides as PDF in doc project folder
  - 5 min per team plus Q&A
  - final solution implemented incl. lessons learnt
  - best results: accuracy

Sentiment Analysis Project (3)

1) text to be analyzed: pre-classified data set
2) design and implement an SA algorithm
    a) baseline: counting sentiment words
    b) sentiment classification with NB
        - features? implement your own idea(s)
3) evaluate implementation
    - 10-fold cross validation (stratified)
    - metric: accuracy
    - compare against
        a) gold standard
        b) baseline
        c) different iterations of your algorithm

Data: Sentiment Classification in Movie Reviews

- Polarity detection
  - is an IMDB movie review positive or negative?
- Movie review data
  - dataset with 1000 pos & 1000 neg reviews: http://www.cs.cornell.edu/people/pabo/movie-review-data/review_polarity.tar.gz
  - background information provided in readme file

ML Algorithm: Naïve Bayes Classifier

Weka http://www.cs.waikato.ac.nz/ml/weka/

- large and popular ML library written in Java
  - current stable version is 3.8.1
- GPL license
- Maven
  - http://search.maven.org/#search%7Cga%7C1%7Ca%3A%22weka-stable%22
- documentation https://www.cs.waikato.ac.nz/ml/weka/documentation.html
  - appendix of data mining book http://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf
  - wiki https://weka.wikispaces.com/Use+Weka+in+your+Java+code
  - javadoc http://weka.sourceforge.net/doc.stable-3-8/