

TACTiCS

Tool for Categorizing and Analyzing Text using Characteristics of Style

Kim Luyckx & Walter Daelemans

CNTS - Language Technology Group, University of Antwerp

We present a user-friendly tool for stylistic analysis of text and prediction of identity, age, personality, or gender of the author based on style characteristics.

1. Context & goals

Computational stylometry

- Meta information in an author's writing style

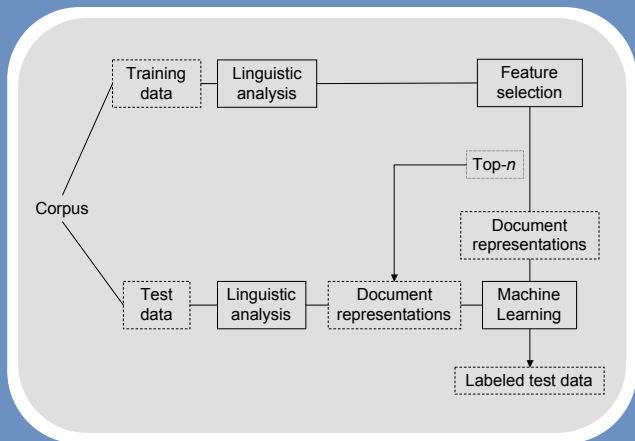

```
<doc id="0">
  <meta>
    <author>0</author>
    <age>25-40</age>
    <education>academic</education>
    <gender>female</gender>
    <personality>INTJ</personality>
  </meta>
  <text>...</text>
</doc>
```
- Machine Learning approach using NLP-based features

Goals of TACTiCS

- Covers essential tasks in computational stylometry
- Includes the use of Shallow Parsing and Machine Learning
- Performs evaluation of the predictions
- User-friendly interface, fit for humanities research and forensics
- Online available (free for academic purposes)
- Easily extensible due to its modular design

2. Design

- Modular system based on methodology that has been tested in previous studies in computational stylometry



- Modules
 1. Convert data to XML format and perform shallow parsing (MBSP or Tadpole)
 2. Feature selection
 3. Create feature vectors
 4. Perform classification (*k*-fold cross-validation)
 5. Evaluate predictions
- Developed in Python



3. Functionality

System with high coverage of tasks, features, and techniques in computational stylometry

- Tasks
 - authorship attribution & verification
 - personality prediction
 - gender prediction
- Feature types
 - n*-grams of chars, words, pos
 - '*n* most frequent *y*'
 - readability, sentence/word length
- Feature selection: χ^2 , Information Gain
- Machine Learning
 - TiMBL, MaxEnt, WEKA
 - k*-fold cross validation
- Evaluation in terms of precision, recall, F-score, accuracy and other metrics

4. Demo

A demo of the TACTiCS system is available from:
<http://www.cnts.ua.ac.be/stylometry/demo.html>

Acknowledgements

Funded by FWO Flanders



Stylometry project: <http://www.cnts.ua.ac.be/stylometry>

Kim Luyckx
<http://www.cnts.ua.ac.be/~kim>
kim.luyckx@ua.ac.be

Walter Daelemans
<http://www.cnts.ua.ac.be/~walter>
walter.daelemans@ua.ac.be