
Development of the opinion polls system

Maciej Wielgosz, Rafał Frączek, Marcin Pietroń
Paweł Rusek, Kazimierz Wiatr

Akademickie Centrum Komputerowe Cyfronet AGH

Objectives

Main objectives

- The work aims to develop an automatic public opinion polling system.

Data sources

- public web portals,
 - discussion forums,
 - electronic mailing lists
 - electronic publications (e. g. pdf)
-

Methods

- ❑ Automatic Internet content retrieval via RSS channels
- ❑ Text preprocessing
 - removal of all special characters,
 - removal of all redundant words,
 - lemmatization
- ❑ Numeric model construction
 - TF-IDF
- ❑ Classification
 - Nearest neighbour (cosine similarity measure)
 - K-means (PCA)