Contents

Feature Engineering

Fine-Grained Opinion Mining from Mobile App Reviews with Word Embedding Features .................................................. 3
Mario Sänger, Ulf Leser, and Roman Klinger

Feature Selection Using Multi-objective Optimization for Aspect Based Sentiment Analysis .................................................. 15
Md Shad Akhtar, Sarah Kohail, Amit Kumar, Asif Ekbal, and Chris Biemann

Feature Selection and Class-Weight Tuning Using Genetic Algorithm for Bio-molecular Event Extraction .................................. 28
Amit Majumder, Asif Ekbal, and Sudip Kumar Naskar

Automated Lexicon and Feature Construction Using Word Embedding and Clustering for Classification of ASD Diagnoses Using EHR .......... 34
Gondy Leroy, Yang Gu, Sydney Pettygrove, and Margaret Kurzius-Spencer

Multi-objective Optimisation-Based Feature Selection for Multi-label Classification .................................................. 38
Mohammed Arif Khan, Asif Ekbal, Eneldo Loza Mencia, and Johannes Fürnkranz

Information Extraction

WikiTrends: Unstructured Wikipedia-Based Text Analytics Framework .... 45
Michel Naim Gerguis, Cherif Salama, and M. Watheq El-Kharashi

An Improved PLDA Model for Short Text ........................................ 58
Chaotao Chen and Jiangtao Ren

Mining Incoherent Requirements in Technical Specifications .............. 71
Patrick Saint-Dizier

Enriching Argumentative Texts with Implicit Knowledge .................... 84
Maria Becker, Michael Staniek, Vivi Nastase, and Anette Frank

Technical Aspect Extraction from Customer Reviews Based on Seeded Word Clustering ........................................... 97
Jean-Marc Davril, Tony Leclercq, Maxime Cordy, and Patrick Heymans
Event Detection for Heterogeneous News Streams
Ida Mele and Fabio Crestani

Twitter User Profiling Model Based on Temporal Analysis of Hashtags and Social Interactions
Abir Gorrab, Ferihane Khoubi, Ali Jaffal, Bénédicte Le Grand, and Henda Ben Ghezala

Extracting Causal Relations Among Complex Events in Natural Science Literature
Biswanath Barik, Erwin Marsi, and Pinar Öztürk

Supporting Experts to Handle Tweet Collections About Significant Events
Ali Hürriyetoğlu, Nelleke Oostdijk, Mustafa Erkan Başar, and Antal van den Bosch

Named Entity Classification Based on Profiles: A Domain Independent Approach
Isabel Moreno, M.T. Romá-Ferri, and Paloma Moreda

Information Extraction from Resource-Scarce Languages

Does the Strength of Sentiment Matter? A Regression Based Approach on Turkish Social Media
Ali Mert Ertugrul, Itir Onal, and Cengiz Acarturk

AL-TERp: Extended Metric for Machine Translation Evaluation of Arabic
Mohamed El Marouani, Tarik Boudaa, and Nourddine Enneya

A Morphological Approach for Measuring Pair-Wise Semantic Similarity of Sanskrit Sentences
Vaishakh Keshava, Mounica Sanapala, Akshay Chowdhu Dinesh, and Sowmya Kamath Shevgoor

Word-Level Identification of Romanized Tunisian Dialect
Chaima Aridhi, Hadhemi Achour, Emna Souissi, and Jihene Younes

Named Entity Recognition in Turkish: Approaches and Issues
Doğan Küçük, Nursal Arıcı, and Dilek Küçük

Natural Language Processing Applications

Simplified Text-to-Pictograph Translation for People with Intellectual Disabilities
Leen Sevens, Vincent Vandeghinste, Ineke Schuurman, and Frank Van Eynde
Contents

What You Use, Not What You Do: Automatic Classification of Recipes . . . . 197
   Hanna Kicherer, Marcel Dittrich, Lukas Grebe, Christian Scheible,
   and Roman Klinger

Constructing Technical Knowledge Organizations
from Document Structures .................................................. 210
   Sebastian Furth and Joachim Baumeister

Applications of Natural Language Techniques in Solving
SmartHome Use-Cases .......................................................... 214
   Parnab Kumar Chanda, Nitish Varshney, and Ashok Subash

A Method for Querying Touristic Information Extracted from the Web . . . . 218
   Ermelinda Oro and Massimo Ruffolo

Towards Generating Object-Oriented Programs Automatically
from Natural Language Texts for Solving Mathematical Word Problems . . 222
   Sourav Mandal and Sudip Kumar Naskar

Authorship Attribution System .................................................. 227
   Oleksandr Marchenko, Anatoly Anisimov, Andrii Nykonenko,
   Tetiana Rossada, and Egor Melnikov

Neural Language Models and Applications

Estimating Distributed Representations of Compound Words
Using Recurrent Neural Networks .............................................. 235
   Natthawut Kertkeidkachorn and Ryutaro Ichise

A Large-Scale CNN Ensemble for Medication Safety Analysis ................. 247
   Liliya Akhtyamova, Andrey Ignatov, and John Cardiff

A Feature Based Simple Machine Learning Approach with Word
Embeddings to Named Entity Recognition on Tweets ........................ 254
   Mete Taşpınar, Murat Can Ganiz, and Tankut Acarman

Challenges and Solutions with Alignment and Enrichment of Word
Embedding Models ............................................................... 260
   Cem Şafak Şahin, Rajmonda S. Caceres, Brandon Oselio,
   and William M. Campbell

Legalbot: A Deep Learning-Based Conversational Agent
in the Legal Domain .................................................................. 267
   Adebayo Kolawole John, Luigi Di Caro, Livio Robaldo,
   and Guido Boella
Using Word Embeddings for Computing Distances Between Texts and for Authorship Attribution ........................................ 274
Armin Hoenen

Combination of Neural Networks for Multi-label Document Classification ......................................................... 278
Ladislav Lenc and Pavel Král

Supporting Business Process Modeling Using RNNs for Label Classification ....................................................... 283
Philip Hake, Manuel Zapp, Peter Fettke, and Peter Loos

A Convolutional Neural Network Based Sentiment Classification and the Convolutional Kernel Representation ...................... 287
Shen Gao, Huaping Zhang, and Kai Gao

Quotology - Reading Between the Lines of Quotations .......................................................... 292
Dwijen Rudrapal, Amitava Das, and B. Bhattacharya

Opinion Mining and Sentiment Analysis

Automatically Labelling Sentiment-Bearing Topics with Descriptive Sentence Labels .............................................. 299
Mohamad Hardyman Barawi, Chenghua Lin, and Advaith Siddharthan

Does Optical Character Recognition and Caption Generation Improve Emotion Detection in Microblog Posts? .................. 313
Roman Klinger

Identifying Right-Wing Extremism in German Twitter Profiles: A Classification Approach ............................................ 320
Matthias Hartung, Roman Klinger, Franziska Schmidtke, and Lars Vogel

An Approach for Defining the Author Reputation of Comments on Products ..................................................... 326
Carlos A. de Sa, Roney L. de S. Santos, and Raimundo S. Moura

Quality of Word Embeddings on Sentiment Analysis Tasks .......................................................... 332
Erion Çano and Maurizio Morisio

Question Answering Systems and Applications

A Hierarchical Iterative Attention Model for Machine Comprehension .......................................................... 341
Zhuang Liu, Degen Huang, Yunxia Zhang, and Chuan Zhang

A Syntactic Parse-Key Tree-Based Approach for English Grammar Question Retrieval .................................................. 353
Lanting Fang, Luu Anh Tuan, Lenan Wu, and Siu Cheung Hui
A Discriminative Possibilistic Approach for Query Translation Disambiguation ........................................ 366
   *Wiem Ben Romdhane, Bilel Elayeb, and Narjès Bellamine Ben Saoud*

Building TALAA-AFAQ, a Corpus of Arabic Factoid Question-Answers for a Question Answering System. .......... 380
   *Asma Aouichat and Ahmed Guessoum*

Detecting Non-covered Questions in Frequently Asked Questions Collections ................................. 387
   *Mladen Karan and Jan Šnajder*

**Semantics-Based Models and Applications**

Vector Space Representation of Concepts Using Wikipedia Graph Structure . . 393
   *Armin Sajadi, Evangelos E. Milios, and Vlado Keselj*

Composite Semantic Relation Classification .................................................. 406
   *Siamak Barzegar, Andre Freitas, Siegfried Handschuh, and Brian Davis*

Vector Embedding of Wikipedia Concepts and Entities ....................... 418
   *Ehsan Sherkat and Evangelos E. Milios*

TEKNO: Preparing Legacy Technical Documents for Semantic Information Systems .................................. 429
   *Sebastian Furth, Maximilian Schirm, Volker Belli, and Joachim Baumeister*

Improving Document Classification Effectiveness Using Knowledge Exploited by Ontologies. .................. 435
   *Zenun Kastrati and Sule Yildirim Yayilgan*

**Text Summarization**

Document Aboutness via Sophisticated Syntactic and Semantic Features . . 441
   *Marco Ponza, Paolo Ferragina, and Francesco Piccinno*

Summarizing Web Documents Using Sequence Labeling with User-Generated Content and Third-Party Sources. . 454
   *Minh-Tien Nguyen, Duc-Vu Tran, Chien-Xuan Tran, and Minh-Le Nguyen*

“What Does My Classifier Learn?” A Visual Approach to Understanding Natural Language Text Classifiers .......... 468
   *Jonas Paul Winkler and Andreas Vogelsang*
Gated Neural Network for Sentence Compression
Using Linguistic Knowledge. .................................................. 480
   Yang Zhao, Hajime Sembro, Xiaoyu Shen, and Akiko Aizawa

A Study on Flexibility in Natural Language Generation
Through a Statistical Approach to Story Generation. ...................... 492
   Marta Vicente, Cristina Barros, and Elena Lloret

Author Index ................................................................. 499
Natural Language Processing and Information Systems
22nd International Conference on Applications of
Natural Language to Information Systems, NLDB 2017,
Liège, Belgium, June 21-23, 2017, Proceedings
Frasincar, F.; Ittoo, A.; Nguyen, L.M.; Métais, E. (Eds.)
2017, XXII, 500 p. 91 illus., Softcover
ISBN: 978-3-319-59568-9