Contents

Natural Language Processing

Automated Detection of Adverse Drug Reactions from Social Media Posts with Machine Learning .................................................. 3

Ilseyar Alimova and Elena Tutubalina

Automated Detection of Non-Relevant Posts on the Russian Imageboard “2ch”: Importance of the Choice of Word Representations ................. 16

Amir Bakarov and Olga Gureenkova

A Morphological Processor for Russian with Extended Functionality .......... 22

Elena I. Bolshakova and Alexander S. Sapin

SyntaxNet Errors from the Linguistic Point of View .......................... 34

Oleg Durandin, Alexey Malafeev, and Nikolai Zolotykh

Size vs. Structure in Training Corpora for Word Embedding Models:
Araneum Russicum Maximum and Russian National Corpus ............... 47

Andrey Kutuzov and Maria Kunilovskaya

Combining Thesaurus Knowledge and Probabilistic Topic Models ........... 59

Natalia Loukachevitch, Michael Nokel, and Kirill Ivanov

Russian-Language Question Classification: A New Typology
and First Results ................................................................. 72

Kirill Nikolaev and Alexey Malafeev

Domain Adaptation for Resume Classification Using Convolutional
Neural Networks .................................................................. 82

Luiza Sayfullina, Eric Malmi, Yiping Liao, and Alexander Jung

Fighting with the Sparsity of Synonymy Dictionaries for Automatic
Synset Induction .................................................................. 94

Dmitry Ustalov, Mikhail Chernoskutov, Chris Biemann,
and Alexander Panchenko

Men Are from Mars, Women Are from Venus: Evaluation and Modelling
of Verbal Associations ......................................................... 106

Ekaterina Vylomova, Andrei Shcherbakov, Yurii Philippovich,
and Galina Cherkasova
Rotations and Interpretability of Word Embeddings: The Case of the Russian Language .................................................. 116
Alexey Zobnin

General Topics of Data Analysis

HuGaDB: Human Gait Database for Activity Recognition from Wearable Inertial Sensor Networks .................................................. 131
Roman Chereshnev and Attila Kertész-Farkas

On Finding Maximum Cardinality Subset of Vectors with a Constraint on Normalized Squared Length of Vectors Sum .................. 142
Anton V. Eremeev, Alexander V. Kelmanov, Artem V. Pyatkin, and Igor A. Ziegler

Using Cluster Analysis for Characteristics Detection in Software Defect Reports .................................................. 152
Anna Gromova

A Machine Learning Approach to Enhanced Oil Recovery Prediction ....... 164
Fedor Krasnov, Nikolay Glavnov, and Alexander Sitnikov

An Approach to Establishing the Correspondence of Spatial Objects on Heterogeneous Maps Based on Methods of Computational Topology ...... 172
Sergey Eremeev, Kirill Kuptsov, and Semyon Romanov

Predicting Winning Team and Probabilistic Ratings in “Dota 2” and “Counter-Strike: Global Offensive” Video Games .................. 183
Ilya Makarov, Dmitry Savostyanov, Boris Litvyakov, and Dmitry I. Ignatov

Bagging Prediction for Censored Data: Application for Theatre Demand .... 197
Evgeniy M. Ozhegov and Alina Ozhegova

Original Loop-Closure Detection Algorithm for Monocular vSLAM ........ 210
Andrey Bokovoy and Konstantin Yakovlev

Analysis of Images and Video

Organizing Multimedia Data in Video Surveillance Systems Based on Face Verification with Convolutional Neural Networks .......... 223
Anastasiia D. Sokolova, Angelina S. Kharchevnikova, and Andrey V. Savchenko

Satellite Image Forgery Detection Based on Buildings Shadows Analysis .... 231
Andrey Kuznetsov and Vladislav Myasnikov
Nonlinear Dimensionality Reduction of Hyperspectral Data
Using Spectral Correlation as a Similarity Measure .......................... 237
Evgeny Myasnikov

Large-Scale Shape Retrieval with Sparse 3D Convolutional
Neural Networks .......................................................... 245
Alexandr Notchenko, Yermek Kapushev, and Evgeny Burnaev

Floor-Ladder Framework for Human Face Beautification ................. 255
Yulia Novskaya, Sun Ruoqi, Hengliang Zhu, and Lizhuang Ma

Array DBMS and Satellite Imagery: Towards Big Raster Data
in the Cloud ................................................................. 267
Ramon Antonio Rodriges Zalipynis, Evgeniy Pozdeev, and Anton Bryukhov

Impulsive Noise Removal from Color Images
with Morphological Filtering ............................................. 280
Alexey Ruchay and Vitaly Kober

Optimization Problems on Graphs and Network Structures

An Exact Polynomial Algorithm for the Outerplanar Facility Location
Problem with Improved Time Complexity ................................ 295
Edward Gimadi

Approximation Algorithms for the Maximum $m$-Peripatetic
Salesman Problem ......................................................... 304
Edward Kh. Gimadi and Oxana Yu. Tsidulko

A Randomized Algorithm for 2-Partition of a Sequence ................. 313
Alexander Kel’manov, Sergey Khamidullin, and Vladimir Khandeev

An Approximation Scheme for a Weighted Two-Cluster
Partition Problem ......................................................... 323
Alexander Kel’manov, Anna Motkova, and Vladimir Shenmaier

Hitting Set Problem for Axis-Parallel Squares Intersecting a Straight
Line Is Polynomially Solvable for Any Fixed Range of Square Sizes . . 334
Daniel Khachay, Michael Khachay, and Maria Poberiy

Polynomial Time Solvable Subclass of the Generalized Traveling Salesman
Problem on Grid Clusters .................................................. 346
Michael Khachay and Katherine Neznakhina
Stabbing Line Segments with Disks: Complexity and Approximation Algorithms .......................... 356
  Konstantin Kobylkin

Analysis of Dynamic Behavior Through Event Data

On the Efficient Application of Aho-Corasick Algorithm in Process Mining ................................. 371
  Andrey M. Konchagin and Anna A. Kalenkova

Social Network Analysis

Health, Grades and Friendship: How Socially Constructed Characteristics Influence the Social Network Structure .................. 381
  Sofía Dokuka, Ekaterina Krekhovets, and Margarita Priymak

Dynamic Semantic Network Analysis of Unstructured Text Corpora ...................... 392
  Alexander Kharlamov, Galina Gradoselskaya, and Sofía Dokuka

Scientific Matchmaker: Collaborator Recommender System .................. 404
  Ilya Makarov, Oleg Bulanov, Olga Gerasimova, Natalia Meshcheryakova, Ilia Karpov, and Leonid E. Zhukov

Author Index .......................... 411
Analysis of Images, Social Networks and Texts
6th International Conference, AIST 2017, Moscow, Russia, July 27–29, 2017, Revised Selected Papers
2018, XXX, 412 p. 105 illus., Softcover
ISBN: 978-3-319-73012-7