Maksim Tkachenko

CONTACT INFORMATION

Homepage: www.mtkachenko.info
LinkedIn: maksim-tkachenko
Google Scholar ID: AXrs46UAAAAJ

EDUCATION

PhD, Singapore Management University, Singapore
School of Information Systems,
2015 - 2019, Full-ride Scholarship Awarded + Competitive Presidential Doctoral Fellowship Awarded
Thesis: Comparison Mining from Text
Supervisor: Hady W. Lauw
Dissertation Committee: Jiang Jing, Zheng Baihua, Xiaoli Li

Specialist, St. Petersburg State University, Russia
Faculty of Mathematics and Mechanics,
Software Engineering Department,
2006 - 2011, Competitive Full-ride Scholarship Awarded
Thesis: Named Entity Recognition Technique Based on Wikipedia
Supervisor: Boris Novikov

RESEARCH INTERESTS

Data Science, Text Mining, Natural Language Processing, Data Mining, Machine Learning, Information Extraction, Statistics, Graphical Models

PUBLICATIONS (PEER-REVIEWED)

[1] Maksim Tkachenko and Hady W Lauw. "CompareLDA: A Topic Model for Document Comparison". In: *Proc. of AAAI Conference on Artificial Intelligence,* **AAAI**, *Full Paper*. 2019.

[2] Maksim Tkachenko, Chong Cher Chia, and Hady W. Lauw. "Searching for the X-Factor: Exploring Corpus Subjectivity for Word Embeddings". In: *Proc. of Annual Meeting of the Association for Computational Linguistics,* **ACL**, *Full Paper*. 2018.

[3] Maksim Tkachenko and Hady W. Lauw. "Comparative Relation Generative Model". In: *IEEE Transactions on Knowledge and Data Engineering,* **TKDE**, *Journal Paper,* 29.4 (2017).

[4] Maksim Tkachenko and Hady W. Lauw. "Plackett-Luce Regression Mixture Model for Heterogeneous Rankings". In: *Proc. of ACM International Conference on Information and Knowledge Management,* **CIKM**, *Full Paper*. 2016.

[5] Maksim Tkachenko and Hady W. Lauw. "A Convolution Kernel Approach to Identifying Comparisons in Text". In: *Proc. of Annual Meeting of the Association for Computational Linguistics and International Joint Conference on Natural Language Processing,* **ACL-IJCNLP**, *Full Paper*. 2015.

[6] Maksim Tkachenko and Hady W. Lauw. "Generative Modeling of Entity Comparisons in Text". In: *Proc. of ACM International Conference on Information and Knowledge Management,* **CIKM**, *Full Paper*. 2014.

[7] Rinat Gareev, Maksim Tkachenko, Valery Solovyev, Andrey Simanovsky, and Vladimir Ivanov. "Introducing Baselines for Russian Named Entity Recognition". In: *Proc. of Computational Linguistics and Intelligent Text Processing,* **CICLing**, *Full Paper*. 2013.

[8] Maksim Tkachenko and Andrey Simanovsky. "Named Entity Recognition: Exploring Features". In: *Proc. of Conference on Natural Language Processing,* **KONVENS**, *Full Paper*. 2012.

[9] Maksim Tkachenko and Andrey Simanovsky. "Selecting Features for Domain-Independent Named Entity Recognition". In: *Proc. of Conference on Natural Language Processing,* **KONVENS**, *Short Paper*. 2012.

[10] Maksim Tkachenko, Alexander Ulanov, and Andrey Simanovsky. "Classifying Wikipedia Entities into Fine-Grained Classes". In: *Proc. of IEEE International Conference on Data Engineering (Workshops),* **ICDEW**, *Workshop Paper*. 2011.

Employment

School of Information Systems, Singapore Management University

Research Scientist 12/2018 (Current Position)
Research Assistant 11/2013 - 07/2015

Leading data infrastructure project, which aims to assemble a comprehensive knowledge base for e-commerce. Analysing the language of online reviews, designing methods for opinion and comparison mining. Co-supervising graduate and undergraduate students.

St. Petersburg State University, Russia

Research Engineer 07/2012 - 11/2013
Junior Researcher 11/2011 - 07/2012

Worked on named entity recognition (NER) and entity linking systems for a number of European languages (e.g., English, French). Designed and validated an active learning scheme for training sequential labeling models. Developed the state-of-the-art named entity recognition system for English. Experimented with word/phrase embedding methods for sequential labeling.

Hewlett Packard Labs, St. Petersburg, Russia

Research Engineer 11/2011 - 11/2013
Research Intern 12/2009 - 11/2011

Joint position with St. Petersburg State University (please refer to the above-mentioned for details). Designed systems for fine-grained text classification. Proposed and validated a fine-grained classifier for Wikipedia.

Lanit-Tercom, Inc., R&D department, St. Petersburg, Russia

Software Engineer 09/2009 - 12/2010

Developing news media information retrieval and opinion mining systems. Worked on low-level language processing for Russian (e.g., part-of-speech tagging, lemmatization).

JetBrains, MPS Team, St. Petersburg, Russia

Software Developer 06/2009 - 09/2009

Developing compilers for scripting languages (i.e., bash and cmd).

Awards & Grants

- 2017-2018 & 2018-2019, Presidential Doctoral Fellow, Singapore Management University (competitive, the recipients of this fellowship are selected from the top 5% of PhD students across SMU). Total funding awarded: 78000 SGD.
- Development Technology of Information Extraction Systems for the Russian Language, 2013-2015. (with V. Ivanov, F. Nikolaev, M. Sidikov, A. Simanovsky, R. Gareev, R. Yamilov and V. Solovyev). The project is supported by Russian Foundation for Basic Research.
- SIGIR Student Travel Grants, 2016. 470 USD.

Academic Service

- PC Member: IJCAI 2021, PAKDD 2021, AAAI 2021, IJCAI 2020, PAKDD 2020, AAAI 2020, SEIM 2020, SEIM 2018, SEIM 2017, SEIM 2016, SYRCoDIS 2014.
- Invited Conference Reviewer: DSAA 2017, PAKDD 2017, DSAA 2016, PAKDD 2015, CIKM 2014, ICDM 2014.
- Invited Journal Reviewer:
  - Transactions of the Association for Computational Linguistics (TACL)
  - Language Resources and Evaluation Journal
  - Journal of the Association for Information Science and Technology (JASIST)

Public Talks

- *Sentiment-Infused Word Embeddings with SentiVec*,
  - VNU Hanoi, University of Engineering and Technology, Vietnam, Dec 2019
  - VNU Ho Chi Minh University of Science, Vietnam, Dec 2019
  - Data Science Meetup by SEEK Asia & Big Data Malaysia, Kuala Lumpur, Malaysia, Mar 2019
- *CompareLDA: A Topic Model for Document Comparison*, Preferred.AI TechFest, Singapore, Aug 2019
- *SentiVec: Sentiment-Infused Word Embeddings*, Data Science Meetup by Singapore Data Science Consortium, Singapore, Jun 2018
- *Learning User Preferences from Multi-Modal Data: Preference Signal from Review Text*,
  - St. Petersburg Academic University, Russia, Mar 2018
  - Computer Science Club of Steklov Institute of Mathematics, St. Petersburg, Mar 2018
  - ITMO University, St. Petersburg, Russia, Mar 2018
  - Avito Data Science Meetup, Moscow, Russia, Feb 2018
  - Sberbank-Technologies, Moscow, Russia, Feb 2018
  - Moscow Institute of Physics and Technology (MIPT), Russia, Feb 2018
- *Language & Artificial Intelligence*, an invited lecture for the Interdisciplinary Electronic Arts Survey (IEAS) course, National University of Singapore (NUS), Sep 2018
- *Comparison Mining*, The Three Minute Thesis (3MT) Competition, Nanyang Technological University, Singapore, Jun 2017

Teaching

- Summer Internship Mentor at Preferred.AI, 2019-2020
  - Neural Network Lab, a visual language for designing and evaluating neural network models.
  - SnappyBuyer, a platform that allows you to make better decisions when buying electronics.
- Co-designed and delivered Course on Web Data Extraction and Regression Analysis (polytechnic and undergraduate levels), Instructor, 2019 (90% of students gave 5 and 4 stars out of 5 to rate the course)
- Computational Thinking (undergraduate level), Workshop Instructor, 2017 (teaching feedback: 3.9/5)

Research Prototypes

- CompareLDA (Comparative Latent Dirichlet Allocation) learns predictive topic distributions that comply with the pairwise comparison observations.
- Plackett-Luce Regression Mixture Model finds preference groups within a population of rankers.
- Tree-SVM learns SVM-based classifiers over tree structures.
- SentiVec learns word embeddings and enriches them with specified sentiment information.
- Venom is a focused crawler framework.
- Simple-ML implements online learning algorithms for classification.

Languages

English (fluent), Russian (native), Chinese (basic)