Anna Potapenko

Curriculum Vitae

57 Orekhovy Blvd, Apt 180
Moscow, Russia
+7 (916) 820 62 70
anna.a.potapenko@gmail.com

Education

2014 – 2018 (current) PhD program in computer science, Faculty of Computer Science, Yandex & National Research University Higher School of Economics, Moscow, Russia.
Research area: Machine learning for natural language processing. Working on word embeddings and topic models for learning semantic representations from texts.
Adviser: Prof. Konstantin Vorontsov.

2009 – 2014 Faculty of Computational Mathematics and Cybernetics,
Lomonosov Moscow State University, Russia. *GPA 5.0 out of 5.0.*
Specialist degree in Applied Mathematics and Computer Science.
Thesis: "Linguistic Regularization of Probabilistic Topic Models".
Adviser: Prof. Konstantin Vorontsov.

Additional

2013 – 2015 **Yandex School of Data Analysis**, Russia. *GPA 5.0 out of 5.0.*
Master's-level program in Computer Science and Data Analysis.

2011 – 2013 Elective course of the English Language for preparing translators in «Applied Mathematics and Informatics», Moscow State University.

Work Experience

April – July 2017 *Software Engineer Intern*, **Google**
Research project on neural dialogue systems in Assistant Team in MTV, CA.

August 2016 – February 2017 *Data scientist*, **Yandex Data Factory**
Building machine learning models for YDF projects, e.g. predicting time series with gradient boosting and neural networks.

April – July 2016 *Software Engineer Intern*, **Google**
Working on distributional semantic models for automatic concept mining from texts in Ads Quality Team in Zurich.

2013 – 2015 *Research engineer*, **Computing Center of Russian Academy of Sciences**
Designing a library for multi-objective topic modeling of large text collections, based on novel additive regularization approach: BigARTM.org.

Teaching Experience

2016-2017 Natural Language Processing, Faculty of Computational Mathematics and Cybernetics at Moscow State University, *Instructor*.

2015-2016 Machine Learning, Yandex School of Data Analysis, *Teaching assistant*.

2014-2016 Machine Learning and Data Mining, Faculty of Computer Science at Higher School of Economics, *Teaching assistant*.

2014-2015 Bayesian Methods in Probabilistic Topic Modeling, Faculty of Computational Mathematics and Cybernetics at Moscow State University, *Instructor*.

Mathematical background

Data Analysis Machine Learning, Probabilistic Graphical Models, Representation and Deep Learning

Text mining Natural Language Processing, Machine Translation, Introduction to Linguistics

General Algorithms and Data Structures, Probability and Statistics, Algebra, and others

Programming skills

Programming C/C++, Python

Scientific MATLAB/Octave, Scikit-learn, Weka

Parallel Hadoop, Spark, Open MP, MPI

Typesetting LaTeX, HTML

Languages

Russian Native

English Advanced *TOEFL iBT 104 out of 120*
Cambridge English: First (FCE) Certificate

German Basic

Publications

In English:

- Potapenko A. A., Vorontsov K. V. *Robust PLSA Performs Better Than LDA.* The 35-th European Conference on Information Retrieval, ECIR-2013, Moscow, Russia, 24-27 March 2013. – Lecture Notes in Computer Science LNCS 7814, Springer-Verlag Germany, 2013. Pp. 784-787.
- Vorontsov K. V., Potapenko A. A. *Tutorial on Probabilistic Topic Modeling: Additive Regularization for Stochastic Matrix Factorization.* Analysis of Images, Social Networks, and Texts. – Springer, 2014, CCIS, vol. 436, pp. 29-46.
- Vorontsov K. V., Potapenko A. A. *Additive Regularization of Topic Models.* Machine Learning Journal, Special Issue "Data Analysis and Intelligent Optimization" – Springer, 2015. Volume 101, Issue 1, Page 303-323. DOI: 10.1007/s10994-014-5476-6.
- Vorontsov K.V., Potapenko A.A., Plavin A.V. Additive Regularization of Topic Models for Topic Selection and Sparse Factorization. The Third International Symposium On Learning And Data Sciences, April 20-22, 2015, Royal Holloway, University of London, UK. – Springer International Publishing Switzerland 2015, A. Gammerman et al. (Eds.): SLDS 2015, LNAI 9047, pp. 193-202.

In Russian:

- Vorontsov K.V., Potapenko A. A. *Regularization, Robustness and Sparsity of Probabilistic Topic Models.* Computer research and modeling, 2012. V. 4, N 4. Pp. 693-706.
- Vorontsov K.V., Potapenko A. A. *Robust Sparse Probabilistic Topic Models.* The 9-th International Conference "Intellectualization of Information Processing" (IIP-2012), Budva, Montenegro. – Moscow: Torus Press, 2012. Pp. 605-608.
- Vorontsov K.V., Potapenko A. A. *Modifications of Generalized EM-algorithm for Probabilistic Topic Modeling.* Journal of Machine Learning and Data Analysis (ISSN 2223-3792), 2013.
- Potapenko A. A. *Sparse Probabilistic Topic Models.* Theses of the 16-th Russian Conference "Mathematical Methods of Pattern Recognition", Kazan. – Moscow: MAKS Press, 2013, P. 89.
- Potapenko A. A. *Regularization of probabilistic topic model for forming topic kernels.* XXI International scientific conference "Lomonosov-2014". – Moscow: Issuing office of MSU, CMC, 2013. Pp. 80-82.
- Vorontsov K. V., Potapenko A. A. *Regularization of probabilistic topic models to improve interpretability and determine the number of topics.* International Conference on Computational Linguistics "Dialogue". – Computational Linguistics and Intellectual Technologies, Moscow, 2014. Pp. 707-719.

Schools and Events

- Deep learning hackathon devoted to Q&A systems and solving The Allen AI Science Challenge (kaggle), February 2016, Moscow, Russia. *Invited talk on "Word embeddings and topic models: bridging the gap".*
- Yandex School of Data Analysis Conference "Machine Learning: Prospects and Applications", Berlin, Germany, October 5-8 2015. *Poster on "Linguistic regularization of topic models".*
- The 5th Lisbon Machine Learning School LxMLS-2015. Topic of the school: Natural Language Understanding. July 16-23, Lisbon, Portugal.
- Visit to Microsoft Research Cambridge, April 2015. *Talk on "Additive regularization of topic models and its parallel implementation BigARTM.org".*
- The 8th Russian Summer School in Information Retrieval RuSSIR 2014, Nizhny Novgorod, Russia, August 18-22 2014. *Poster on "Additive Regularization for Learning Interpretable Topic Models".*

Honours and Awards

- Ilya Segalovich Fellowship in Computer Science, Yandex, 2015, 2016.
- Advanced Research Fellowship, Higher School of Economics, 2014 - 2018.
- Best Master's Thesis Award, Faculty of Computational Mathematics and Cybernetics, Moscow State University, 2014.
- Governmental Award for outstanding achievements in research and study, 2012, 2013, 2014.
- Lobachevsky Award for excellent studies, Moscow State University, 2014.