

Anna Potapenko

Curriculum Vitae

Orekhovy Blvd 57-180
115682, Moscow, Russia
☎ +7 (916) 820 62 70
✉ anna.a.potapenko@gmail.com
📄 apotapenko.com

Education

- 2014 – 2018 **PhD program in Computer Science, Yandex & National Research University**
(current) [Higher School of Economics](#), Moscow, Russia.
Thesis: "Embedding words and documents with topic modeling techniques".
Adviser: Prof. Konstantin Vorontsov.
- 2013 – 2015 **Yandex School of Data Analysis**, Russia. *GPA 5.0 out of 5.0.*
Two years advanced program in Computer Science and Data Analysis.
- 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 Informatics.
Thesis: "Linguistic Regularization of Probabilistic Topic Models".
Adviser: Prof. Konstantin Vorontsov.

Work Experience

- Nov. 2017 – *Academic guest*, **ETH Zurich**
- Apr. 2018 Research on sparsity and interpretability of word embeddings,
supervised by Prof. Thomas Hofmann at Data Analytics Lab.
- April – July *Software Engineer Intern*, **Google**
2017 Research project on neural dialogue systems in Assistant Team in MTV, CA.
- Aug. 2016 – *Data scientist*, **Yandex Data Factory**
- Feb. 2017 Building machine learning models for [YDF projects](#), e.g. predicting time series with
gradient boosting and neural networks.
- April – July *Software Engineer Intern*, **Google**
2016 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

- 2017 Natural Language Processing, [online course on Coursera](#), *Main author and instructor.*
- 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, Linear Algebra

Programming skills

- Languages C/C++, Python
- Libraries TensorFlow, Numpy/Scipy, Scikit-learn, NLTK, Gensim
- Typesetting L^AT_EX

Languages

- Russian Native
- English Advanced
- German Basic

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

Publications

In English:

- A. Potapenko, A. Popov, and K. Vorontsov *Interpretable probabilistic embeddings: bridging the gap between topic models and neural networks*. In: Artificial Intelligence and Natural Language (AINL-2017), CCIS, Springer.
- 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, 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. *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

- [Rep4NLP workshop at ACL 2017](#), July 30 - August 4, Vancouver, Canada. *Poster on “Regularized Topic Models for Sparse Interpretable Word Embeddings”*.
- [DataFest⁴](#): data science conference and workshops, Moscow, Russia, February 11-12, 2017. *Invited talk on “Vector representations of words and documents”*.
- [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 Conference “Machine Learning: Prospects and Applications”, October 2015, Berlin. *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”*.
- [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.