Anton Karazeev

Moscow, Russia
+7-977-490-21-83 • anton.karazeev@phystech.edu
https://www.linkedin.com/in/akarazeev

EXPERIENCE

Laboratory of Neural Networks and Deep Learning October 2017 — Present
Junior Researcher
Currently responsible for preparing practical and theoretical assignments for the course of Reinforcement Learning and theoretical assignments for the course of Natural Language Processing with the number of 100+ enrolled students each.

Laboratory of Functional analysis of the Genome June 2016 — Present
Research Assistant
Research on protein function analysis.
Text mining, Natural language processing, Keyword extraction, Machine learning algorithms. As an intermediate result the new method of keywords extraction using Information Theory proposed (ResearchGate).

Sberbank-Technology August — October 2017
Data Scientist
Responsible for Natural Language Processing projects. Participated in preparing the datasets and building baselines for competition Sberbank Data Science Journey which is based on SQuAD. Developed an analogue of Amazon Mechanical Turk to improve experience of colleagues who evaluated the quality of collected datasets (Python, Flask).

HiQE Group March — June 2017
R&D Data Scientist
Negotiated with IBM engineers and applied some of the IBM Watson's services in tasks of signal processing. The system of baby cry recognition was built.

EDUCATION

Moscow Institute of Physics and Technology 09.2014 — 08.2018 (expected)
Department of Innovation and High Technologies,
Undergraduate student (B.Sc.)

- Computer Science, Physics
- Diploma - "Advanced toolkit for biomedical texts processing"

PROJECTS

Frontopolar, Moscow February - June, 2017
Applied Reinforcement Learning for Stock Trading. State of the art results were reached.
Different approaches were tested including Q-learning and Recurrent Reinforcement Learning. References are listed here.

Contributed to Open source:

- **Gensim** - fixed issue #671
- **yandexdataschool/Practical_RL** - PR #12
- **My projects on GitHub**

SKILLS

- **Russian**: native, **English**: fluent, **German**: basics (A2)
- **Programming languages**: Python, C/C++, bash, R, experienced with SQL
- **Python libraries**: numpy, sklearn, pandas; **for NLP**: NLTK, Gensim; **for Deep Learning**: TensorFlow, PyTorch
- Experimented with RaspberryPi and Arduino. Projects are located on GitHub

TEACHING

Deep Reinforcement Learning October — Present 2017
Course at MIPT, based on rll.berkeley.edu/deeprlcourse/
Practical assignments

Deep Learning in Natural Language Processing March — Present 2017
Course at MIPT, based on cs224n.stanford.edu
Practical assignments

PUBLICATIONS

Medium Story August, 2017
"Generative Adversarial Networks (GANs): Engine and Applications"

Moscow Conference on Computational Molecular Biology July 27 - 30, 2017
Moscow, Russia
"Advanced Parser for Biomedical Texts", Poster, Thesis

ADDITIONAL EDUCATION

Deep|Bayes Summer School, Moscow August 26 - 30, 2017
"Summer school on Bayesian Methods in Deep Learning"

Bioinformatics Summer School, Moscow July 31 - August 5, 2017
"Big Data in Bioinformatics"

Natural Language Processing (based on cs224d.stanford.edu) 2016
by DeepHack Lab

Supercomputer technologies for atomistic modelling 2015
by Igor Morozov (IHED RAS)
Molecular Dynamics - program written in C using OpenMP framework for parallel computing. Used VMD for visualization.

HACKATHONS

LauzHack, EPFL, *"NN:Nerds" team member* November 11 - 12, 2017
1st place in challenge by SGS, Presentation
Solution allows quick access to the main concepts found in documents. Responsible for development of telegram-bot (Python) and processing documents using IBM Watson service for Natural Language Understanding. Devpost.

mABBYYlity, Phystechpark October 7 - 8, 2017
4th place, "App in the Restaurant" iOS application, Demo, Presentation
App allows to recognise entities from restaurant menus using smartphone's camera and translates them. ABBYY Real-Time Recognition SDK, ABBYY Lingvo API and Spoonacular API were used. Responsible for backend (Python).

Neurocampus, Skolkovo Moscow School of Management September 22 - 24, 2017
2nd place, "S.o.S. - Sense of Speech" telegram-bot, @SenseOfSpeech_bot
Solution allows to extract emotions from user's recorded speech. Also it helps to train selected emotion with samples from TED talks.
Speech Emotion Recognition (SER) module was used as a core for telegram-bot based system to help users improve speech during performances. Responsible for development (Python).

Bioinformatics Summer School, Moscow August 3 - 4, 2017
"Prediction of Experimental Metadata from Gene Expression"
Used Machine learning algorithms to predict phenotype by gene expression. Distinguish with high accuracy samples of male and female tissues of Mus musculus organism. Datasets from Gene Expression Omnibus were used. Project.

BioHack 2017, Saint Petersburg March 3 - 5, 2017
Text Mining, parsing the records from PubMed and UMLS.
Analysis of research trends of chemical compounds and diseases during period of 1990-2015 using parsed information from PubMed database. Project.