Alexander Rakhlin

+7(916)577-4958 | rakhlin@gmx.net | GitHub | Kaggle | LinkedIn

I am a Machine Learning Engineer and Data Scientist with over 10 years of experience in data analysis and research. Currently I specialize in medical and healthcare applications of Machine Learning.

EXPERIENCE AND PROJECTS

Return to Diabetic Retinopathy Detection

December 2016 – PRESENT

The project achieves state-of-the-art performance of sensitivity 99%, specificity 71% and AUC 0.97. bioRxiv 225508

Machine Learning projects on Kaggle

November 2013 – PRESENT

Flavours of Physics: Finding $\tau \rightarrow \mu\mu\mu$, 5th of 673. Jul 2015–Oct 2015. In search for a rare decay phenomenon – charged lepton flavour violation. Technologies used: Fully-Connected Neural Networks, Gradient Boosted Trees, model ensembling, GPU, HDF5. GitHub. Presentation for "Heavy Flavour Data Mining workshop". The sponsors: CERN, Kaggle, Intel, Yandex.

Facebook V: Predicting Check Ins, 38th of 1212 teams (top 10%). May 2016–Jul 2016. Predicted which place a person would like to check in to. Technologies used: Pandas, XGBoost, Scikit-learn. The sponsor: Facebook.

Yelp Restaurant Photo Classification. 22nd of 355 teams (top 10%). Dec 2015–Apr 2016. Predicted attribute labels for restaurants using user-submitted photos. LinkedIn publication. GitHub. Caffe, Python, Scikit-learn, Pandas, H5py, Theano, Keras, XGBoost.

Ultrasound Nerve Segmentation, 41st of 923 teams (top 5%). May 2016–Aug 2016. Identified nerve structures in ultrasound images to improve pain management through the use of indwelling catheters that block pain at the source. Python, Pandas, Keras

Sea Lion Population Count, 25th of 385 teams (top 7%). Mar 2017 – Jun 2017. Developed algorithms which accurately count the number of sea lions in aerial photographs. Keras Deep Learning environment. The sponsor: NOAA Fisheries. GitHub.

Diabetic Retinopathy Detection, 131st of 661 (top 25%). Feb 2015–Jul 2015. LinkedIn publication. Built an automated system for Diabetic Retinopathy Detection. Technologies: Fully-Connected Neural Networks, Theano, Keras, ensembling, GPU, HDF5, AWS. GitHub. The sponsors: California Healthcare Foundation.

Telstra Network Disruptions, 81st of 1004 teams (top 10%). Nov 2015–Feb 2016. Predicted service faults on Australia's largest telecommunications network. XGBoost, t-SNE.

Otto Group Product Classification Challenge, 218th of 3514 (top 10%). Mar 2015 – May 2015. Built a predictive model which is able to distinguish between main product categories. Technologies used: Fully-Connected Neural Networks, Keras, XGBoost, model ensembling, R, GPU (Keras), HDF5. GitHub. The sponsors: Otto Group.

Other Kaggle projects:

American Epilepsy Society Seizure Prediction Challenge (Nov 2014) • NIPS 2017: Adversarial Attacks and Defenses (Oct 2017) • Text Normalization Challenge – English/Russian Language (Nov 2017) • DecMeg2014 - Decoding the Human Brain (Apr 2014) • Grasp-and-Lift EEG Detection (Jun 2015) • The Allen AI Science Challenge (Feb 2016, top 25%) • Plankton image identification (Mar 2015, top 25%) • Rossmann Store Sales (Dec 2015, top 10%), daily sales prediction • NIH Seizure Prediction (Dec 2017, top 16%) • Rental Listing Inquiries (Apr 2017, top 3%).

Algo trading – Founder

December 2007 – PRESENT

Developer of market models and algorithms for derivative market on Moscow Exchange. Implemented and put into work an automated system for trading futures and options on equities index. Matlab/R/C#

Uniastrum Bank, Utrade.ru – Deputy Chief for operations on international markets.

July 2002 – October 2007

- Design, testing and implementation of trading strategies for investment funds.
- Brokerage services for the Bank clients.
- Accounting and trade automation. Localization of Interactive Brokers trading platform. C++/VB/Java

Education

National Research University of Electronic Technology (MIET)
MS in Computer Science and Microelectronic Devices, 1994

Awards and achievements

- Kaggle Master. A status awarded to some of the best Data Scientists in the world who have consistently submitted high-ranking solutions to the predictive modeling challenges hosted on kaggle.com, 2015.
- 190 (of 75,000) in global Kaggle ranking, as of January 2018
- Physics Prize: HEP meets Machine Learning Award. CERN, Universität Zürich, Yandex, Intel. December 2015. link

Publications

- Shvets, A., Rakhlin, A., Kalinin, A., Iglovikov, V.: Automatic Instrument Segmentation in Robot-Assisted Surgery Using Deep Learning. bioRxiv p. 275867 (2018)
- Rakhlin, A., Shvets, A., Iglovikov, V., Kalinin, A.: Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis. bioRxiv p. 259911 (2018)
- Iglovikov V., Rakhlin A., Kalinin A., Shvets A.: Pediatric Bone Age Assessment Using Deep Convolutional Neural Networks. bioRxiv p. 234120 (2017)
- Rakhlin, A.: Diabetic retinopathy detection through integration of deep learning classification framework. bioRxiv p. 225508 (2017)

Professional Certifications, Development and Training

- **edX**, CS1156x: Learning From Data, 2013
- **Stanford Online**, Statistical Learning, 2014
- **Coursera**, Neural Networks for Machine Learning,2017
- **Coursera**, Pattern Discovery in Data Mining, 2015
- **Coursera**, Cluster Analysis in Data Mining, 2015
- **Coursera**, Data Visualization, 2015
- **Coursera**, Natural Language Processing, 2012
- **Coursera**, Probabilistic Graphical Models, 2012
- **Coursera**, Introduction to Systematic Program Design 2013

Research & Development

• Machine Learning • Deep Learning • Medical Imaging • Data Clustering • Neural Networks • Support Vector Machines • Gradient Boosted Trees

Technologies

Frameworks, libraries, tools: Theano, Caffe, Keras, XGBoost, scikit-learn, neon, NumPy, SciPy, t-SNE, AWS, GPU,
Programming Languages: Python, Matlab, C#, R
IDEs: PyCharm, Anaconda, Microsoft Visual Studio, RStudio
Version Control Systems: Git, GitHub
Operating Systems: Windows, Linux (Ubuntu)

Knowledge/Skills

Machine Learning • Deep Learning • Data Clustering • Data preprocessing • Neural Networks (Fully Connected, CNN, RNN, Echo State) • Support Vector Machines • Gradient Boosted Trees • AWS • GPU • HDF5

Open Source Projects

- ICIAR 2018 Challenge on Breast Cancer Histology Images. link
- MICCAI 2017 Robotic Instrument Segmentation. link
- NIPS 2017 Adversarial contests. link
- Sea Lion Population Count. link
- Sentiment analysis. CNNs for Sentence Classification. link
- Diabetic Retinopathy Detection. link

Presentations

Presentation for "Heavy Flavour Data Mining workshop", February 18-19 2016, Zurich. link, link

Volunteer Experiences

Coursera - Community Teaching Assistant.
April 2015 – PRESENT. Supports students learning by clarifying points, explaining concepts, and addressing misunderstandings on the forums