Anton Dergunov
Curriculum Vitae

+44 77 999 41 638
anton.dergunov@gmail.com
www.linkedin.com/in/anton-dergunov/
https://anton-dergunov.github.io/blog/
https://github.com/anton-dergunov

Highlights and Skills

- **Machine Learning**: Extensive experience and deep interest in applying machine learning techniques for information retrieval, recommendation systems, and natural language processing. Proficient in Python, Jupyter Notebooks, NumPy, Pandas, SciKit-Learn, PySpark, PyTorch, LightGBM, and Azure Machine Learning. Skilled in linear models, SVM, Decision Trees, Ensemble Learning, Gradient Boosting, Neural Networks, Deep Learning, Large Language Models, Graph Neural Networks, CNN, RNN, LSTM, BERT, Knowledge Representation, model interpretability, exploratory data analysis, and working with eyes-off data. Successfully developed and deployed models in production environments and built ML pipelines, demonstrating significant improvements in key metrics.
- **Software Development**: Comprehensive experience in all phases of large software development projects, including design, prototyping, research, review, documentation, planning, and implementation. Proficient in C, C++, Java, Python, Bash, and SQL. Skilled in Linux, Windows, and cross-platform development. Focused on software quality, from design and TDD to formal software verification. Expertise in parallel, concurrent, and distributed software development, with a strong background in performance analysis and optimization of both serial and parallel software. Mentored and supervised colleagues and interns.

Employment

Senior Applied Scientist, Microsoft   May 2018 - Present
Applying machine learning in Microsoft Office products.

- People Search:
  - Built from scratch the pipeline for collecting ''eyes-off'' training and testing data, monitoring its quality, and training ranking and classification models with offline evaluation, producing a model ready for product integration. Optimized the pipeline's performance to process large volumes of click log data. The model is used for people ranking in many Microsoft products (such as Teams, Outlook, Bing, Windows). Deployed several models using this pipeline to production, demonstrating online metrics improvements (such as up to a 4% increase in click-through rate and improvements in time to success). (Azure Machine Learning, Python, PySpark, LightGBM, ONNX, Gradient Boosting)
  - Implemented the Personalized PageRank algorithm over organization people graphs for this people ranker, resulting in a 2.15% increase in NDCG@1 and a 1.21% increase in NCDG@3. Presented this work at an internal conference and tested other graph-based approaches produced by different teams.
  - Conducted multiple online A/B experiments, analyzing metrics to test new features or resolve issues.
- Document recommendation for Microsoft Outlook:
  - Enhanced the relevance of recommendations in the Meeting Insights feature of Microsoft Outlook (paper, press release, (documentation) using ML techniques and various heuristics. This feature recommends emails and documents relevant to meetings. (Python, SciKit-Learn, NumPy, Pandas, PyTorch, LightGBM, Gradient Boosting, C#)
  - Enhanced the data extraction pipeline, increasing the amount of training data by ~50x, which improved model performance (AUC up by 3%, AUPRC up by 12%). Further data analysis improved model performance (AUC up by 3%, AUPRC up by 4%, NDCG@3 up by 11%) and enabled twice as many features.
  - Developed and validated new features for this machine learning model by identifying and optimizing new signals (e.g., social, content similarity, key phrases).
  - Conducted extensive experiments using the BERT model and related neural network architectures to improve content similarity features in Meeting Insights: verified the model on multiple datasets, performed error analysis and comparison with existing features, and distilled BERT into simpler models. Integrating the model showed a 3% increase in NDCG@3. (BERT, Deep Learning, PyTorch)
  - Explored various data sets, such as user interaction logs, manually annotated data, heuristically labelled ''eyes-off'' data using weak supervision and their combinations for model training (Snorkel). Conducted A/B testing to validate improvements.
  - Prototyped a complete pipeline for data extraction, feature extraction, training, and validation of a document-to-document recommendation system. Assisted a colleague in implementing an email-to-email recommendation prototype.
  - Developed a tool for explaining ML models predictions, aiding in error analysis and accelerating root cause detection. Contributed to other internal debugging tools. Provided supervision and mentorship for 3 interns and support for several colleagues, serving as a go-to person for the products we were implementing.

Lead Application Developer, JP Morgan   May 2017 - May 2018
Web services implementation.

- Implemented backend systems for monitoring engineering practices and source code repository commits. (Java, Spring, Python, SQL, MongoDB, MariaDB, Oracle Database) Conducted phone screening interviews.

Employment (cont.)

Software Engineer, Yandex January 2016 - May 2017

Feature engineering for Internet video search and machine learning.

- Implemented new video characteristics for the MatrixNet machine learning algorithm, improving video ranking quality in Yandex Video Search. Several factors led to enhanced video ranking quality, as demonstrated by various metrics. (C++, Python, MapReduce, Machine Learning)
- Developed a machine learning pipeline and conducted large-scale web text classification experiments using logistic regression, LSTM, word2vec, and other algorithms. (Natural Language Processing, Deep Learning)
- Utilized the developed infrastructure for YouTube video classification, improving video ranking quality and click-through rate.
- Collaborated across teams to improve integration of Video Search results on the main Yandex Web Search page.

Senior Software Development Engineer, Intel February 2006 - December 2015

Applied research in performance analysis and implementation of performance tuning tools.

- Gained expertise in performance analysis of serial and parallel software; algorithmic, threading and machine-dependent performance optimization; CPU and hardware architecture; parallel, concurrent and distributed software research and development (MPI, OpenMP).
- Developed a capability to represent knowledge about performance issues in applications (using Python code) and their automatic detection for Intel VTune Amplifier and Intel Advisor products. (Knowledge Representation)
- Defined metrics and methodologies to identify parts of an application that are best suited for execution on CPUs or for offload to a coprocessor, and implemented a tool for automatic detection.
- Full-cycle development (design, prototype, research, review, documentation, planning, implementation, etc.) of the threading and performance profiler Intel VTune Amplifier from scratch, including major data visualization and data provider components. (C++, Java, Python, OpenMP, SQLite) Mentorship (5 collegues).
- Implemented a new approach for source-level performance information visualization in Intel VTune Amplifier, unlocking new usage models for senior performance engineers.

Lecturer, Research Advisor, Lobachevsky State University of Nizhni Novgorod 2009 - 2012

Research and teaching courses in Artificial Intelligence and Machine Learning.

- Supervised 3 bachelor and 5 master students on topics including artificial intelligence, machine learning, natural language processing, software verification, parallel programming, and software engineering.
- Developed and taught courses ''Artificial Intelligence'' (2010, 2011), ''Databases and Expert Systems'' (2009-2012).

Software Development Engineer, Mera Networks October 2003 - February 2006

Telecommunication software implementation.

- Developed software for emulating cellular base stations. (C++, Java, Bash, networking) Mentored 3 colleagues.
- Contributed to a major redesign of the software's data model, significantly improving execution speed and memory consumption, allowing support for previously unsupported data volumes.

Personal Projects

- Small repository of various Jupyter notebooks: https://github.com/anton-dergunov/ml-playground

Education

PhD in Theoretical Computer Science, Lobachevsky State University of Nizhni Novgorod July 2008 - July 2011

Thesis: Models, methods and tools of knowledge representation for improving performance of MPI applications.
Implemented results from the thesis in Intel Trace Analyzer and Collector and its Performance Assistant capability.

Specialist Degree in Applied Computer Science, Lobachevsky State University of Nizhni Novgorod September 2001 - May 2006

Thesis: Neural network model for knowledge representation and extraction.

Publications

- Anton Dergunov. Proving Termination and Computational Complexity of Computer Programs. Abstract of poster for ''CSEDays. Algorithms and Complexity'' summer school.
- Anton Dergunov. Generalized algebraic data types in Haskell. The Monad.Reader Issue 22, pp. 5-39. Etc.
- Anton Dergunov. Specification and automatic detection of performance problems in message passing (MPI) applications. Radioelectronic and Computer Systems, 2012, 7(59), pp. 217-222.
- Anton Dergunov. Expert methodology for performance analysis of communicating processes in MPI applications. In proceedings of conference ''Parallel Computational Technologies'', Chelyabinsk (Russia), 2012., pp. 134-144 (in Russian).
- Personal blog on machine learning: https://anton-dergunov.github.io/blog/