GitHub: MichaelSolotky
LinkedIn: michael-solotky

Michael Solotky

Gdańsk, Poland
jaworomichal128@gmail.com

Education

- **Master of Science** in **Applied Mathematics and Informatics**, GPA 3.90 / 4.0
  Higher School of Economics : Faculty of Computer Science — Sep 2019 – Jun 2021
  Joint programme with **Yandex School of Data Science**
- Bachelor of Science in Applied Mathematics and Computer Science, GPA 3.89 / 4.0
  Lomonosov Moscow State University — Sep 2015 – Jun 2019
  Faculty of Computational Mathematics and Cybernetics

Experience

- **Software Development Engineer at Amazon** — Aug 2021 – now
  Alexa TextToSpeech C | C++ | Python | Bash | Perl | Continuous Deployment
  - **Decreased latency** of a Deep Learning model for homograph disambiguation by **56%**
  - **Urgently fixing bugs** with wrong pronunciation helping to deliver projects on time
  - Extending functionality of an internal library for integration testing in Speech Synthesis
- **Research Science Intern at Yandex** PyTorch | NumPy | h5py | Pyplot | LaTeX — Sep 2020 – Jun 2021
  - **Quality validation** of existing methods for **uncertainty estimation** on complex tasks
  - Finding **theoretical foundations** for various methods of **uncertainty estimation in Deep Learning**
  - **Results** are described in the **Master's thesis**
- **Machine Learning Engineer Intern at Yandex** — Jun 2019 – Sep 2019
  Machine Translation department TensorFlow | MapReduce | SciPy | Pyplot
  - Conducted experiments to improve quality and diversity of translations
  - Analyzed baseline approaches and found some basic mistakes that they make
  - Implemented several successful models in company's internal machine learning library
  - **Achieved improvement in quality and diversity** by internal company's metrics and by commonly used machine translation metrics: about **10% of max-BLEU growth** and about **60% of self-BLEU diversity growth**
- **Software Engineer Intern at Yandex** — Jun 2018 – Oct 2018
  Voice Technology department C++ | Python | MapReduce | Protobuf
  - Implemented several methods of probability smoothing in language models for Speech Recognition
  - Conducted experiments on quality measurement to find the best model among all
  - Implemented an optimal algorithm for training n-gram language models in C++ using MapReduce, which **reduced training time by 3 times and slightly increased quality measure**

Projects

- **BigARTM** C++ | Boost | Protobuf | Travis CI | AppVeyor — Jan 2017 – May 2019
  Open Source library for topic modelling with support of multiple regularization
  - Development and support of a tool for parallel calculation of pairwise word statistics such as frequency of mutual occurrence, PMI in large text corpora in conditions of low RAM
  - **Wikipedia full-text processing takes less than 8 Gb of RAM now compared to at least 20 Gb needed before (code sample, documentation)**

Technical Skills

- **Languages used at work**: C++, Python, C, Bash, Perl
- **Basic knowledge**: SQL, Assembly language
- **Technologies/Libraries**: MapReduce, Protobuf, C++ Boost, Make, NumPy/SciPy, Sklearn, Pandas, CVXPY
- **Deep Learning frameworks**: PyTorch, TensorFlow, Keras, MXNet
- **Tools**: Git, Subversion, UNIX/Linux, Continuous Deployment, GDB, Docker, Travis CI, AppVeyor, LaTeX