GitHub: MichaelSolotky
LinkedIn: michael-solotky

Michael Solotky

Gdańsk, Poland
jawormichal128@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 – Present · 11mo
  Alexa TextToSpeech `C` `C++` `Python` `Bash` `Perl` `CI/CD` `MXNet`
  - **Reduced 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 **making it simple** to execute various new testing scenarios
- **Research Science Intern at Yandex** `PyTorch` `NumPy` `Pyplot` `LaTeX` — Sep 2020 – Jun 2021 · 9mo
  - Comparing existing methods for **uncertainty estimation** on large-scale 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 · 3mo
  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
  - **Increased quality and diversity** by internal company's metrics and by commonly used machine
    translation metrics: **10% of max-BLEU growth** and about **60% of self-BLEU diversity growth**
  - Implemented several models in company's internal machine learning library
- **Software Engineer Intern at Yandex** — Jun 2018 – Oct 2018 · 3mo
  Voice Technology department `C++` `Python` `MapReduce` `Protobuf`
  - Implemented several methods of probability smoothing in language models for
    Automatic 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**

PROJECTS

- **BigARTM** `C++` `Boost` `Protobuf` `Travis CI` `AppVeyor` — Jan 2017 – Jun 2018
  Open Source library for topic modelling
  Developed a tool for parallel calculation of pairwise word statistics (**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