

✉ mihlushi@yandex.ru

🌐 <https://github.com/MichaelSolotky>

in <https://www.linkedin.com/in/michael-solotky/>

Michael Solotky

Moscow, Russia

☎ +7 967 291-08-66

EDUCATION

- MSc student in Applied Mathematics and Computer Science
Higher School of Economics : Faculty of Computer Science September 2019 – June 2021
Joint programme with [Yandex School of Data Science](#)
- BSc in Applied Mathematics and Computer Science, GPA 4.89 / 5.0
Lomonosov Moscow State University September 2015 – June 2019
[Faculty of Computational Mathematics and Cybernetics](#)

Continuing with the PhD degree afterwards

EXPERIENCE

- **Machine Translation department of Yandex** June 2019 – September 2019
Software Engineering Intern (Machine learning engineer)
- **Voice Technology department of Yandex** June 2018 – October 2018
Software Engineering Intern (Back-end developer)
 - Implemented several methods of probability smoothing and their modification in language models for Automatic Speech Recognition
 - Conducted experiments on quality measurement to find the best model among all
 - Implemented an optimal algorithm for constructing n-gram language models in C++ using MapReduce, **which decreased wall time by at least 3 times and slightly increased quality measure comparing to the baseline**
 - Wrote a complete framework with a set of operations available from CLI

PROJECTS

- **BigARTM (C++ Boost/STL, Protobuf, Travis, AppVeyor)** January 2017 – present
Open Source library for topic modelling that supports multiple regularization
🌐 github.com/bigartm/bigartm
 - Developed a tool for parallel calculation of pairwise word statistics such as frequency of mutual occurrence, PMI in text corpus of unlimited size
Wikipedia full-text processing takes 6 hours on octa-core intel core i5 8th gen
 - Responsible for parsing of input data

OTHER EXPERIENCE

- **Primality test (C++)**
Implementation of Miller test for deterministic checking of large numbers for primality as part of the [Kaspersky Lab](#) information security course
Wall time on prime numbers of length 100 is about 4 seconds
🌐 github.com/MichaelSolotky/sandbox/tree/master/Cpp_old_tasks/Primality_tests
- **ML (NumPy, Scipy)**
Implementation of various ML algorithms from scratch
🌐 github.com/MichaelSolotky/sandbox/tree/master/ML

TECHNICAL SKILLS

- **Used at work:** C++, Python, C, Bash
- **Basic knowledge:** SQL, Assembly language
- **Technologies:** MapReduce, Protobuf, C++ Boost, CMake, Make, SciPy, Scikit-learn, NumPy, Pandas
- **Deep Learning frameworks:** PyTorch, TensorFlow
- **Tools:** Git, Subversion, UNIX/Linux, Travis, AppVeyor, L^AT_EX