

EDUCATION

BSc in Applied Mathematics and Computer Science, IV year

- [Lomonosov Moscow State University](#) September 2015 - June 2019
Faculty of Computational Mathematics and Cybernetics

Continuing with the Master's degree afterwards, expected graduation in 2021

EXPERIENCE

- **Voice technology department of [Yandex](#)** June 2018 - October 2018
Software Engineering Intern
 - Implemented several methods of probability smoothing and their modification in language models for Automatic Speech Recognition
 - Carried out 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 surpassed baseline by up to 5 times in wall time on checked datasets and slightly in quality**
 - 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
🔗 github.com/bigartm/bigartm
 - Developed a tool for parallel calculation of pairwise word statistics such as frequency of mutual occurrence, PMI in text collections of unlimited size
Wikipedia full text processing takes 6 hours on a laptop with quad core AMD A6

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:** C++ Boost, CMake, Make, Protobuf, MapReduce, SciPy, Scikit-learn, NumPy, Pandas, Pytorch
- **Tools:** Git, Subversion, UNIX/Linux, Travis, AppVeyor, L^AT_EX