

ОБРАЗОВАНИЕ

Бакалавр, специальность: прикладная математика и информатика

- **МГУ имени М.В. Ломоносова** Сентябрь 2015 – Июнь 2019
Факультет Вычислительной Математики и Кибернетики, ср. балл 4.89 / 5.0

Планирую продолжать учёбу в магистратуре, ожидаемый год окончания: 2021

ОПЫТ РАБОТЫ

- **Отдел машинного перевода. Яндекс** Июнь 2019 – Сентябрь 2019
Стажёр-разработчик машинного обучения
 - .
- **Группа разработки голосовых технологий. Яндекс** Июнь 2018 – Октябрь 2018
Стажёр-разработчик бэкэнда
 - Реализовано несколько методов сглаживания вероятностей в языковых моделях для распознавания речи
 - Проведены эксперименты по сравнению качества для нахождения лучшей модели среди использованных
 - Реализован оптимальный алгоритм построения n-граммных языковых моделей на C++ с использованием MapReduce, **время работы которого минимум в 3 раза меньше базовой реализации, а также он немного превосходит по метрике качества базовую реализацию на некоторых датасетах**
 - Написан фреймворк с операциями, доступными из терминала

ПРОЕКТЫ

- **BigARTM (C++ Boost/STL, Protobuf, Travis, AppVeyor)** Январь 2017 – настоящее время
Библиотека с открытым кодом по тематическому моделированию, поддерживающая одновременное использование множества регуляризаторов
🌐 github.com/bigartm/bigartm
 - Разработан и реализован алгоритм параллельного сбора статистики со-встречаемостей пар слов, positive PMI на корпусах неограниченного размера
Обработка полного текста англоязычной Википедии на 8 ядерном процессоре intel core i5 8th gen за 6 часов
 - Отвечаю за парсинг входных данных

ДРУГОЙ ОПЫТ

- **Тест простоты чисел (C++)**
Реализация теста Миллера для детерминированной проверки больших чисел на простоту в рамках курса [Лаборатории Касперского](#) «C++ и проблемы безопасности»
Время работы на простых числах длины 100 примерно 4 секунды
🌐 github.com/MichaelSolotky/sandbox/tree/master/Cpp_old_tasks/Primality_tests
- **Машинное обучение (NumPy, Scipy)**
Реализация различных алгоритмов машинного обучения с нуля
🌐 github.com/MichaelSolotky/sandbox/tree/master/ML

ТЕХНИЧЕСКИЕ НАВЫКИ

- **Языки. Использовались в работе:** C++, Python, C, Bash
- **Языки. Базовые знания:** SQL, Assembly language
- **Технологии:** MapReduce, Protobuf, C++ Boost, CMake, Make, SciPy, Scikit-learn, NumPy, Pandas
- **Библиотеки глубокого обучения:** PyTorch, TensorFlow
- **Инструменты:** Git, Subversion, UNIX/Linux, Travis, AppVeyor, L^AT_EX