

Дегтярев Максим Эдуардович

Контактная информация:

- Email: max.degtyareff1234567890@gmail.com
- Email: prod.soft@live.ru
- Телефон: +7-919-888-07-78
- Город: Москва
- GitHub: <https://github.com/maxdegtyarev?tab=repositories>
- Актуальная версия резюме: <https://github.com/maxdegtyarev/resume>

О себе

Нахожусь в процессе поиска стажировки (или позиции Junior) по направлениям, связанным с машинным обучением и искусственным интеллектом, дата-инженерией. Главная цель - учиться новому и набираться опыта в решении реальных проблем.

Ключевые навыки и умения

МАТЕМАТИКА И АЛГОРИТМЫ

Могу самостоятельно разбираться с формулировками алгоритмов, где требуются знания высшей математики. В свободное время прохожу курсы на Coursera, Stepik. Всегда открыт новому.

SQL & NO-SQL

Писал запросы с использованием PostgreSQL, MongoDB, HBase.

DATA SCIENCE

Sklearn, pandas, numpy, seaborn, XGBoost, SparkML. Начинаю писать диплом в области NLP, а, как следствие, начинаю разбираться с NLP и Deep Learning.

Имею сертификаты Cousera (Data Science):

<https://www.coursera.org/account/accomplishments/records/952YXLK789PY>

<https://www.coursera.org/account/accomplishments/records/E5U7SANTH27Z>

BIG DATA

Прошел курс молодого бойца - Spark (+-medium), Hive (medium), Yarn, ClickHouse

Сертификат NewProlab (Data Engineer стек) :

<http://lk->

[de.newprolab.com/certificate/maxim.degtyarev/bd5f01a705c498b039b796ca4a22b52bdfe66c4ef78bc3c4c738a182202bc4013941418bae70309dba51fabd5576045b650aee16edf09832bf7892eeffb4eade](http://lk-de.newprolab.com/certificate/maxim.degtyarev/bd5f01a705c498b039b796ca4a22b52bdfe66c4ef78bc3c4c738a182202bc4013941418bae70309dba51fabd5576045b650aee16edf09832bf7892eeffb4eade)

ЗНАНИЕ АЛГОРИТМОВ И СТРУКТУР ДАННЫХ

Закончил курс по алгоритмам на Stepik: <https://stepik.org/cert/210041>

ЯЗЫКИ ПРОГРАММИРОВАНИЯ

C/C++; Java; Scala (вместе с фреймворком Spark); Python

Опыт работы

МЕГАФОН | С 14.02.2020 ПО НАСТОЯЩИЙ ДЕНЬ | СТАЖЕР В ОТДЕЛЕ БОЛЬШИХ ДАННЫХ

В данной компании я получил первый опыт работы с большими данными. Освоился с экосистемой Hadoop (в основном, Spark + Hive), основным языком, который использую - Scala + некоторые фишки из Java и Python. Писал и запускал задачи на production кластере. Основной род деятельности - написание пайплайнов и запуск моделей.

За период стажировки сумел в кратчайший срок (2 недели) войти с нуля без знания стека больших данных в достаточно нетиповой проект, самостоятельно осваивал документацию, искал курсы и литературу, что очень помогло.

В процессе работы над проектом, сумел сократить использование дискового пространства в 1,5 раза, также на данный момент являюсь единственным членом команды, который экспериментирует с HBase, пытаюсь внедрить его в наши с ментором проекты. Текущая цель - добиться уменьшения времени выполнения задач, избавившись от операций Join.

Образование

БАКАЛАВРИАТ | В ПРОЦЕССЕ | МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ

В данный момент:

- Курс : 3
- Бюджетная форма обучения
- Специализация: Фундаментальная информатика и информационные технологии

Курсовые работы:

- Разработка калькулятора обыкновенных дробей (26.12.2017)
- Разработка компьютерной игры "The Snake" с искусственными противниками (07.06.2018)
- Поиск кратчайшего пути в графах (26.05.2018)
- Реализация динамических структур данных (18.12.2018)
- Решение теоретических и прикладных задач компьютерной алгебры (04.06.2019)
- Разработка мессенджера с использованием PgSql и Java

СЕМЕСТРОВЫЙ КУРС ПО МАШИННОМУ ОБУЧЕНИЮ ОТ MAIL.RU | ОКОНЧИЛ 23.12.2019 | МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ

В начале 5 семестра поступил по конкурсному отбору (экзамены на знание Python, высшей математики, алгоритмов и структур данных) на данный курс, в процессе прохождения курса решал задачи регрессии, классификации с использованием следующих алгоритмов:

- Линейной и логистической регрессией
- Решающими деревьями и KNN
- Лесами решающих деревьев
- Градиентным бустингом (XGBoost)

Все эти алгоритмы использовал в стандартной реализации sklearn.

Отдельно реализовал метод К-средних для решения задач кластеризации. Также изучал разделы обработки языков (стемминг, лемматизацию, tf-idf, word2vec). В процессе решил задачу поиска самых схожих текстов среди большого объёма статей.

Визуализировал данные в matplotlib, seaborn.

Работал с данными в pandas и numpy.