

Соколов Александр

24.06.1977

Москва

sokaa2011@gmail.com

+7-925-599-1913

<https://www.kaggle.com/sokolovaleks>

<https://github.com/alex-sokolov2011>

DATA SCIENTIST

Перешёл в отрасль Data Science из сферы образования. Нацелен на долгосрочную работу. Качественный разведывательный анализ данных (EDA). Внимание к потенциальным даталикам. Уделяю внимание CV. Хорошая визуализация данных (примеры на kaggle). Дополнител. курсы по SQL.

СТЭК

Python(Base (Pandas, Numpy, tqdm) – ML (SciKit-Learn, LightGBM, Catboost, shap, LAMA, AutoWoE) – DL (TensorFlow, Keras - RL(GYM) - NLP(pymorphy2, NLTK) - NV(albumentations)) – Visualization (Matplotlib, Seaborn) - Parsing(Bs4) – Geo (Shapely) – PROD (streamlit, flask) – DB (pySpark)), DB(Hadoop, Hive, sqoop, SQL, Bash), Linux, Git, heroku, IDE (VS Code, JN, Zeppeline, PyCharm)

проекты и достижения

- соревнование AlfaBattle2.0 - бронзовая медаль. Построил модель кредитного скоринга на данных карточных транзакций клиента. ML(LightGBM, CatBoost), DL(TF, GRU, BIGRU). Защитил дипломный проект по этой теме и дополнительно: использовал отбор признаков с помощью feature и target permutation, сделал сравнительный анализ моделей ML по метрикам f1, ROC PRC и развернул прототип на heroku на оптимальной по качеству и требуемым ресурсам модели. янв.-апр. 2021
- Computer Vision. Построил CNN модель классификации автомобилей по фотографиям для компании занимающейся перепродажей поддержанных авто. TOP10. В проекте использовал: transfer learning на Xception, EfficientNet[B0, B3], оптимизаторы (Adam, Adamax Nadam), callback (EarlyStopping, LearningRateScheduler), fine-tuning, Batch-нормализацию с отключением баеса на предыдущем слое, Test Time Augmentation. Основной стек: tensorflow, keras, albumentations. дек. 2020
- Parsing+ML. Спарсил все объявления с сайта auto.ru Москвы и построил модель предсказания рыночной стоимости автомобиля по его характеристикам для компании занимающейся перепродажей поддержанных авто. TOP10. Определил, что метрика MAPE несостоятельна для оценки эффективности модели для этой задачи и ввел дополнительные метрики. В проекте использовал: методы отбора (correlation matrix, ANOVA F-test, Student t-test, feature_importances_RF + noise), sklearn(LinearReg, RandomForest, GradientBoostingReg, ExtraTreeReg, BaggingReg), blending портфолио пэт-проектов: <https://github.com/alex-sokolov2011/pet-projects> сен. 2020

образование

- колледж Ораним (Израиль) – учитель математики старших классов 2017
- Open University, UK (МИМ «ЛИНК», Новосибирск) - MBA 2007
- Новосибирский Государственный Университет 1999
Механико-математический факультет (средний балл 4,7 + ФМШ)
- SkillFactory, онлайн-курс “Специализация Data Science” (18 месяцев) 2021
- NVIDIA (Image Segmentation with TensorFlow) 2020
- DataCamp (Joining Data in SQL, Introduction to SQL) 2020
- Coursera (ВШЭ (Python), НГУ+2ГИС (Введение в данные, Исследование статистических взаимосвязей) 2019-2020

электронные сертификаты в профиле LinkedIn: <https://www.linkedin.com/in/sokaa2011>

опыт работы

- сеть школ ОПТ Гутман (Нетания, Израиль) - учитель математики старших классов янв. 2017
 - получил академическую степень преподавателя в колледже в Израиле (на иврите) июл. 2019
 - успешно прошел стажировку в школе в течение учебного года (на иврите)
 - получил лицензию учителя математики старших классов
 - наставник команды учеников соревнований по робототехнике "First Lego League"

опыт работы

- директор по продажам (коммерческий директор) в компаниях: май. 2002
- 2ГИС (B2B, B2C, продажа рекламы в городских информационных справочниках с картой города) 2gis.ru дек. 2016
- NPM-Group (Новосибирскпродмаш) производитель оборудования для розлива пива, воды и молока - B2B) npmgrouр.ru
- Парфюм Новосибирск ЗАО (FMCG дистрибьютор продукции Компаний Schwarzkopf, SCA, Калина, Garnier, Loreal, Reckitt Benckiser, SCJohnson, Unilever, Colgate-Palmolive, Johnson&Johnson и других известных ТМ на территории Сибири – Новосибирская, Томская области и Кузбасс)
- Дистрибьюторская Менеджмент Система (ЗАО ДМС) (эксклюзивный FMCG дистрибьютор продукции Компании Проктер энд Гембл (P&G) на территории Западной Сибири – Новосибирская, Томская, Омская, Тюменская области, ХМАО, ЯНАО, Алтай и Кузбасс)
- Coca-Cola HBC Russia (производство и FMCG дистрибуция безалкогольных напитков)