Aleksandr Sokolov

24.06.1977
Antalya, Turkey

sokaa2011@gmail.com
https://t.me/aleks_2011
https://www.kaggle.com/sokolovaleks
https://github.com/alex-sokolov2011

DATA SCIENTIST

I am enthusiastic data scientist with 1+ year of commercial experience in the field and a master's degree in math, transferred from business roles in retail and tech. Many different industrial projects in the role of DS. My previous team noted my qualities beneficial for projects such as strong business expertise, Exploratory Data Analysis (EDA), emphasis on potential data leaks, focus on CV(cross-val), good visualization of data reports, good command of SQL.

stack

Python(Base (Pandas, Numpy, Math, SciPy, 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)), DE(Hadoop, Hive, sqoop, SQL, Bash, pySpark, OpenShift, Jenkins, Sonar Qube), Linux, BitBucket, Git, heroku
IDE: VS Code, Jupyter Notebook, Zeppeline, PyCharm. JIRA, Confluence, Slack.

work experience

- Data Scientist – Data Scientist – Sber. (Moscow, Central Office – Corporate Investment Business - Intelligence control system - Artificial intelligence team) — Apr. 2022 / May. 2021
    - bringing several projects into production (focus on TS, NLP, RecSys), more details about the projects can be obtained during the interview

projects and achievements

- A bronze medal in AlfaBattle2.0 competition. I built a credit scoring model based on the clients' card transaction data (ML (LightGBM, CatBoost), DL (TF, GRU, BIGRU). I successfully defended my thesis project on this topic and additionally used feature selection aided by feature and target permutation, I made a comparative analysis of ML models on f1 and ROC PRC metrics, and deployed a prototype on heroku on the model of optimal quality and required resources. — Jan.-Apr. 2021

- Computer Vision. I built a CNN model of classifying cars by photos for a company engaged in the resale of used cars. TOP10. In the project I used: transfer learning on Xception, EfficientNet[B0, B3], optimizers (Adam, Adamax Nadam), callback (EarlyStopping, LearningRateScheduler), fine-tuning, Batch-normalization with baes off on the previous layer, Test Time Augmentation. Stack: tensorflow, keras, albumentations. — Dec. 2020

- Parsing+ML. I have parsed all the ads from auto.ru website in Moscow and built a model of forecasting a market value for a car based on its characteristics for a company engaged in resail of used cars. TOP10. I determined that MAPE metrics is inconsistent in regard to evaluation of the model performance and introduced additional metrics. I used the following feature selection methods (correlation matrix, ANOVA F-test, Student t-test, feature_importances_RF + noise), sklearn(LinearReg, RandomForest, GradientBoostingReg, ExtraTreeReg, BaggingReg), blending — Sep. 2020

Pet projects portfolio: https://github.com/alex-sokolov2011/pet-projects

education

- Oranim College (Israel) – High-school math teacher — 2017
- Open University, UK (International Institute of Management LINK, Novosibirsk) – MBA — 2007
- Novosibirsk State University, Mathematics and Mechanics Faculty (4.7 grade point average + Physics and Mathematics School) — 1999

- SkillFactory, Data Science Specialization online course (18 months) — 2021
- NVIDIA (Image Segmentation with TensorFlow) — 2020
- DataCamp (Joining Data in SQL, Introduction to SQL) — 2020
- Coursera (Higher School of Economics) (Python), NSU+2GIS (Intro to data science)) — 2019-2020

E-certificates are available in LinkedIn profile: https://www.linkedin.com/in/sokaa2011

work experience

- High-school math teacher in ORT Guttman school network (Netanya, Israel)   Jan. 2017
  • obtained an academic degree in teaching from a college in Israel (in Hebrew)   Jul. 2019
  • successfully completed an internship at the school during academic year (in Hebrew)
  • licensed as a high school math teacher
  • mentor of the students' team in the robotics competition "First Lego League"

- Sales Director (Commercial Director) in the following companies: May. 2002
  • 2GIS (B2B, B2C, advertising sales in city information guides containing city maps)   Dec. 2016
  2gis.ru
  • NPM-Group (Novosibirskprodmash) manufacturer of filling equipment for beer, water and milk - B2B) npmgroup.ru
  • Parfum Novosibirsk CJSC (FMCG distributor of Schwarzkopf, SCA, Kalina, Garnier, Loreal, Reckitt Benckiser, SCJohnson, Unilever, Colgate-Palmolive, Johnson & Johnson and other well-known TM in Siberia – Novosibirsk region, Tomsk regions and Kuzbass)
  • Distribution Management System (DMS CJSC) (exclusive FMCG Distributor of Procter & Gamble (P&G) products in West Siberia – Novosibirsk, Tomsk, Omsk and Tyumen regions, Khanty-Mansiysk Autonomous District, Yamalo-Nenets Autonomous District, Altai and Kuzbass)
  • Coca-Cola HBC Russia (Coca-Cola HBC Russia (production and FMCG distribution of nonalcoholic beverages)