Vladislav Serkov

Moscow, Russia | +79853446858 | vladserkoff@gmail.com
github.com/vladserkoff | linkedin.com/in/vladserkoff

Profile

Data scientist with hads-on experience in machine learning, statistical analysis, data visualization and big data. On one hand, I am capable of delivering results relevant to current business needs. And on the other hand, interested in a research for new data- and ML-driven products. My favorite tool for day-to-day work is Python, but I am also familiar with R.

Work experience

Data scientist at DCA: October 2017 - Present

DCA is a programmatic advertising platform and is one of the largest DMPs in Russia, seeing data from millions of sites every day. Here I build end-to-end systems that help to make sense of the data in unsupervised way: **1.** Automatic segment creation based on target keywords; **2.** Web pages categorization; **3.** Audience look-alike segmentation; **4.** Banner CTR prediction.

Data analyst at R&D at Flocktory: November 2014 - September 2017

Designed and built data tools and products using Flocktory's data on activity of visitors of the majority of Russia's largest e-commerce sites. Flocktory sees tens of GBs of data per day. This causes major challenges in extracting value and insights from it, thus a large variety of tools is used in the process. Tools and products include (but not limited to): **1.** Forecasting customers' future purchasing behavior; **2.** Matching product catalogs (taxonomies) of different shops onto a universal taxonomy; **3.** Identifying patterns in goods for promotions, e.g. the products that are purchased repeatedly; **4.** Building and maintaining dashboards and the data for the dashboards both for internal and external use.

Pricing analyst at Mamsy: March 2013 - November 2014

Mamsy is a e-commerce retail based on flash-sales model. Every day there's a different assortment of products and prices are adjusted every time a product enters the assortment. In Mamsy I've decided on buying and selling prices based on expected costs, analyzed profitability of certain goods and brands, monitored and analyzed competitors' pricing.

Technical skills and programming languages

- **Statistical skills**: regression analysis, hypothesis testing, data visualization, association rules
- **Machine learning**: data cleaning and preparation, feature engineering, regression, classification, cross-validation
- **Databases**: Impala, Hive, HBase, MongoDB, PostgreSQL, Vertica etc.

- **Python**: data munging with **pandas**, machine learning with **scikit-learn**, **XGBoost** and **LightGBM**, web scraping with **requests**, plotting with **matplotlib** and **seaborn**.
- **R**: data munging with **data.table** and **dplyr** , dashboards with **shiny**, plotting with **ggplot2**, association rules with **arules**.
- **Other tools**: **SQL** for structured data, **S3** and **kafka** for raw and stream data, parallel processing with **dask** library, **Linux** cli, cloud computing with **AWS** and **Azure**, version control with **git**, deploying to production with **Docker**.

Education

National Research University Higher School of Economics

Economist, with World Economy as the main field of study, Specialist: September 2008 - August 2013