Konstantin Burkin

konstantin.burkin.m@gmail.com | kburkin | kozersky | konstantinburkin.github.io

SUMMARY: Data Scientist in tech group of the largest fastfood chain.
3+ years of production experience.
Red diploma alumnus of Lomonosov Moscow State University.

WORK EXPERIENCE

- **Middle Data Scientist** *2023 – Present* • **Technologii i tochka**
  - Project: NLP of reviews
  - Stack: Python, ruBERT, Yandex-GPT
    - Automated suggestions for app improvement based on scrapped reviews from AppStore and PlayMarket.
    - Preselected informative reviews for each function of app and aggregated them using Yandex-GPT.

  - Project: New products sales
  - Stack: Python, PyTorch, CatBoost, Airflow, MLflow
    - Developed graph neural network model to predict sales of new products with accuracy of 50 WAPE.
    - Improved WAPE by 22% using CatBoost

  - Project: Restaurants order proposal
  - Stack: Python, PyTorch, SQL, CatBoost, SARIMA, Airflow, MLflow, Git, DVC
    - Feature engineering and tech support for models predicting regular and promo products.
    - Developed baseline model for predictions of unpopular products

- **ML researcher** *2021 – 2023* • **Higher School of Economics University**
  - Project: Prediction of outcomes for cardiovascular patients based on clinical data.
  - Stack: Python, Git, Bash, scikit-learn, imblearn, MLxtend, CatBoost, Pandas, NumPy
    - Demonstrated biomarkers' predictive capabilities (up to 5% AUROC increase).
    - Determined 2 primary predictors by feature selection algorithms: SHAP and FFS, retaining F_2>0.6 and AUROC>0.8.
    - Published results at the conference and earned grant support of federal academic program.

  Intern

    - Medical data preprocessing, imputation, and analysis; ML model training for combined target prediction.
    - Increased Recall by 9% by tuning models with respect to F_2 metric.
    - Presented project results at the conference.

EDUCATION

- **PhD program in Machine Learning** *2023 – Present* • **Higher School of Economics University**
  - Project: Parsing unstructured medical data

- **Program: Neural networks in research** *2022 – 2023* • **Lomonosov Moscow State University**
  - Scholarship: top-25 based on ML competition and academic results
  - Stack: Python, PyTorch, scikit-learn, MLxtend

- **BSc & MSc in Fundamental and Applied Chemistry** *2017 – 2023* • **Lomonosov Moscow State University**
  - GPA: 4.97/5, Red diploma
  - Academic council Scholarship: top-10 MSU students for scientific achievements

INDIVIDUAL PROJECTS

- **Delivery Club Sales Prediction**
  - Compared effectiveness CatBoost and LGBM models for weekly sales prognosis.
  - Compared SARIMA and LSTM for univariate time series. SARIMA outperformed LSTM in accuracy and training time.
  - Stack: Python, scikit-learn, CatBoost, Pandas, PyTorch, Pandas, NumPy, Plotly