Alexey Za
Senior Senior ML/AI Engineer

SUMMARY

Data Scientist/Machine Learning Engineer with over 5.5 years of experience specializing in predictive modeling, analytics, computer vision, and offering a robust background in computer science and software engineering. Technical proficiency in Python and a range of data science tools including Numpy, Pandas, and Machine Learning libraries such as Scikit-learn. Expert in deploying ML models to cloud environments like Azure and AWS. Experience includes implementing custom ML solutions for fraud detection in banking, credit risk analysis in finance, energy consumption forecasting, and telecom subscriber insights. Strong track record in data engineering with Apache Airflow and Spark, and backend development with FastAPI. Practical DevOps and MLOps integration expertise using Docker, Jenkins, and MLflow, and experience with databases like PostgreSQL and Redis. Recognized for enhancing model performance through feature engineering, optimizing hyperparameters, and ensuring model explainability.

TECHNICAL SKILLS

Main Technical Skills	Python, NumPy, Pandas, Matplotlib, Seaborn
Programming Languages	Python
Python Libraries and Tools	Matplotlib, NumPy, Pandas, Seaborn
AI & Machine Learning	NumPy
Data Analysis and Visualization Technologies	Pandas

WORK EXPERIENCE

Machine Learning engineer, ML Banking Transaction Analysis Platform

Duration: 11.2023 – Present

Summary: ML platform for banking operations, processing large volumes of transaction data, applying fraud detection, anomaly detection, and client segmentation to improve customer experience.

Responsibilities: Defining business requirements, data collection and processing, developing and optimizing machine learning models, addressing imbalanced datasets, implementing algorithms for anomaly detection, performing client segmentation, designing dashboards, integrating interpretability layers with SHAP, ensuring model integration with real-time transaction systems, monitoring production model behavior.

Technologies: Python, Pytorch, Pandas, Numpy, LightGBM, CatBoost, Scikit-learn, Statsmodels, SHAP, Optuna, MLflow, Seaborn, Kafka, Apache Spark, Azure ecosystem, Redis, Docker, MongoDB, Git, GitLab

Data Scientist / Machine Learning Engineer, Credit Platform Analytics and Decision Engine

Duration: 10.2022 – 11.2023

Summary: Finance management and credit analytics platform for the construction industry, enhancing credit score predictions and solving optimization problems for marketing.

Responsibilities: Collecting business requirements, experimenting with statistical and ML approaches, developing clusterization models, setting up MLflow for tracking, developing an analytical layer on Azure Synapse, initiating data and ML practices, setting up a DWH, creating PowerBI dashboards, and setting up inference pipelines.

Technologies: Python, Numpy, Pandas, SciPy, Plotly, Matplotlib, XGBoost, SHAP, Prophet, Tensorflow, Keras, Azure, MongoDB, PostgreSQL, SQL Server, PowerBI, MLflow, Docker, Git, Bitbucket

Data Scientist / Machine Learning Engineer, Energy Consumption Forecasting

Duration: 08.2021 – 10.2022

Summary: Energy consumption forecasting for home devices to optimize bills and CO2 savings.

Responsibilities: Setting up model metrics monitoring, creating data preprocessing pipelines, applying various models for prediction, enhancing production models, conducting A/B testing, exploring data characteristics, handling data and concept drifts, creating stakeholder reports, using MLflow for tracking, and running Apache Airflow DAG.

Technologies: Python, Pandas, Scikit-learn, Optuna, PyTorch, SciPy, Statsmodels, Seaborn, Plotly, SHAP, Prophet, FastAPI, XGBoost, LightGBM, PostgreSQL, Apache Airflow, MLflow, Docker, Git, GitLab

Data Scientist / Machine Learning Engineer, Telecom Subscriber Insights

Duration: 04.2020 – 08.2021

Summary: Churn prediction and customer segmentation for telecom subscribers to inform retention strategies and improve customer satisfaction.

Responsibilities: Monitoring model metrics, creating preprocessing pipelines, developing segmentation layers, creating model ensembles, experimenting with data balancing techniques, adding explainability, conducting EDA, handling data drift, reporting to stakeholders, deploying models on Azure ML, implementing CI/CD for model updates.

Technologies: Python, Pandas, Scikit-learn, Hyperopt, SciPy, Statsmodels, Seaborn, SMOTE, Plotly, SHAP, CatBoost, XGBoost, LightGBM, Spark, Azure ecosystem, Oracle, Jenkins, Docker, Git, GitLab

Data Scientist / Machine Learning Engineer, Anomaly Detection on Passport Photo

Duration: 07.2019 – 04.2020

Summary: Created a service for anomaly detection in passport photos and matching faces from passports with selfies, optimizing AWS resources during model training.

Responsibilities: Collecting and annotating data, developing preprocessing and augmentation pipelines, experimenting with anomaly detection models, setting up AWS infrastructure for model training and monitoring, optimizing training costs, deploying models, and supporting model testing and pre-production phase.

Technologies: Python, PyTorch, TorchVision, Tensorboard, OpenCV, Pillow, Matplotlib, AWS ecosystem, MongoDB, Redis, Docker, Git, GitLab

EDUCATION

- **Computer Science and Software Engineering**