

Sergazy Nurbavliyev, Ph.D.

◇ sergazy.nurbavliyev@gmail.com ◇ (385)-389-7833

◇ linkedin.com/in/sergazy ◇ github.com/sernur

◇ Salt Lake City, UT

SUMMARY

Senior Machine Learning Scientist with over **5 years of experience** in delivering high-impact solutions in **Machine Learning (ML)**, **Deep Learning (DL)**, **Reinforcement Learning (RL)**, and **Large Language Models (LLMs)**. Proven expertise in architecting scalable ML models, optimizing pricing strategies, and enhancing customer segmentation. Adept at mentoring teams, collaborating on cross-functional projects, and deploying AI/ML models in production environments. Committed to advancing AI applications and driving organizational success.

SKILLS

- **Programming Languages:** Python, R, SQL
- **ML Frameworks:** PyTorch, TensorFlow, Keras
- **ML Algorithms:** Reinforcement Learning, Deep Learning, Transformers, CNN, Tree based algorithms
- **Natural Language Processing:** LLMs, Transformer Models (BERT, GPT), Multimodal Embeddings
- **Cloud Platforms:** Google Cloud Platform (GCP), Azure
- **MLOps/DevOps Tools:** Docker, Jenkins, Airflow, MLflow, Kubeflow
- **Statistical Methods:** Bayesian and frequentist analysis, A/B testing, causal modeling
- **Optimization Techniques:** Linear programming, high-dimensional optimization, parallel computing
- **Databases:** MySQL, PostgreSQL, MongoDB
- **Languages:** English, Kazakh (Native), Russian, Turkish, Turkmen

PROFESSIONAL EXPERIENCE

Beyond (formerly Overstock.com)

Feb 2023 – Present

Senior Machine Learning Scientist

Salt Lake City, UT

Sponsored Ads Products Optimization (SPA):

- Developed and deployed advanced **Click-Through Rate (CTR)** and **Conversion Rate (CVR)** models using XGBoost.
- Optimized the SPA ranking algorithm and increased ad click predictions by **15%** and conversion rates by **5%**, boosting sponsored ad revenue by **\$5 million annually**.

Machine Learning Coupon Personalization (MLCP):

- Led cross-functional projects to develop personalized email coupon algorithm using causal ML models.
- Scaled the model to serve **40 million customers**, resulting in a **3.6% increase in revenue** and a **15% increase in profit**, confirmed by A/B testing on **20 million customers**.

Dynamic ML Pricing Optimization (MLPO):

- Directed the development of a pricing optimization algorithm processing over **2 million SKUs**.
- Oversaw the creation of a predictive model and a parallel-based optimization method processing over **2 million SKUs**, leading to improvements in key financial metrics, including a **\$13.5 M** revenue increase.

Team Mentoring and Hiring:

- Participated in the hiring process, interviewing and onboarding new data science team members.
- Mentored an intern in developing Bayesian models for A/B testing targeting skewed revenue-based KPIs, resulting in a 20% to 40% reduction in sample size requirements and significant experiment cost savings.

Automated Site Sale Optimization:

- Developed and deployed a machine learning algorithm using deep neural networks and transformer-based models to predict item quantities sold for millions of products.
- Created optimization methods to maximize revenue and profit under budget constraints, integrating advanced optimization techniques and parallel processing.
- Improved prediction accuracy by **10%** based on MAPE and MAE, enhancing the effectiveness of sales forecasts, reduced manual processing time by **80%** through full automation
- Increased annual revenue by **6% (\$150 million)** and profit by **4% (\$10 million)** through optimized discount allocations, as demonstrated by A/B testing.
- Wrote production code and deployed models using Docker, Jenkins, and Airflow on local servers and Google Cloud Platform (GCP).
- Participated in A/B testing for new algorithms, including test design, method development, sample splitting, metric selection, power calculation, and results analysis.

Customer Lifetime Value (CLV) Estimation:

- Developed and implemented CLV estimation model combining churn probability prediction with future spend forecasting using gradient boosting, improving accuracy isotonic regression.
- Enabled the business to optimize marketing strategies by identifying high-value customers, resulting in more precise marketing budget allocations, increased retention, and sales.

Probabilistic SKU Selection:

- Developed a probabilistic model for SKU selection that optimized pricing and discounts by integrating business requirements, improving transparency and flexibility, and reducing processing costs by **20%** .

Product Deduplication Detection:

- Implemented a deduplication model analyzing over **3 million SKUs** using image hashing, decreasing duplicate listings by **95%** and improving operational efficiency by **30%**.

EDUCATION

- **Ph.D. in Mathematics**, University of Utah, Salt Lake City, UT
 - **M.Sc. in Statistics**, University of Utah, Salt Lake City, UT
- Dec 2020
Dec 2019

PROJECTS

- Multimodal Retrieval-Augmented Generation System2024
- Designed and developed a system integrating text and video data using multimodal embeddings from **Large Language Models (LLMs)** like BERT and GPT, and **Large Vision-Language Models (LVLMs)** such as CLIP.
 - Collected and preprocessed a dataset of synchronized text and video content, implementing automated video preprocessing pipelines using OpenCV and FFmpeg, which reduced preprocessing time by **40%**.
 - Extracted and aligned text and video embeddings into a shared vector space, enabling cross-modal retrieval with enhanced accuracy.
 - Implemented an efficient retrieval system using vector databases like **Faiss**, improving retrieval accuracy from vector stores by **30%**.
 - Deployed the system using Docker, created RESTful APIs for interaction, and ensured scalability and performance through optimization and parallel processing techniques.