Svetlana Vinogradova, PhD

(857) 264-8057 | kintany@gmail.com | linkedin.com/in/svetlana-phd

SUMMARY

Data Scientist with extensive experience in developing and deploying ML models across academia and industry. Expert in recommendation systems, causal inference methods, and experimentation, with a strong emphasis on writing production-ready code. Proficient in classical machine learning techniques and cutting-edge Generative AI.

EXPERIENCE

Senior Data Scientist, ML Algorithms
July 2024 – Present

Nuna Health
San Francisco, CA

- Leading the development of the new ML-based trend algorithms and surfacing insights for users and healthcare providers

Lead Data Scientist
September 2022 – June 2024

Zwift
Long Beach, CA / remote

- Lead the development and deployment of customer-facing product features, including introducing the racing score as part of a company-wide initiative for a **personalized recommendation system**. Conducted initial research, statistical modeling, and implemented the machine learning model in production, resulting in significantly enhanced user engagement. Collaborated with engineers and product teams, leading the data science aspect of the project.
- Developed a machine learning model for customer segmentation, leveraging behavior data from 1 million users on a virtual cycling gaming platform, aiding the business in expanding its recreational cycling enthusiast audience.
- Enhanced data integrity across multiple data streams by initiating and establishing a comprehensive data quality and governance review process, covering data validation, anomaly detection, and consistency checks

Lead Data Scientist / Data Science Manager
August 2019 – September 2022

InsideTracker
Cambridge, MA

- As a founding data scientist, led a team of 3 data scientists setting the strategic direction for the team and working cross-functionally and collaborating closely with company leadership to develop and deliver metrics, attribution models, and actionable business analyses
- Drove end-to-end data science projects, creating machine learning models with features based on blood biomarkers, genomic, wearable, and customer behavior data. Delivered algorithms vital to the company's flagship product InnerAge2.0. Wrote production-ready code in Python and deployed models to production.

Research Scientist
August 2016 – August 2019

Harvard Medical School / Dana Farber Cancer Institute
Boston, MA

- Developed novel machine learning methods to study gene expression and epigenetic signatures, introducing a computational approach, Qllelic, that accurately accounts for technical noise and established a new gold standard for analyzing allelic imbalance
- Co-authored an R package for differential allele-specific analysis and co-authored 5 peer-reviewed articles

Research Scientist
August 2010 – July 2016

Institute for Information Transmission Problems

- Developed a machine learning approach to incorporate RNA probing data into RNA secondary structure prediction and genome-wide search for structured RNAs
- Published 3 papers in peer-reviewed journals

EDUCATION

Lomonosov Moscow State University
2010 – 2016

Ph.D. in Mathematical Biology and Bioinformatics

Lomonosov Moscow State University
2005 – 2010

Bachelor and Master of Science in Computational Biology

CERTIFICATIONS

Deep Learning Specialization (5 courses)
2018

DeepLearning.AI

Technical Skills

Languages: Python (pandas, scikit-learn), PySpark, R (tidyverse, Quarto, Shiny apps), SQL, Java
Deep Learning and Generative AI:: TensorFlow, PyTorch, Hugging Face Transformers, Large Language Models
Developer and Engineering Tools: Git, Docker, Databricks, AWS, Google BigQuery, VertexAI