

SARVAR KHAMIDOV

Brooklyn, NY, 11221 | (617)529-9890 | sk10583@nyu.edu | [linkedin.com/in/sarvarkhamidov/](https://www.linkedin.com/in/sarvarkhamidov/)

EDUCATION

Master of Science in Biostatistics (Cumulative GPA: 4.0) New York University, School of GPH	Expected May 2026 New York, NY
Bachelor of Science in Data Science , Minor in Information Sciences and Technology Pennsylvania State University	May 2023 University Park, PA

RELEVANT WORK EXPERIENCE

<i>AI Engineer Co-op, AbbVie</i> (Florham Park, NJ)	Aug 2025 – Present
<ul style="list-style-type: none">Engineered an LLM powered pipeline to automate SAP generating, cutting drafting time and reducing manual errorBuilt a vector database to enable RAG, improving contextual accuracy of SAP drafts and aligning outputs with regulatory standards	
<i>ML Research Assistant, Feng Lab</i> (New York, NY)	Jan 2024 – Present
<ul style="list-style-type: none">Evaluated and benchmarked lightweight CNNs and ResNet architectures against large-scale multimodal models (CLIP, Qwen-VL) to identify optimal trade-offs between accuracy, efficiency, and computational cost on the TinyImageNet dataset.Fine-tuned CLIP's ViT-B/32 image encoder by unfreezing the final transformer blocks, achieving a +10% improvement in top-1 accuracy over zero-shot baselines.	
<i>Programmer Intern, Regeneron</i> (Warren, NJ)	May 2025 – Aug 2025
<ul style="list-style-type: none">Developed a diagnostic software to support quality checks across 10+ Regeneron's Medidata EDC clinical studies.Optimized compliance report software boosting efficiency across 15+ database builds.Implemented validation scripts ensuring data integrity prior to deployment (data analysis / database lock).	
<i>Data Scientist, EDA Clinical</i> (Pittsburgh, PA)	Jun 2023 – Aug 2024
<ul style="list-style-type: none">Developed survival analysis models to predict patient outcome by integrating EHR data for regulatory submissions.Automated CDISC-compliant data pipelines (SDTM/ADaM) in R, and SAS, cutting processing time by 40%.Analyzed clinical and real-world data, improving treatment efficacy by 15% in key patient groups.	
<i>Software Engineer, Penn State University</i> (University Park, PA)	Nov 2022 – Jun 2023
<ul style="list-style-type: none">Deployed scalable applications (educational platforms) supporting around 82,000+ students.Integrated data analysis techniques to optimize performance, reducing processing time.Designed a data pipeline and dashboard using LDAP database for university's budget planning.	
<i>Quality Assurance Intern, Numerix LLC</i> (New York City, NY)	May 2022 – Aug 2022
<ul style="list-style-type: none">Developed and implemented 60+ automated tests scripts using Selenium, improving testing efficiency by 30%.Identified and executed manual and automated test cases to validate functionality, performance of applications.Standardized QA documentation and reporting practices, enhancing communication between QA and Dev teams.	

OTHER WORK EXPERIENCE

<i>Biostatistics Research Assistant, Mount Sinai Hospital</i> (New York, NY)	Oct 2024 – May 2025
<ul style="list-style-type: none">Applied survival analysis and machine learning to clinical research study on lung transplant mortality.Developed R Shiny App for interactive visualization of summary of patient outcomes.Delivered statistical consulting services for clinical research projects.	

SKILLS

- Software: Docker | R | **SAS** | HPC | CUDA | **PyTorch** | LangChain | LangGraph | Chroma
- Programming Languages: **Python** | SQL | JavaScript | React | Git
- Certifications & Training: CDISC | Clinical SAS | AWS |

PERSONAL PROJECTS / COMPETITIONS

<i>Smart Health Tracker App (Personal Project)</i>	May 2025 – Present
<ul style="list-style-type: none">Developed a mobile application (React Native + Node.js) integrating wearable device APIs (Garmin, Apple Health) to monitor health metrics.Implemented predictive analytics using Python to detect anomalies in heart rate and sleep patterns.	
<i>ENAR DataFest: NHANES Hypertension Analysis (Team Project)</i>	Feb 2025
<ul style="list-style-type: none">Analyzed 39K+ NHANES records using multivariable logistic variable and Random Forests.	

- Achieved AUC 0.79 and identified depression as a key predictor of uncontrolled hypertension despite medication.