Angela A. Boakye Danquah

Woodbridge, VA | angela.bdanquah@gmail.com | 571.431.9681 | [LinkedIn](#) | [GitHub](#) | [Website](#)

SUMMARY

Data Scientist with expertise in large-scale data management, statistical modeling, and Python-based analytics. Proven track record of processing terabytes of data, building optimized databases and analytical pipelines, and developing reproducible workflows on high-performance computing systems and cloud infrastructure. Skilled in applying statistical models and quantitative methods to complex problems, creating interactive visualizations and dashboards, and communicating analytical findings clearly to both technical and non-technical stakeholders.

EXPERIENCE

Computational Research Scientist   *Dec 2023 - Present*

Data Analytics Center |Research Computing | ITS

University of Virginia | Charlottesville, VA

- Partner with faculty to optimize data pipelines processing billions of records across TB-scale research datasets on high-performance computing infrastructure.
- Build and maintain databases with optimized indexing, querying, and aggregation pipelines, improving scalability and query performance for research analytics workloads.
- Create interactive dashboards and publication-quality visualizations using Plotly, Tableau, and Python libraries to communicate complex statistical analysis results to diverse stakeholders including researchers, administrators, and funding agencies.
- Develop efficient, maintainable, and reusable Python code implementing modular design patterns and comprehensive documentation for data processing, analysis, and team collaboration.
- Troubleshoot, debug, and enhance existing data processing tools and workflows, including multi-GPU training and distributed systems optimization.
- Apply strong quantitative and problem-solving skills using statistical models to analyze complex patterns in time-series and longitudinal data.
- Architect secure, cloud-based data systems and develop analytical solutions tailored to client requirements for highly sensitive data.

SELECTED TECHNICAL PROJECTS

- **BlueSky Social Network Analysis** — Built ~7B-record (~3 TB) NoSQL database; engineered sharding, indexing & aggregation utilities for high-volume network science research
- **HIPAA-Compliant AWS Application** — Architected secure cloud platform for rare cancer research lab to extract PHI from diverse file formats; designed VPC isolation with dual-VPN connectivity (UVA Health VPN + site-to-site); built React frontend and serverless backend (Lambda, S3, DynamoDB); integrated Textract and Comprehend Medical for AI-powered document processing; coordinated with IT, HIT compliance, and IRB for security adherence
- **ECG Heart Classification** — Adapted foundation transformer models + custom CNNs + LSTM (ResNet/VGG) for 1-D ECGs; trained on 300K ECGs achieving 0.89 specificity / 0.91 sensitivity

EDUCATION

Master of Science in Data Science | University of Virginia | Charlottesville, VA   May 2023
Bachelor of Arts in Statistics & Economics | University of Virginia | Charlottesville, VA   May 2021

TECHNICAL SKILLS

Programming & Analysis: Python, SQL, R, Bash, Git

Machine Learning & AI: PyTorch, TensorFlow, Scikit-learn, JAX, Hugging Face, NLP, Computer Vison

Infrastructure: High-performance computing (SLURM), AWS/Azure, Docker, Git

Data Engineering & Cloud: AWS, Azure, Containerization (Docker/Apptainer), Apache Spark, Hadoop, Dask, Airflow, MapReduce, Relational Databases, NoSQL Databases, Distributed Computing

Tools & Other: Tableau, CI/CD, data visualization, dashboard creation, ETL pipelines

LEADERSHIP & PROFESSIONAL ENGAGEMENT

- Organizer of Journal Club at UVA Research Computing, facilitating technical discussions on data science methodologies
- Contributor to research publications, grant proposals, and collaborative projects across multiple disciplines
- Professional memberships: Society for Industrial and Applied Mathematics (SIAM), American Statistical Association (ASA), Women in HPC, Virginia (WHPCVA)