Joslin Ishimwe

531-739-2469 | ishijoslin@gmail.com | linkedin.com/in/joslin-ishimwe/ | github.com/ishjosl

Education

University of Nebraska-Lincoln — Lincoln, NE
Master of Science in Computer Science, Bioinformatics — *Aug. 2024 – May 2026*

University of Nebraska-Lincoln — Lincoln, NE
Bachelor of Science with Distinction — *Aug. 2019 – May 2023*

Experience

Machine Learning Researcher — Aug. 2024 – Present
University of Nebraska-Lincoln — *Lincoln, NE*

- Accomplished 60% reduction in TEM analysis cost (saving 200 research-hours/month) by building PyTorch CNN that boosted vesicle detection accuracy to 40% (vs 24% baseline
- Scaled genomic ETL to 10TB/day on AWS EMR (20 nodes), enabling 35% faster insights for 15 cross-validation teams
- Discovered 3 patent-pending protein biomarkers via clustering models to be published in Bioinformatics Journal

Research Data Scientist — Jun. 2023 – Jul. 2024
Elemental Enzymes (Corteva Partner) — *St. Louis, MO*

- Achieved 100% FDA audit compliance for agrochemical launches by automating 200+ test cases with Scikit-learn, cutting release cycle from 6 weeks to 3 days.
- Generated $120k client savings via Matplotlib dashboards that reduced stakeholder decision latency from 5 days to 4 hours.
- Collaborated with regulatory teams and external partners to design Part 11-compliant ML workflows.

Software Engineer Intern — May 2024 – Aug. 2024
Sand Technologies (AWS Partner) — *St. Louis, MO*

- Optimized AI/IoT workflows with Docker/Kubernetes microservices, reducing cloud spend by $85K annually.
- Delivered full-stack Spring Boot/React platform serving 50K+ users with 95% unit tests, coverage, resulting in zero critical incidents over 3 months.
- Automated CI/CD pipelines with Jenkins, integrating Selenium/Robot Framework to cut manual QA by 40%.

Undergraduate Data Scientist — Jun. 2021 – May 2023
University of Nebraska-Lincoln — *Lincoln, NE*

- Designed PySpark ML model predicting plant growth under stress with 50% higher accuracy vs. baseline, achieving statistical significance (p¡0.01)
- Created NLP tool extracting insights from plant science literature, improving literature review efficiency by 40%.

Projects

FederatedGenomics | *PyTorch, Spark, TF Privacy* — June 2025 – Present

- Cut Federated Learning comms cost by 40% via PruneFL while maintaining 95% accuracy on TCGA Pan-Cancer data and reducing data leakage risk by 60% with $\epsilon = 0.3$ differential privacy (Tensorflow privacy)

DataMover | *Spigot API, Java, Maven, TravisCI, Git* — May 2023 – Dec. 2023

- AI-driven schema mapping tool for SQL/NoSQL migrations, increasing efficiency by 30%.

Technical Skills

Languages: Java, Python, C/C++, R , SQL (Postgres)
ML Frameworks: PyTorch, Tensorflow, Scikit-learn, Distributed Systems Training (PySpark MLlib
Data Science: Statistical inference, Hypothesis Testing, A/B Testing, FDA 21 CFR Part 11, Experimental Design
Cloud: AWS(SageMaker, Glue, EMR), Docker, Kubernetes, Serverless (Lambda).
Other: Git, Jenkins, REST APIs, Airflow, JavaScript, React, Visualizations (seaborn, Matplotlib, Plotly).
Courses: Machine learning, DSA, Database, Design and Analysis Algorithm, Advanced statistics, Data and Network Security.