Shiuli Subhra Ghosh

Engineer, Data Scientist, Energy Systems Analyst

Richmond, USA | linkedin.com/in/shiuli | ghoshshiulisubhra16@gmail.com | EU Blue Card Eligible

Professional Summary

Engineer, Data Scientist, and Energy Systems Analyst with 5+ years of experience in machine learning, causal inference, statistical modeling, forecasting, and power transmission analytics. Experienced building AI-driven forecasting, ETL, dashboard, and causal inference solutions for energy systems and markets.

Education

Rensselaer Polytechnic Institute — **Troy, NY**
Master of Science in Electrical Engineering - GPA: 3.82/4 — *Dec 2024*

Chennai Mathematical Institute — **Chennai, India**
Master of Science in Data Science - GPA: 9.19/10 — *Aug 2022*

National Institute of Technology, Durgapur — **Durgapur, India**
Bachelor of Technology in Electrical Engineering - GPA: 8.62/10 — *June 2018*

Professional Experience

Dominion Energy Virginia — **February 2025 - Present**
Engineer — *Richmond, VA*

- Designed an enterprise hierarchical forecasting platform covering 300+ transmission assets, improving MAPE by 20% and supporting outage planning, renewable integration, and data center interconnection flexibility using Python, Databricks, SARIMA, Prophet, LSTM, and Transformers.
- Developed probabilistic risk frameworks supporting 50+ monthly outage scenarios, reducing decision turnaround by 30% and post-contingency load relief warnings by 10% through PSSE and TARA outage reliability analysis.
- Clustered 300+ AI data center load profiles to support demand estimates for 50+ new substations, reducing planning uncertainty by 25% and improving interconnection study quality and infrastructure prioritization through Gaussian Processes and Python time-series methods.
- Integrated PI Web API, WBIS, and Snowflake into an on-prem SQL Server database, improving data accessibility by 50% across 6.1M daily telemetry records and reducing data modeling interpolation errors through automated ETL pipelines.
- Created a Gen-AI Agent using Claude Opus 4.7 and domain-specific training data, automating 40% of manual data entry and enabling instant querying of on-prem databases for transmission expansion projects.
- Deployed Azure DevOps and Power BI dashboards tracking 10+ customized forecast and congestion KPIs for 5+ stakeholders, reducing communication redundancy by 50% and improving planning and operations visibility through CI/CD and dashboard automation.

Dominion Energy Virginia — **May 2024 - August 2024**
Summer Research Intern — *Richmond, VA*

- Built version-controlled Python ML pipelines for hierarchical load forecasting, reducing model deployment time by 20%, improving MAPE by 15%, and strengthening governance across 3 hierarchy levels and 4 regions through MLflow and Grafana dashboards.
- Implemented MinT, Hybrid-MinT, and Bottom-up forecast reconciliation algorithms, reducing hierarchical forecast incoherency by 100% across 3 transmission levels and improving system-level consistency for operational planning and outage studies.

Information Systems Group, Rensselaer Polytechnic Institute — **August 2022 – December 2024**
Graduate Research Assistant — *Troy, NY*

- Applied Cyclic-LiNGAM, C-PATH, and CCI causal inference methods to analyze 1,000+ cascading failure scenarios, improving failure-risk prediction by 28% and strengthening grid resilience research and transmission risk assessment.
- Developed an unsupervised critical cascade identification framework using PyWhy causal discovery algorithms, achieving 45% regret improvement versus GNN and influence graph baselines for interpretable power transmission risk analysis.

Imec Belgium — **January 2022 - June 2022**
Research Intern — *Leuven, Belgium*

- Implemented YoloV5, RetinaNet, EfficientNet, Noise2Noise, and PyTorch for CD-SEM defect classification, achieving 95% accuracy across 10,000+ images and 10+ defect classes while improving semiconductor inspection reliability and speed.

INESC TEC — **September 2021 - December 2021**
Research Intern — *Porto, Portugal*

- Designed an Online Learning framework for the Energy Data Market, improving allocation efficiency by 25% across 30 simulated market participants and strengthening incentive fairness, data sharing, and collective decision-making through adaptive gradient-based Python algorithms.
- Implemented scikit-learn quantile regression for energy demand and renewable generation forecasting, reducing pinball loss by 20% across 10+ scenarios and improving market pricing, dispatch planning, and incentive design.

Legato Health Technologies — **July 2021 - September 2021**
Summer Research Intern — *Bangalore, India*

- Developed the Anthem AutoML framework from scratch, evaluating 30+ models across 10 customer databases and supporting healthcare AI model development workflows through automated machine learning framework design.
- Developed an automated model bias detection prototype for Anthem, improving fairness validation by 30% across 30+ models and 10 customer databases while supporting healthcare compliance, audits, and fairness monitoring through statistical hypothesis testing.

Jindal Stainless Limited — **July 2018 - January 2021**
Associate Manager — *Odisha, India*

- Developed a time-series forecasting model for transformer relay data, improving MTTR and MTBF metrics by 20% across 200+ MV transformers and switchgears, enhancing maintenance planning and reducing downtime risk.
- Developed a predictive manpower allocation model, reducing overtime costs by 12% annually and delivering 0.5M cost reduction across 3 cold rolling units while improving workforce utilization and labor planning.
- Developed grade-wise power consumption prediction models achieving 95% accuracy across 15+ stainless steel grades, supporting energy cost optimization and CPP/internal microgrid production planning through scikit-learn and Pandas/NumPy pipelines.

Notable Achievements

- **Publications & Patent:** Author of **1 US Patent** and **10+ peer-reviewed IEEE conference and journal papers** in power systems, causal inference, and machine learning. Full list available on Google Scholar.
- **Peer Reviewer:** Served as a technical reviewer for **IEEE TENCON 2020, IEEE T&D**, and **General Meeting 2026**.
- **Dominion Energy Diversity Scholarship (2024):** Awarded for outstanding academic performance and demonstrated commitment to diversity, equity, and inclusion during the summer internship program.
- **J. Baliga Award:** Awarded for exceptional GPA, academic background, and faculty nominations in the first year of the Ph.D. program at Rensselaer Polytechnic Institute.

Technical Skills

Programming and Analytics: Python, MATLAB, R, SQL, Julia, Statistical Modeling, Machine Learning, Causal Inference, Time-Series Forecasting, Probabilistic Forecasting, Quantile Regression, Online Learning

Machine Learning and Python Libraries: PyTorch, Pandas, NumPy, scikit-learn, TimeGPT, StatsForecast, PyWhy, SARIMA, Prophet, LSTM, Transformers, Gaussian Processes

Energy Systems and Power Transmission: PSSE, PLEXOS, TARA Outage Reliability Analysis, PI System, PI Web API, Power Transmission Planning, Load Forecasting, Renewable Generation Forecasting, Data Center Load Forecasting, Congestion Risk Analysis

Data Engineering and Databases: Databricks, PySpark, MySQL, SQL Server, Snowflake, WBIS, ETL Pipelines, On-Prem Databases, Data Integration, Telemetry Processing

MLOps, DevOps, and Visualization: Azure DevOps, MLflow, Git, Docker, Power BI, Tableau, Grafana, CI/CD, Forecast Monitoring Dashboards, Data Drift Monitoring

Research and Modeling Tools: Cyclic-LiNGAM, C-PATH, CCI, MATPOWER, GNN Models, Influence Graphs, MinT, Hybrid-MinT, Bottom-up Reconciliation