Josh Meehl

joshua@meehl.org | 320.224.5017 | www.linkedin.com/in/meehl | github.com/meehljd | mage.meehl.org

PROFESSIONAL EXPERIENCE

Bioinformatic Data Scientist, *LGC (Diagnostics+Genomics R&D)* 2022.03 – present | Alexandria, MN

- *Protein Engineering:* Developed an ML-based pipeline for enzyme development using LLMs and optimization algorithms.
  - Evaluated protein language models (PLMs), derived from LLMs, for functional prediction of enzyme sequences.
  - Analyzed zero-shot prediction methods using marginal likelihoods between mutant and wild-type sequences.
  - Implemented Bayesian optimization and Gaussian processes to efficiently explore the fitness landscape.
  - Reduced the expected number of discovery iterations by 2.2x, projected to save over $580,000 per discovery campaign.
- *Highly-Multiplex NGS Workflow*: Analyzed genotyping panels targeting 200-5,000 loci, utilizing PCR and NGS technologies.
  - Developed a GCN-based primer-dimer prediction model in PyG, reducing primer-dimer read counts by over 35%.
  - Optimized primer design objective function, enhancing functional performance and reducing loci read variance by 20%.
- *Oligos Failure Model:* Enhanced a model for predicting oligos production failures due to biochemical and process factors.
  - Built a model validation suite, addressing issues with data leakage, distribution shift, and class imbalance.
  - Transitioned from MLE to Bayesian inference methods due to poor model fit, utilizing causal DAGs for insights.
  - Identified key product families with distinct production failure profiles, achieving a Brier score of less than 0.1 on PPCs.
  - Deployed to pricing team, leading to calibrated pricing and improved margins on over $37M/year in product sales.
  - Deployed to production, flagging risky orders for manual review and intervention, reducing failures by over 15%.
  - Extended to a multilevel model for heterogeneous product mix and adding Gaussian processes for non-stationarity.
- Automated 7 NGS bioinformatic pipelines in Apache Airflow, reducing R&D labor by 2 FTEs under ISO 13485 control.
- Designed an ML pipeline in AWS to predict positive drug screens from LCMS data, reduced labor by 2.5 FTEs.

Systems Engineer III, *LGC (Diagnostics+Genomics R&D)* 2017.08 – 2022.02 | Alexandria, MN

Evaluated a portfolio of scientific instruments spanning mechanical, electrical, software, and biochemical subsystems.

- Conducted DOE to pinpoint failures in SARS-CoV-2 PCR diagnostic workflow, resulting in $60,000/day savings.
- Engineered computer vision system for precise reagent dispense in PCR instruments, reducing mis-dispenses by 85%.
- Created PCR assay scoring algorithm employing semi-supervised learning techniques to increase scoring consistency.
- Led regression and integration testing for software systems encompassing UI, controls, analytics, and protocol systems.

Lead Energy Engineer, *Vidaris (now SOCOTEC)* 2008.07 – 2017.08 | New York, NY

Conducted computational modeling and simulation of thermal and energy systems for complex building projects.

- Key project included: Carnegie Hall, Moderna Labs, UPenn Nanotech Lab, and 432 Park Ave (tallest US residential tower).
- Achieved annual energy savings exceeding $2 million with optimized energy strategies for clients.
- Specialized in advanced control systems and electric-thermal cogeneration. Developed DevOps tools for model validation.

For additional prior roles, refer to linkedin.com/in/meehl

EDUCATION

M.S. Analytics, *Georgia Institute of Technology* 2021.01 – 2024.05 | Atlanta, GA

Track: Computational Data Analytics (Data Science & AI)
Practicum: *Efficient Protein Engineering Variant Screening using LLMs and Bayesian Optimization*

B.S. Mechanical Engineering, *Rose-Hulman Institute of Technology* 2003.09 – 2007.05 | Terre Haute, IN

Minors: Economics; Management Studies

CREDENTIALS & AWARDS

- LGC Value Awards: Brilliance 2021; Curiosity 2018
- Licensed Professional Engineer, 2013 - 2024

SKILLS

Programming Languages
Python, R, SQL, Spark

Vizualizations & Dashboards
Matlibplot, Seaborn, Streamlit, Plotly

Frameworks & Libraries
Sklearn, PyTorch, Airflow, JAX, AWS

Programming Skills
Git, Docker, OOP, APIs, Agile, TDD