Josh Meehl

www.meehl.org joshua@meehl.org 320.224.5017

Bioinformatics Data Scientist with 17+ years of experience with background in computational biology and machine learning. Specializes in developing predictive models for biological datasets, driving innovation and substantial cost savings in biotech. Seeking a Computational Scientist role to spearhead cutting-edge innovations in life sciences.

PROFESSIONAL EXPERIENCE

Bioinformatics Data Scientist, *LGC (Diagnostics+Genomics R&D)* 03.2022 – present | Alexandria MN

Protein Engineering Workflow Project:

- 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 discovery iterations by 2.2x, projected to save over $580,000 per campaign for lab teams.

Genotyping Next-Generation Sequencing (NGS) Workflow Development:

- Developed a primer-dimer (PD) prediction model using a graph-based GCN in PyG, reducing PDs by over 35%.
- Optimized primer design objective function, enhancing performance and reducing loci read variance by 20%.

Oligos Synthesis Failure Prediction Model:

- Designed Bayesian inference model to predict failure likelihood of DNA oligo sequences with a Brier score of 0.1.
- Validated models and addressed issues with data leakage, distribution shift, and class imbalance.
- Deployed to marketing, 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%.

Machine Learning Engineering Projects:

- Designed an ML pipeline in AWS to predict positive drug screens from LCMS databases, reduced labor by 2.5 FTEs.
- Automated 7 NGS bioinformatics pipelines in Airflow utilizing a worker cluster, reducing R&D labor by 2 FTEs.
- Engineered ISO 13485 compliant NGS QC analysis applications for clinical and biopharma customers.

Systems Engineer III, *LGC (Diagnostics+Genomics R&D)* 08.2017 – 02.2022 | Alexandria MN

Evaluated a portfolio of scientific instruments with mechanical, electrical, software, and chemical subsystems.

- Conducted design of experiments to pinpoint failures in SARS-CoV-2 diagnostic workflow, saving $60,000/day.
- Engineered computer vision system for precise dispensing in PCR instruments, reducing mis-dispenses by 85%.
- Created PCR assay scoring algorithm using Davies-Boulding Index and human evaluation scores.
- Led integration testing for software systems encompassing UI, controls, analytics, and protocol systems.

Lead Energy Engineer, *Vidaris (now SOCOTEC)* 07.2008 – 08.2017 | New York NY

Built computational simulation models, saving clients over $2 million per year.

Previous roles (see linkedin.com/in/meehl) 06.2007 – 07.2008 | New York NY

EDUCATION

M.S. Analytics, *Georgia Institute of Technology* 01.2021 – 05.2024 | Atlanta GA

B.S. Mechanical Engineering, *Rose-Hulman Institute of Technology* 09.2003 – 05.2007 | Terre Haute IN

CREDENTIALS & AWARDS

- LGC Value Awards: Brilliance 2021; Curiosity 2018

SKILLS

Programming Languages: Python, R, SQL, Spark

Frameworks & Libraries: Sklearn, PyTorch, JAX, PyG, Matplotlib, Streamlit, Apache Airflow

Programming Skills: Git Version Control, Object-Oriented, APIs, Agile, Testing, Data Pipelines

Computing Environments: Linux, Docker, Cloud Computing, AWS

Comuptaional Biology: Multi-Omics (Genomics, Proteomics, Transcriptomics), LCMS, PCR, NGS