CHRISTIAN D. BASILE Resume

Brooklyn, NY | 585-490-2017 | christian.daniel.basile@gmail.com | [Website](#) | [GitHub](#) | [LinkedIn](#)

PROFILE SUMMARY

Computational Scientist with a strong foundation in differential equations, machine learning, and deep learning. Former PhD candidate in Biophysics who, following a pivotal personal chapter in 2024, emerged with renewed focus and purpose. Seeking a full-time role to apply analytical rigor and technical expertise in research-driven or AI-focused domains. Proficient in Julia, Python, MATLAB, SQL, and R, with a proven record of building predictive models in healthcare and engineering. Continuously expanding skills through advanced training in reinforcement learning, optimization, and large language models.

KEY SKILLS & COMPETENCIES

Programming: Julia, Python, MATLAB, R, SQL, C++, JavaScript, HTML, CSS

Data Processing & Management: Large-Scale Datasets, Data Cleaning & Transformation, PySpark, Ray/Dask

Machine Learning: PyTorch, TensorFlow/Keras, scikit-learn, XGBoost, LightGBM, Hyperopt/Optuna, Imbalanced Learning (Imblearn), RNN/LSTM, Generative Models (GANs), Large Language Models (Hugging Face, LoRA, QLoRA, RAG), Bayesian Modeling (PyMC), Reinforcement Learning, Time Series Analysis, Outlier & Anomaly Detection

NLP: GPT, BERT, Transformers, Tokenization, Sentence Transformers, NER, Text Classification, Text Generation, Embeddings

Audio Processing: Whisper (Speech-to-Text), TTS, Audio Classification

Computer Vision: YOLO, OpenCV, Image Classification, Object Detection

Clustering: K-Means, Hierarchical, DBSCAN, GMM, Spectral Clustering, Cluster Evaluation

Dimensionality Reduction: PCA, t-SNE, UMAP, Autoencoders, LDA, Applied to Genomics & High-Dimensional Data

Computational Biology Tools: BLAST, SPAdes, Biopython, Prokka, PlasmidFinder, QUAST, MOB-suite, Bioconductor, Seurat, Scanpy

LLM Tools & RAG: Prompt Engineering, LangChain, RAG Pipelines, LlamaIndex, FAISS, Semantic Search

Pipelines & Workflow Tools: scikit-learn Pipelines, MLflow, PyTorch Lightning, Airflow, Cron, Nextflow

Data Analysis: Feature Engineering, Statistical Modeling, Anomaly Detection, Clustering

Tools & Platforms: AWS, S3, Git/GitHub, Linux, MongoDB, PostgreSQL, Tableau, Flask, Django, Node.js, Express, FastAPI, Streamlit, Gradio, Docker, GraphQL, Matplotlib, Seaborn, Plotly, ggplot2,

Specializations: Differential Equations, Deep Learning, Multi-Objective Optimization, Sensitivity Analysis & Parameter Estimation (Chemical/Biological Models), Statistics

Predictive Modeling: Drug Discovery, Dose Prediction, Cost & Risk Segmentation, Fraud Detection, Disease Classification, Tumor Response, Plasmid Detection, Genomic Data Mining

Optimization: Algorithm Development, Multi-Objective & Derivative-Free Optimization

Interpretability: SHAP, Model Stability

Software Development: Algorithm Optimization, RESTful API Development, Prototyping

Collaboration: Agile Teams, Scientific Communication, Cross-Functional Work

PROFESSIONAL EXPERIENCE

PhD Candidate | University of Rochester (Chemical Engineering) | Rochester, NY | Jan 2024 – Dec 2024

- Developed a workflow that can predict plasmid count and spatial arrangement using short DNA sequencing data applied to bacterial systems.
- Applied principles from machine learning, Bayesian statistics, and multi-objective optimization (Pareto frontier).
- Developed a Graph Neural Network for predicting gene regulatory links from transcriptomic and perturbation data, improving over the rule-based inference in cBONITA.

Research Data Scientist | Blue Cross Blue Shield (Florida Blue) | Philadelphia, PA | Nov 2022 – Jan 2024

- Developed predictive models to estimate monthly costs per member (PMPM) and segmented members into risk categories using insurance claims data.
- Implemented sensitivity analysis, feature engineering, and SHAP analysis to identify optimal models.
- Successfully deployed models into production and crafted real-time prediction scoring scripts.
- Designed and implemented a model to evaluate member adherence to diabetes management protocols.
- Employed advanced techniques to detect potential cases of medical insurance fraud.

Research Scientist | Genentech/Ascent (Contract) | Lawrenceville, NJ | July 2021 – July 2022
- Built ML models using an internal QSP model to predict drug treatment arm responses, implementing a multi-objective approach for model selection.
- Conducted interpretability analysis using SHAP.
- Developed algorithms for optimal model selection in a multi-objective context.
- Collaborated with QSP/AI scientists to validate and enhance workflows.

Algorithm Developer Scientist | Bristol Myers Squibb/Spectraforce (Contract) | Lawrenceville, NJ | Aug 2019 – July 2021
- Implemented and optimized a 757-differential equation model in Julia, significantly improving performance over Matlab by a factor of 5.
- Optimized Matlab algorithms for steady-state ODE model solutions and implemented them in Julia.
- Developed a script for adaptive model reduction.
- Automated the generation of analytical Jacobians for ODE models.
- Applied global sensitivity analysis to enhance model calibration.
- Developed a PBPK app utilizing a model to help the discovery team test different scenarios.

Data Scientist/Engineer | Clarivate Analytics | Philadelphia, PA | Oct 2018 – April 2019
- Constructed Python algorithms for data pipeline automation on Linux servers.
- Conducted literature reviews and applied statistics/machine learning to enhance a research management system.

Consultant Data Scientist (Contract) | Silicon Power Corporation | Malvern, PA | April 2018 – Oct 2018
- Developed algorithms for importing and analyzing time series static transfer data.
- Used deep learning and computer vision to classify electrical products and integrate metadata into internal databases.

Clinical Researcher/Data Scientist | Hospital of UPenn | Philadelphia, PA | March 2016 – Dec 2017
- Built ML models to classify Parkinson's, Alzheimer's, and healthy controls using EMG data.
- Presented findings to medical professionals and researchers.

Control Engineer | Metso Automation | Lansdale, PA | Sep 2013 – Dec 2014
- Installed and maintained control systems in power plant settings.
- Built mathematical models for fossil fuel cells and implemented equipment failure detection solutions.

Research Scientist | Bioinformatics, Columbia University | New York, NY | March 2013 – Aug 2013
- Constructed intracellular cellular heart models and optimized parameters using sensitivity analysis and optimization algorithms.
- Collaborated with IBM Computational Biology on a parameter estimation challenge.

EDUCATION & PROFESSIONAL CERTIFICATES

M.S. Computer Science (in progress) | Georgia Tech, Georgia, GA | Expected 2026
M.S. Chemical Engineering | University of Pittsburgh, Pittsburgh, PA | 2012
B.S. Chemical Engineering | University of Rochester, Rochester, NY | 2010

POSTER PRESENTATIONS & ACKNOWLEDGEMENTS

- Absolute Rank Correlation Coefficients and Coefficient of Variation as Useful Methods for Global Sensitivity Analysis to Guide Parameter Selection for QSP Models, AAPS Pharm-Sci, Aug 2021.
- Summary of the DREAM8 Parameter Estimation Challenge: Toward Parameter Identification for Whole-Cell Models, PLOS Computational Biology, May 2015.
- A Crowdsourced Approach to Whole-Cell Model Parameter Estimation, May 2013.