Dr. César A. Ojeda M.

Generative AI & Machine Learning

Berlin, 10407
+49172 3937366
cesarali07@gmail.com
Linked

Personal Data

Name **César Ali Ojeda Marin**

Place of Birth **Lecherias, Venezuela**

About Me

AI researcher with 10+ years of experience in generative modeling, stochastic processes, and machine learning. Contributed pioneering work on prior-fitted neural operators for MJPs, SDEs, and point processes, with publications at top ML conferences. Lead in pharmacometric applications, developing Transformer models for precision dosing, and serving as AI advisor to the PharMetrX graduate program. Expert in multimodal flow matching and large-scale generative modeling for scientific data (e.g., LHC particle jets). Currently seeking industry roles to advance AI for drug development and scientific R&D.

Experience

2022–Present **Postdoctoral Researcher & AI Advisor**, *University of Potsdam*, Potsdam, Germany

- Senior researcher in stochastic process modeling, co-developing prior-fitted neural operator frameworks for Markov jump processes, stochastic differential equations, and point processes.
- Lead researcher on pharmacokinetics: designed the Amortized In-Context Mixed-Effect Transformer (AICMET) for personalized dose–response forecasting, unifying compartmental priors with amortized inference.
- Deployed large pretrained pharmacokinetic models as a GPU-enabled service on a Kubernetes cluster, integrating AI agents to automate deployment and provide a working web-accessible application.
- Serve as AI Advisor for the PharMetrX graduate program, guiding interdisciplinary research at the interface of pharmacometrics and machine learning.
- Expert in large-scale generative modeling for science, introducing multimodal flow matching methods for particle-cloud data at the Large Hadron Collider (LHC), capable of training on millions of jets.

2019–2022 **Postdoctoral Researcher**, *Technical University of Berlin*, Berlin, Germany

- Advanced generative modeling with Schrödinger bridges, developing neural methods for efficient reverse drift estimation and applying them to genetic data (single-cell RNA).
- Derived a stochastic differential equation for variational training dynamics in Bayesian neural networks and proposed StochControlSGD, a robust optimizer with adaptive, parameter-wise learning rates.
- Designed neural dynamic topic models decoupling topic activity from proportions, achieving state-of-the-art results on UN debates, NeurIPS, and ACL datasets.
- Introduced symbolic schema networks for pretrained language models, uncovering interpretable structures and improving commonsense reasoning in GPT-like models.

2014–2019 **R&D Assistant / Doctoral Researcher**, *Fraunhofer IAIS*, Sankt Augustin, Germany

- Conducted applied research in machine learning for industry partners in the automotive, finance, and job-matching sectors.
- Designed deep generative models (recurrent marked point processes, Wasserstein GANs) for modeling large-scale client–server systems.
- Proposed novel methods for explainable AI, including semantic stochastic paths in latent spaces to analyze classifier decisions.
- Applied NLP and multimodal ML in business process optimization and recommender system development.

2011–2013 **R&D Assistant**, *Forschungszentrum Jülich*, Jülich/Aachen, Germany

- Implemented a parallel tempering algorithm in C with MPI and OpenMP.
- Performed molecular dynamics simulations in LAMMPS to study hydrogen bonding between ionic liquids and cellulose, with applications to biofuel research.
- Developed Python scripts to analyze hydrogen bond dynamics and lifetimes across multiple ionic liquid–cellulose systems.
- Explored coarse-graining approaches and theoretical aspects of free energy calculations in non-equilibrium simulations.

Education

2014–2019 **Dr. rer. nat. (magna cum laude)**, *University of Bonn*, Bonn, Germany
Dissertation: *Approximate Inference for Representation Learning and Stochastic Processes.*

2011–2013 **MSc in Simulation Sciences**, *RWTH Aachen University*, Aachen, Germany
Master's Thesis: *Attention Dynamics in Complex Networks.*

2005–2010 **Licentiate in Physics**, *Universidad Simón Bolívar*, Caracas, Venezuela
Undergraduate Thesis: *Models for the Catalytic Oxidation of CO on Surfaces.*

Selected Publications

- D. Faroughy, M. Opper, C. Ojeda. *Multimodal Generative Flows for LHC Jets.* arXiv:2509.01736, 2025.
- C. Ojeda, W. Huisinga, P. Kavwele, N. Hartung. *Amortized In-Context Mixed Effect Transformer Models: A Zero-Shot Approach for Pharmacokinetics.* arXiv:2508.15659, 2025.
- P. Seifner, K. Cvejoski, D. Berghaus, C. Ojeda, R.J. Sánchez. *In-Context Learning of Stochastic Differential Equations with Foundation Inference Models.* arXiv preprint arXiv:2502.19049, 2025.
- D. Berghaus, K. Cvejoski, P. Seifner, C. Ojeda, R.J. Sánchez. *Foundation Inference Models for Markov Jump Processes.* Advances in Neural Information Processing Systems (NeurIPS), vol. 37, pp. 129407–129442, 2024.
- R. Sanchez, L. Conrads, P. Welke, K. Cvejoski, C. Ojeda. *Hidden Schema Networks.* Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL), 2023.
- K. Cvejoski, R.J. Sánchez, C. Ojeda. *Neural Dynamic Focused Topic Model.* Proceedings of the 37th AAAI Conference on Artificial Intelligence (AAAI-23), 2023.
- C. Ojeda, K. Cvejoski, B. Georgiev, C. Bauckhage, J. Schuecker, R.J. Sánchez. *Learning Deep Generative Models for Queuing Systems.* Proceedings of the AAAI Conference on Artificial Intelligence, 2022.

- N. Malem-Shinitski, C. Ojeda, M. Opper. *Flexible Temporal Point Processes Modeling with Nonlinear Hawkes Processes with Gaussian Processes Excitations and Inhibitions*. Proceedings of the 38th International Conference on Machine Learning (ICML), 2021.

Skills

Technologies	Cloud (AWS: Sagemaker, Amplify, EC2, Lambda, S3), Kubernetes (GPU-enabled clusters), Databases (MongoDB, MySQL), CI/CD (AWS CDK, GitLab, GitHub, Docker), HPC (MPI, OpenMP, CUDA), Simulation Software (LAMMPS)
Machine Learning	PyTorch & PyTorch Lightning, HuggingFace Transformers, Neural Operators, Diffusion Models, Flow Matching, Bayesian Inference
Programming	Python (10+ years), C++ (5 years), Java (2 years)
Languages	English (fluent), German (advanced), Spanish (native), French (intermediate)