Hugo Cisneros

Lead ML Engineer with PhD background in AI, specializing in production ML systems and data platforms.

Personal Data

Website: https://hugocisneros.com
LinkedIn: https://www.linkedin.com/in/hugo-cisneros-04347212b/
GitHub: https://github.com/hugcis/
Google Scholar: Scholar Profile

Work & Research Experience

Mar 2023 - *Current*

Inicio (Startup) *Lead Machine Learning Engineer*, Paris

- **Led core logic team** (5 engineers + QA) through company growth from **5-person startup to 35 employees** post-seed funding; drove technical direction for data engineering, ML/AI modeling, and geoprocessing infrastructure
- **Scaled renewable energy site selection platform** to process **entire countries in <24h** (from 1h/city), scanning **10TB of geospatial data** with **100+ simultaneous constraints** using distributed AWS infrastructure (ECS, Aurora PostgreSQL/PostGIS)
- Built **end-to-end ML pipelines**: computer vision for urbanism map segmentation (**SAM - Segment Anything Model**, classical CV); **LLM-based document parsing** and knowledge graph extraction; project success prediction models. Tech stack: Python, Rust, PyTorch, AWS
- Developed **grid intelligence platform** extracting structured data from **thousands of regulatory PDFs**, creating comprehensive map of solar projects in development across **France, Italy, and UK** for **~100 enterprise users**
- Mentored junior engineers and interns on ML best practices, code review, and system design

Completed PhD (2019-2023, see Education below)

Apr - Nov 2019

CIIRC (Czech Institute of Informatics, Robotics and Cybernetics) *Research Intern*, Prague

Under the supervision of Tomas Mikolov (Facebook AI Research).

- Studied emergence, complexity and spontaneous organization in complex systems and their applications to Artificial intelligence.
- Built a neural network-based complexity metric for measuring emergence in cellular automata and other dynamical systems (implemented in **C**) which led to a peer-reviewed publication. [GitHub]

Mar - Sep 2018

INRIA and CNRS (LIMSI) *Research Intern*, Paris

Under the supervision of Xavier Tannier and Ioana Manolescu.

- Built a pipeline generator for extracting and integrating multiple data sources with **Natural language processing (NLP) and data processing algorithms** for data journalism. [GitHub]
- Collaborated with journalists from *Le Monde (Les Décodeurs)* on automating their data processing pipelines and using NLP for their investigations.
- Reviewed literature on **machine learning in graphs**, **automatic knowledge base construction** and **natural language processing for fact checking**.

Jun 2017 - Mar 2018

Aiden.ai *Software Engineering and Machine Learning Intern*, London

- Built an **AI powered virtual colleague** for Marketing analysts based on **Natural Language Processing** intent recognition (LUIS)
- Implemented **machine learning pipelines** with Python for marketing data forecasting, classification and user clustering (scikit-learn, Keras, FB Prophet)
- Participated in implementing the chat interface and the Natural Language recognition system with **Typescript**

Sep 2016 - Feb 2017

ENS Ulm, Kastler-Brossel Laboratory *Research Assistant*, Paris

Light control and propagation in amplified multimode fibers

- Implemented and optimized **finite elements simulations** with Python and Matlab
- Performed **high performance computing** on large distributed clusters
- Worked with PhD candidate on building a tool for optimizing the propagation of a light beam in optical fibers [Report]

Education

May 2023 - Nov 2019	PhD, **INRIA, CIIRC CTU** (Czech Technical University in Prague), Paris & Prague ***Unsupervised learning with Complex Systems and Evolution*** Under the supervision of Tomas Mikolov and Josef Sivic. Topics: complex dynamical systems, self-organization, artificial evolution, artificial intelligence. · Developed and trained **deep learning models** (PyTorch, custom architectures) for studying emergence and complexity in dynamical systems · Built **highly optimized C code** for large-scale simulations of cellular automata and complex systems · Conducted **distributed computing** experiments using **SLURM clusters** and **OpenMPI** for parallel processing · Published **4 peer-reviewed papers** in conferences and journals (CoLLAs, ALIFE, IEEE SSCI) · Supervised Master-level theses and internship projects
Sep 2019 - Sep 2018	MVA Master in Machine Learning and Applied Mathematics, **ENS Paris Saclay**, Paris Relevant Coursework: Convex Optimization, Probabilistic Graphical Models, Computer Vision, Reinforcement Learning, Deep Learning, Speech and Natural language processing, Kernel Methods, Biostatistics, Theoretical Foundations of Deep Learning — (GPA: 16.2 / 20)
Sep 2018 - Sep 2015	Master of Science in Engineering, **Mines ParisTech**, Paris Specialization: Computer Science — (3.7 GPA) Relevant Coursework: Machine Learning, Probabilities, Statistics, Programming
Aug 2015 - Sep 2013	Preparatory class for *Grandes Ecoles* **Lycée Stanislas** (Paris) MPSI and MP* Bachelor's Degree in Mathematics and Physics, national competitive exam for entering engineering school.
Aug 2013	Scientific Baccalauréat (High school diploma in Maths, Physics and Life Sciences) - High distinction

Publications

Cisneros, H. **Unsupervised Learning in Complex Systems**. Thesis.
Link: https://arxiv.org/abs/2307.10993

Herel, D., Cisneros, H., & Mikolov, T. **Preserving Semantics in Textual Adversarial Attacks**. Pre-print.
GitHub repo: https://github.com/DavidHerel/semantics-preserving-encoder

Cisneros, H., Sivic, J. & Mikolov, T. **Benchmarking Learning Efficiency in Deep Reservoir Computing**. First Conference on Lifelong Learning Agents (CoLLAs 2022).
GitHub repo: https://github.com/hugcis/benchmark_learning_efficiency

Cisneros, H., Sivic, J. & Mikolov, T. **Visualizing computation in large-scale cellular automata**. Artificial Life Conference Proceedings 32, 239–247 (2020).

Cisneros, H., Sivic, J. & Mikolov, T. **Evolving Structures in Complex Systems**. in 2019 IEEE Symposium Series on Computational Intelligence (SSCI) 230–237 (IEEE, 2019).
GitHub repo: https://github.com/hugcis/evolving-structures-in-complex-systems.

Projects

Mar - Aug 2021	Participated in the Open-endedness evolution challenge at the GECCO 2021 conference competition track. Developed an open-ended algorithm based on Neural Cellular Automata in **Pytorch** within the game Minecraft. Finished second place. [GitHub][Blog post]
Jun - Aug 2018	Participated in the n2c2 shared task of Harvard Medical School *Cohort Selection for Clinical Trials* in a joint team from AP-HP and LIMSI. Implemented **weakly-supervised and transfer learning methods for Medical NLP** (Keras). Finished 2nd among 30 teams. [Preprint]

Jan 2018	Built a NLP based tool for discovering and matching similar arXiv papers based on similarity measures including **word embeddings-based similarities** of their abstract and **co-authorship graph distance**. [GitHub]
Feb 2017	Implemented a multi-currency blockchain in Python with a team of 9 people (Cryptography, network programming, team software development)

Programming Skills

ML/AI:	PyTorch, Langchain, TensorFlow, Keras
Languages:	Python, C, Rust, Java, TypeScript, Matlab, Scala, Ruby, C++
Data:	SQL (PostgreSQL), Airflow / Dagster
Tools & Platforms:	Git, Docker, AWS (ECS, S3, Aurora), Terraform, LaTeX

Languages

English: Fluent
French: Native
Spanish: Intermediate
Japanese: School level

Interests and Activities

· Creative coding, Live Coding, Generative Art, genocto.xyz
· Technology, Open-Source, Programming
· Mathematics, Statistics and Probabilities
· Running, Hiking, Fencing, Piano, Guitar