

WORK EXPERIENCE

- NASA** | **Machine Learning and Natural Language Processing Specialist**, 2021-Present, New York, NY
Designed, trained and evaluated machine learning models that expand the NASA / Smithsonian Astrophysical Observatory's [Astrophysics Data System](#) services with automatic classification and labeling capabilities.
Organized and ran community challenges for the [Workshop for Artificial intelligence for Scientific Publications](#), building and curating novel tasks, datasets and baselines. Work done in Python with Huggingface.
- Apple** | **Machine Learning Engineer**, Siri Speech team, Summer 2018, Cheltenham, UK
R&D of binary neural network architectures for speech keyword detection on low power devices. Demonstrated feasibility of binary weights for high accuracy, low resource models by reducing required memory size by 50%.
Work done in Python and Keras/TensorFlow using Oracle Grid Engine.
- AT&T** | **Research Scientist**, Data Science and AI Research Lab, Summer 2017, New York, NY
R&D of generative adversarial networks (GANs) over text data, improving response speed for an adaptive dialog assistant system. Generated tailored answers to customer queries, significantly reducing average speed of answer and handle time for care agents. Work done in Python and Keras/TensorFlow using Hadoop.
- Hunter** | **Adjunct Lecturer**, Hunter College, 2016-2021, New York, NY
Teaching of advanced algorithms and data structures, their design and analysis. Supervision of programming project for CS majors and minors. Conception and creation of course syllabus, materials, and examinations. Mentored undergraduates and helped prepare them for the workforce.
- CUNY** | **Research Assistant**, [Graduate Center](#) of the City University of New York, 2012-2022, New York, NY
Collaborated with Pr. Michael Mandel at the [Brooklyn College Speech Lab](#) on adversarial attacks and defenses of neural networks, and the use of recurrent neural networks for audio source separation.
Collaborated with Pr. Andrew Rosenberg at the Queens College speech lab on the use of reservoir networks and prosodic information for conflict detection and NLP. Performed sentiment analysis on user discussions on Reddit.
Work done in PyTorch with CUDA backend.
- CNRS** | **Research Engineer**, [Institution de la Vision](#), Fall 2011, Paris, France
Conception and implementation of a rat brain model for research on the neural bases of active exploratory behavior in rats. Validated the hypothesis by improving the neural network model's performances by evolution and genetic algorithm. Work done in C++ using the [SFERES](#) framework.

EDUCATION AND DEGREES

- Ph.D.** | Doctorate in Computer Science at the CUNY Graduate Center in New York, NY
Focus on neural network theory applied to speech, reservoir and adversarial networks
Member of the [IGERT From Data to Solutions](#) program of Columbia University
- M.S.** | Master's Degree in Computer Science from Sorbonne Université, France
Specialized in Artificial Intelligence and Decision Theory
- B.S.** | Bachelor's in Math and Computer Science, from Sorbonne Université, France
- Bilingual English/French

SELECTED WORKS & PUBLICATIONS

- 2024 | [INDUS: Effective and Efficient Language Models for Scientific Applications](#)
Proceedings of the 2024 Conference on Empirical Methods in NLP: Industry Track
- 2024 | [Enriching a Time-Domain Astrophysics Corpus with Named Entity, Coreference and Astrophysical Relationship Annotations](#)
Proceedings of the 2024 International Conference on Computational Linguistics, Language Resources and Evaluation
- 2023 | [Function of Citation in Astrophysics Literature \(FOCAL\): Findings of the Shared Task](#)
Proceedings of the 2nd Workshop on Information Extraction from Scientific Publications
- 2023 | [Experimenting with Large Language Models and vector embeddings in NASA SciX](#)
- 2022 | [Improving astroBERT using Semantic Textual Similarity](#)
Astronomical Data Analysis Software and Systems 2022
- 2022 | [Overview of the First Shared Task on Detecting Entities in the Astrophysics Literature \(DEAL\)](#)
Proceedings of the first Workshop on Information Extraction from Scientific Publications

SELECTED WORKS & PUBLICATIONS (CONTINUED)

- 2022 [Finite Gaussian Neurons: Defending against adversarial attacks by making neural networks say "I don't know"](#)
Ph.D. thesis presented to the CUNY Graduate Center Computer Science Program
- 2021 [Building astroBERT, a Language Model for Astronomy & Astrophysics](#)
Astronomical Data Analysis Software and Systems XXXI
- 2018 [Enhancement of Spatial Clustering-based Time-frequency Masks using LSTM Neural Nets](#)
- 2018 [Improved MVDR Beamforming Using LSTM Speech Models to Clean Spatial Clustering Masks](#)
- 2017 [Combining Spatial Clustering with LSTM Speech Models for Multichannel Speech Enhancement](#)
Presented at the 2017 Mid-Atlantic Student Colloquium on Speech, Language and Learning
- 2016 [Linguistically-Motivated Features for Language Recognition](#)
- 2015 [Speech Lab @ Queens College: Language Recognition Evaluation 2015](#)
Part of the 2015 NIST Language Recognition Evaluation Plan
- 2014 [Reservoir Computing: A New Paradigm for Neural Networks: A Survey](#)
Second Examination of the CUNY Graduate Center Computer Science Program
- 2013 [Let Me Finish: Automatic Conflict Detection Using Speaker Overlap](#)
Proceedings - InterSpeech 2013
- 2012 [Exploratory Behaviour Depends on Multisensory Integration during Spatial Learning](#)
Artificial Neural Networks and Machine Learning - ICANN 2012