

ANTONIN VIDON

[Mail](#) | [LinkedIn](#) | [GitHub](#) | [Website](#)

EDUCATION

Columbia University MS in Data Science Relevant coursework includes Neural Networks and Deep Learning, Exploratory Data Analysis and Visualization, Algorithms for Data Science, Computer Systems for Data Science, Advanced Deep Learning, Reinforcement Learning	New York, USA Dec 2022
École Polytechnique MS in Applied Mathematics Relevant coursework includes Data Analysis and Unsupervised Learning, Machine and Deep Learning, Statistical Modeling, Statistics, Learning Theory, Regression theory and applications, Rare Event Simulation, Stochastic Models in Finance, Micro & Macroeconomics	Palaiseau, France Aug 2021
Lycée Hoche Preparatory class for nationwide competitive exam (MPSI/MP*) Undergraduate coursework includes Mathematics, Physics and Computer Science (Python and SQL)	Versailles, France Jul 2018

WORK EXPERIENCE

Hyperscience Machine Learning Engineer (Unstructured NLP) • Co-developing a ML pipeline performing a large range of NLP tasks over unstructured documents	New York, USA Jan 2023 - Present
Hyperscience Machine Learning Engineer Intern (Unstructured NLP) • Implemented clustering based approaches to improve reading order of segments within complex documents: invoices, bills, receipts,... • Annotated dataset of challenging documents (e.g. multiple columns, tables) and finetuned Faster R-CNN based layout detection model • Improved accuracy of context based downstream tasks in product pipeline: +3.9% for auto-annotation and +3.8% for extraction	New York, USA Jun 2022 - Aug 2022
Huawei Reinforcement Learning Research Intern • Implemented recurrent A2C and DQN models to solve navigation tasks from raw visual information in an interactive environment • Developed and trained a deep generative model to imitate "expert-like" navigation behavior on different types of surfaces • Modeled posterior distribution of future trajectories by combining the imitation prior with a flexible task specific goal likelihood	London, United Kingdom Mar 2021 - Aug 2021
Bain & Company Associate Consultant Intern • Conducted market potential analyses based on financial datasets of onshore/offshore wind turbine manufacturers • Co-designed a transformation program for a large European manufacturer's consisting of a 2B\$ SG&A reduction plan over 3 years	Paris, France Jun 2020 - Aug 2020
Junior Enterprise of École Polytechnique Treasurer & Project Manager • In charge of a 200,000€ budget, accounting, payroll, and tax payment • Drafted proposals for potential clients and supervised diverse data-oriented projects (e.g. web app development, market analysis) • Awarded "Outstanding Investment" for involvement in École Polytechnique's student community	Palaiseau, France Mar 2019 - Dec 2020
Shanghai Jiao Tong University Teaching Assistant in Physics • Taught Electromagnetism and Thermodynamics to Chinese undergraduates preparing France's Grandes Écoles entrance exams • Purchased, welded, and miniaturized boxes for the controlled steering of a D.C.-motor drive unit • Designed and built a lab equipment database to improve department's inventory management	Shanghai, China Oct 2018 - Mar 2019

ACADEMIC & PERSONAL PROJECTS

Department of Electrical Engineering of Columbia University Image-to-Image translation with cGAN • Implemented U-Net generator and discriminator and conducted ablation experiments on reconstruction task for Facades dataset • Pretrained downsampling path of generator on ImageNet and finetuned whole generator on Country211 dataset for colorization task	New York, USA Apr 2022 - May 2022
Department of Electrical Engineering of Columbia University Phase recognition in hernia surgery • Used MobileNetV2 as backbone to design phase recognition architectures that classify frames from surgery videos among 14 phases • Modeled label correlation with surrounding frames labels as well as surgery progression with LSTM blocks and reached 80% accuracy	New York, USA Sep 2021 - Dec 2021

Department of Electrical Engineering of Columbia University
Squeeze and Excitation Networks [🔗](#)

New York, USA
Sep 2021 - Dec 2021

- Implemented ResNet, ResNeXt and InceptionV3 in TensorFlow as well as Squeeze and Excitation blocks
- Reduced classification error using correlation modules on CIFAR-10, CIFAR-100, and Tiny ImageNet by .5 to 4.5% for ResNet/NeXt
- Performed analysis of ratio, stage integration, activation distributions and inference time with Squeeze and Excitation blocks

Department of Statistics of Columbia University
Energy consumption and human development [🔗](#)

New York, USA
Sep 2021 - Dec 2021

- Conducted analysis of the cross directional causality between energy consumption, GDP, years of schooling and life expectancy
- Built an interactive component to visualize the evolution of the energy mix across time for various HDI index ranges (D3)

Classification for Breast Histopathology [🔗](#)

Feb 2021 - Dec 2021

- Conducted exploratory data analysis of patches scanned at x40 (e.g, class balance, kernel density of tissue color in HSV space)
- Oversampled cancerous patches and selected XGBoost as best state-of-the-art classifier based on cross-validation: 81.4% acc.
- Built from scratch and trained customized versions of ResNet18, ResNet34 and ResNet50: 85.8% best test acc.

Department of Applied Mathematics of École Polytechnique
Improving patient care of young women with breast cancer

Palaiseau, France
Sep 2020 - Dec 2020

- Processed results from surveys made at different stages of illness and built life trajectory related scores to analyze clinical pathways
- Performed t-SNE, PCA and clustering for young patients under 45 to demonstrate the need for an age specific treatment

Department of Physics of École Polytechnique

Palaiseau, France

Integration of physical models into voxel-based video games

Sep 2019 - Jun 2020

- Implemented thermal model of corrosion, diffusion, and passivation of metallic voxel on Unity engine in C#
- Built gameplay to interact with these models in order to enhance pedagogical and recreational features of the game
- Ranked 1st/112 capstone projects and awarded “Excellence prize” during a public prototype exhibition

Department of Computer Science of École Polytechnique

Palaiseau, France

COVID19 Retweet Prediction [🔗](#)

Sep 2020 - Dec 2020

- Carried out thematic clustering and differential prediction of number of retweets with Gradient Boosting and Quantile regression
- Performed text embedding with Bidirectional Encoder Representations (BERT, Google) for deep prediction with TensorFlow

Other projects:

- Developed a Shiny App to create animated gifs from a dynamic interface with R [🔗](#)
- Implemented an Importance Sampling algorithm to estimate value at risk for complex portfolios

SKILLS & INTERESTS

- IT: Python (PyTorch and TensorFlow), R, Java, SQL, C++/C#, LaTeX
- Languages: French (native), English (fluent), German (intermediate)
- Major interest: Piano (15 years) - Numerous performances & CEM diploma (highest non-professional degree of Music Theory)