# Marvin Martin (He/Him/His)

**SE ML / Machine Learning Engineer**

New York, NY 10011 • marvtin@bu.edu • +1 (917) 815-9864

## EDUCATION

### Boston University
Boston, MA
M.Sc. in Computer Information Systems  01/2021-09/2021
- Major: Data Analytics, Grade A (4/4)
- Relevant Coursework:
  Analysis of Algorithms, Big Data Analytics,
  Advanced Programming Techniques,
  Machine Learning, and Deep Learning.

### Ecole Centrale d'Electronique
Paris, FR
M.Sc. in Computer Information Systems  09/2018-12/2020
- Major: Data Science, Grade A (Top 5%, 16.8/20)
- Relevant Coursework:
  Advanced Database, DevOps, Web Development,
  Cloud computing, Mathematics for Data Science,
  Calculus, Algebra, Physics, and Signal Processing.

### INSEEC London
London, UK
B.Sc. in Computer Science  01/2019-05/2019
- Relevant Coursework:
  Java, Probability and Statistics, and Public Speaking.

### Heriot-Watt University
Edinburgh, UK
B.Sc. in Computer Science  01/2018-06/2018
- Relevant Coursework:
  Electronics, C Language, and Business Management.

## EXPERIENCE

### Vic.ai
New-York, NY, USA
Machine Learning Engineer  11/2021-Now
Vic.ai builds AI-based autonomous accounting systems.
- SOTA NLP model development.
- MLOps - Build CI/CD/CT training pipelines.
- Support 20,000 models training daily on AWS (+800%).

### Centre Securité Requin (CSR)
Réunion, FR
Machine Learning Engineer Consultant  11/2020-Now
CSR is a public research center that studies shark attacks.
- Build Neural Network for underwater shark classification.
- Research related to Adversarial Domain Adaptation.
- Create pipelines using Nifi to classify videos in real time.
- Investigate model reliability based on weather conditions.

### Heuritech
Paris, FR
Training Data Operator Intern  04/2020-08/2020
Heuritech builds visual recognition and trend forecasting
technologies for fashion industry.
- Developed Python internal automation tools.
- Ran data engineering tasks to improve datasets.
- Trained Deep ConvNets from containers.
- Active Learning research on acquisition functions.
- Training metrics investigation.
- Prioritized strategies for error and bias reduction.

### Lectra, Innovation Lab
Bordeaux , FR
Lab Intern  05/2019-06/2019
Lectra develops systems for the textile industry.
- Explored Generative Adversarial Network for fashion.
- Collected web data using Data Crawling techniques.

## LINKS

Website   LinkedIn   Github   Medium

## PUBLICATIONS

**In progress**: ADA-Shark: A framework for shark detection employing underwater cameras and domain adversarial neural network (Boston University & CSR).

## PROJECTS

**Domain adversarial neural network** [Report]  [Code]

Reproducibility study of the papers: Ganin Y., et al. Domain-adversarial training of neural networks, and Pei Z., et al. Multi-adversarial domain adaptation. I increased by 0.025 the target accuracy, of the MNIST→MNISTM task.

**PySpark neural network from scratch** [Article]  [Code]

Spark implementation of a shallow neural network to classify handwritten digits. The training can run in the Cloud with large dataset. Gradients are computed via MapReduce jobs.

**Emotion based Art Generator** [Article]  [Code]

Conditional generative adversarial neural network (CGAN) trained to generate $64 \times 64$ painting images for a selected style (portrait, abstract, landscape or flower-painting) and emotion (positive, neutral, or negative).

**LogNice** [Code]

Python distributed code executor for an online coding competition platform. Code submissions are queued with RabbitMQ and Celery, then ran in a docker container. Scores are stored using a Redis database and sent to the user using SocketIO.

## CERTIFICATIONS

- Workera Software Engineer Deep Learning Assessment
- Deep Learning Specialization  by DeepLearning.AI.
- GANs Specialization  by DeepLearning.AI.
- Deep Learning and Computer Vision A-Z .

## SKILLS

- **Languages** : Python, Java, Bash, R, C, JavaScript.
- **Libraries** : Tensorflow, Pytorch-Lightning, Skit-Learn, Wandb, OpenCV, Pandas, Numpy.
- **Tools** : Git, Circle-ci, Docker, Sentry, Datadog.
- **Cloud** : GCP, AWS (S3, Lamda, Batch, Step Functions, EFS), Ray, Hadoop, Spark.
- **Databases** : MongoDB, PostgreSQL, FireBase, HDFS.

## LANGUAGES

**English**: Fluent   **French**: Native   **Spanish**: Novice

## ACTIVITIES & LEADERSHIP

**Comedian, Stuntman**: Performed in Timéo in front of 1000 people, 3 times a week as a skateboarder (2016).
**Sports**: Skateboard (Instructor), Surf, and Slackline.
**Achievements**: 300km in 3 days on a longboard.