

Daniel C. FERREIRA

daniel.ferreira.1@gmail.com

github.com/dcferreira | www.dcferreira.com

Based in Vienna, Austria | Born in 1992

HIGHLIGHTS

- Over 5 years of experience in Machine Learning (academia & industry)
- Strong mathematical background
- Curious and fast learner
- Most interested in NLP and cybersecurity topics

WORK EXPERIENCE

- 2019–2022 Machine Learning Engineer at [Cyan Security](#)
- The position involved a mix of data engineering, ML engineering, MLOps, and research. Topics included website categorization (using multilingual text and images), DNS tunneling detection, and IoT security. Heavy emphasis on developing production-ready ML models, and following MLOps best practices. Using state-of-the-art Machine Learning models (e.g., Transformers).
- 2016–2019 Researcher in the Communications Networks group, TU Wien, integrated in the [Big-DAMA project](#).
- Research topics mostly related to how to represent network traffic for detecting attacks at the network level (i.e., which features to use). Experimented with both classical features, and feature learning approaches using Deep Learning techniques, and in particular representing traffic in 2-dimensional spaces.
- 2016 Junior Researcher at Priberam Labs, integrated in the [SUMMA project](#).
- Worked in a massive project for automating media monitoring. In particular, tackled a *named entity recognition* problem, using both classical methods, and Deep Learning approaches for NLP.

EDUCATION

- 2015 MSc in [Applied Mathematics](#) at [Instituto Superior Técnico](#)
MAJOR: Computation
THESIS in cross-lingual classification (grade: 19/20)
– I wanted to classify news documents in German, given an English-only training set. I proposed two novel approaches to find word representations in a bilingual space, one using CCA and another using an original formulation. Details can be found [here](#).
Advisors: [André F. T. MARTINS](#), [Mariana S. C. ALMEIDA](#), [M. Rosário OLIVEIRA](#)
- 2013 BSc in [Applied Mathematics and Computation](#) at [Instituto Superior Técnico](#)
- 2009 Certificate of Proficiency in English (CPE) by the [University of Cambridge](#)

HIGHLIGHTED PUBLICATIONS

- 2019 [Extreme Dimensionality Reduction for Network Attack Visualization with Autoencoders](#), IJCNN 2019
- 2019 [Towards modular and programmable architecture search](#), NeurIPS 2019
- 2017 [A meta-analysis approach for feature selection in network traffic research](#), ACM SIGCOMM 2017 Reproducibility Workshop
- 2016 [Jointly Learning to Embed and Predict with Multiple Languages](#), ACL 2016

SOFTWARE EXPERIENCE

Professional proficiency	Python, Linux, Pandas, Tensorflow, Keras, scikit-learn, NumPy, Docker, Databricks, Spark, LaTeX, Wireshark
Some professional experience	JavaScript, PyTorch, Scapy, GCP, AWS, Azure, SQL
Hobby experience	R, Go, Rust, C, Photoshop, Inkscape

HIGHLIGHTED PROJECTS

- Developed [Traffic Flow Mapper](#): a prototype to visualize network traffic flows in real time in a 2D plot
- Developed [mdcgenpy](#): a generator of random clustered data, for evaluation of clustering algorithms
- Developed the [NTARC database](#) of network traffic research, for meta-analysis purposes.
- Heavily contributed to [DeepArchitect](#): a framework for neural network architecture search
- Participated in Capture the Flag (CTF) competitions with [WE_OWN_YOU](#)
- Developed task-oriented [multilingual word embeddings](#): useful for text tasks involving multiple languages

RELEVANT COURSES

2022	Databricks Academy's Optimizing Apache Spark on Databricks
2022	Databricks Academy's Data Engineering with Databricks
2021	TryHackMe's CompTIA Pentest+ Pathway (did not attempt the actual certification)
2021	Udemy's JavaScript - The Complete Guide
2014-2018	Lisbon Machine Learning School (participant in 2014, and mentor in the following years)
2015	Coursera's Software Security
Machine Learning	Lisbon Machine Learning School (participant in 2014, and mentor in years 2015, 2016, 2017, 2018); Deep Learning for Visual Computing (2019); DeepLearn Bilbao (2017); Statistical Learning (2015); Statistical Methods in Data Mining (2014)
Security	Network Security - Advanced Topics (TU Wien, 2018); PhD School on Traffic Monitoring and Analysis (2017); Advanced Internet Security (TU Wien, 2017); Software Security (2015)