

# LEONARDO V. CASTORINA

✉ leonardo.castorina@outlook.com 📞 REDACTED ONLINE 🌐 universvm 📄 in/leonardo-castorina

## EDUCATION

### The University of Edinburgh

- Ph.D. Biomedical Artificial Intelligence (CDT) Sep 2021 - Present
- MScR. Artificial Intelligence (Distinction) Sep 2020 - Sep 2021
- BSc. (Hons) Biochemistry Sep 2016 - May 2020

## EXPERIENCE

**Microsoft Research (Redmond)** – Machine Learning Consultant Nov 2023 - Apr 2024

**Microsoft Research (Redmond)** – Research Scientist Intern Jun 2023 - Sep 2023

- Statistically analysed 30K TCR repertoires to identify key binding chemical patterns.
- Developed robust statistical models for TCR-MHC binding by approximating interaction energy.

**NEC Labs Europe** – Machine Learning Consultant Feb 2023 - Apr 2023

**NEC Labs Europe** – Research Scientist Intern Oct 2022 - Feb 2023

- Developed GNN models for TCR-pMHC binding prediction integrating 3D and biological features.
- Implemented distance-based analysis algorithm to evaluate generalization abilities of models.

**Osmitau Technologies** – CEO & Co-Founder Mar 2019 - Jun 2020

- Co-founded while completing my BSc. to develop a Machine-Learning drowsiness detection tool.
- Led the tool development, managed finances, and pitched to automotive manufacturers.
- Co-authored a white paper on automotive safety assessing the state of the field. [🔗 Read Here](#)

**IBM** – Software Engineer Intern (Extreme Blue) Jun 2019 - Sep 2019

- Built an adaptive radius to search for vulnerable missing people using LSTM/CNN, OpenStreetMap saving the UK Police ~1.5 hours per search.

**P&G** – R&D Intern (Personal Healthcare) Jul 2018 - Sep 2018

- Performed market and consumer research to investigate the health benefits of precision vitamins.

**Swiss Institute of Bioinformatics** – Intern May 2017 - Jul 2017

## PROJECTS

**TIMED-Design** Deep Learning for de novo Protein Design [🌐 GitHub](#) [🔗 Try it out](#)

- Designed, benchmarked, and improved 3D CNN models by 10+% for inverse protein folding.
- Implemented user-friendly UI to use models, analyse predictions, and Monte Carlo sampling.
- Winner of the 2023 TensorFlow Community Spotlight Prize out of 1K+ projects.

**Aposteriori** Protein Structures Voxelisation for Deep Learning [🌐 GitHub](#)

**TEDx** AI and Healthcare: The Next Frontier [📺 Coming Soon](#)

**Royal Society Summer Science Exhibition** Programming Proteins [📺 YouTube](#)

## LANGUAGES

English (Fluent), Italian (Fluent), French (DELF B2), Sicilian (Fluent)

## PATENTS & PUBLICATIONS

**Leonardo V Castorina**, F. Grazioli, P. Machart, A. Moesch, and F. Errica. *Assessing the Generalization Capabilities of TCR Binding Predictors via Peptide Distance Analysis*. *bioRxiv*, 2023. doi: 10.1101/2023.07.29.551100. URL <https://www.biorxiv.org/content/early/2023/07/31/2023.07.29.551100>

**Leonardo V. Castorina**, R. Petrenas, K. Subr, and C. W. Wood. *PDBench: Evaluating Computational Methods for Protein Sequence Design*. *Bioinformatics*, 2023a. ISSN 1367-4811. doi: 10.1093/bioinformatics/btad027. URL <https://doi.org/10.1093/bioinformatics/btad027>

F. Grazioli, P. Machart, A. Mösch, K. Li, **Leonardo V. Castorina**, N. Pfeifer, and M. R. Min. *Attentive Variational Information Bottleneck for TCR-peptide Interaction Prediction*. *Bioinformatics*, 2022. ISSN 1367-4803. URL <https://doi.org/10.1093/bioinformatics/btac820>. btac820

**Leonardo V. Castorina**, K. Subr, and C. W. Wood. *TIMED-Design: Efficient Protein Sequence Design with Deep Learning*. *UNDER REVIEW AT PEDS*, 2023b. doi: 10.5281/zenodo.6997495. URL <https://doi.org/10.5281/zenodo.6997495>

B. M. Li, **Leonardo V. Castorina**, M. d. C. Valdés Hernández, U. Clancy, S. J. Wiseman, E. Sakka, A. J. Storkey, D. Jaime Garcia, Y. Cheng, F. Doubal, M. T. Thrippleton, M. Stringer, and J. M. Wardlaw. *Deep Attention Super-Resolution of Brain Magnetic Resonance Images Acquired Under Clinical Protocols*. *Frontiers in Computational Neuroscience*, 16, 2022. ISSN 1662-5188. doi: 10.3389/fncom.2022.887633. URL <https://www.frontiersin.org/articles/10.3389/fncom.2022.887633>

**Leonardo V. Castorina**, R. Petrenas, K. Subr, and C. W. Wood. *TIMED & PDBench: Developing and Evaluating Computational Methods for Protein Sequence Design*. In *APFED-22*, page 51. Advances in Protein Folding, Evolution, and Design, 2022

**Leonardo V. Castorina**, B. M. Li, A. Storkey, and M. C. Valdés-Hernández. *Metrics for Quality Control of Results From Super-Resolution Machine-Learning Algorithms*, 2021. URL <https://datashare.ed.ac.uk/handle/10283/3933>

J. Love, **Leonardo V. Castorina**, T. G. O’Leary, W. Gong, and G. G. Chiarella. *Information Carrier Object and System for Retrieving Information* Patent US11132086B1, 2021. URL <https://patents.google.com/patent/US11132086B1/en>

## BLOG POSTS

**[Towards Data Science] How to Solve the Protein Folding Problem: AlphaFold 2**  Medium

- Awarded Medium Boost prize. Explored AF2 dimensionality and layer functions with visual aids.

**[Towards AI] Latent Diffusion Explained Simply (with Pokémon)**  Medium

- Featured in Towards AI. Used Pokémon to illustrate the concept and applications of diffusion.

**[Better Humans] Obsidian Tutorial for Academic Writing**  Medium

- Featured in the Obsidian Roundup. 50K views within a year of publishing.

**[BH] How to Boost Your Productivity for Scientific Research Using Obsidian**  Medium

- Winner of Obsidian October 2022 Written Content. Received \$400+ and merchandise.
- Featured in the Obsidian Roundup. Received 4K views within one week of publishing.

**[Better Humans] 20+ MacOS Apps to Boost Your Productivity**  Medium

- Named Top 1500 June Article in Medium. Received 4K views within the first month of publishing.

## SKILLS

**Tools:** Python, PyTorch, TensorFlow, Keras, SciKit-Learn, NumPy, Git, Jupyter, Pandas, Streamlit.

**Machine Learning:** Convolutional Neural Networks (CNN), Variational Auto Encoders (VAE), Graph Neural Networks (GNN), Generative Adversarial Networks (GAN), Transformers.