Carlos Mougan

https://cmougan.github.io/
https://www.linkedin.com/in/carlosmougan/

Email : carmougan@gmail.com
Mobile : +34 620 618 275

Professional Summary

I am currently a Marie Curie Research Fellow in the NoBias European ITN and I provide applied skills support advisor at the Alan Turing Institute. I am passionate about predictive modelling and its impact on society. At the moment, I am researching model monitoring (distribution shift and xAI) and AI alignment (ethics and political philosophy). I am very fortunate to have pursued my passions at world-class research and public institutions. In the past, I have also been a statistician at the European Central Bank a consultant at Deloitte, and a visiting researcher at Schufa and BBC.

Experience

- **Marie Curie Research Fellow** — Southampton, United Kingdom
  University of Southampton — *Feb 2021 - Now*
  - **Phd candidate**: Fellowship granted by the European Commission to the Innovative Training Network of NoBias. **Publications**: NeurIPS'23 [1], AAAI'23 [2], AIES'23 [3], NeurIPS'23 (w) [4], NeurIPS'22 (w) [5], ECMLPKDD'21 (w) [6], IFC [7], MDAI 21 [8], EWAF'23 [9]
  - **Schufa Research Visitor**: Publication at NeurIPS'22 Workshop [5]
  - **University of Pisa Research Visitor**: Publication at AIES'23 [3]
  - **BBC DataLab**: Aligning and monitoring BBC sounds RecSys with editorial values
  - **Senior Tutor**: Teaching courses on Data Science and ML for industry professionals, consultants and startups
- **The Alan Turing Institute** — London, United Kingdom
  Applied Skill Team — *April 2022 - Now*
  - **Advisory Support**: I provide academic and industry advisory support in the Applied Skill Team for the Turing Data Study Groups and Turing Internship Network. With a focus on research and industry collaboration, I contribute to the development and implementation of cutting-edge projects between academia and industry.
  - **Principal Investigator**: During this time I lead the publication at NeurIPS'23 Datasets track [1]
- **European Central Bank** — Frankfurt, Germany
  Statistician - Directorate of Statistics — *Jan 2020 - Jan 2021*
  - **Statistician**: Monitoring the quality of statistics on the balance of payments and international investment position, as well as methods and sources in the EU Member States
  - **Machine Learning Engineer**: Developed multiclassification neural network for the National Central Banks and European System of Central Banking
  - **Consulting Management**: End to end management of a Data Science consulting project.
- **Barcelona SuperComputing Center** — Barcelona, Spain
  Applied Learning Methods — *April 2019 - Aug 2019*
  - **Research Data Scientist**: This research line explores the use of Learning Techniques in different domains, from data center optimization to cancer genomics, leveraging different techniques from statistical Machine Learning to state of the art Deep Learning and Neural Networks, and different programming frameworks.
- **CSIC-IIIA** — Barcelona, Spain
  National Research Council - Artificial Intelligence Institute — *Jan 2019 - July 2019*
  - **Deep Learning Research Assistant**: Improving and applying Generative Adversarial Networks (GANs) to approach large scale combinatorial optimization problems.
- **Deloitte Robotics** — Madrid, Spain
  Technological Consultant — *Dec 2017 - August 2018*
  - **Team Leader**: Coordinated a team of two developers, a machine learning engineer, and two data analysts applying Cognitive Automation (CA) to the Audit Process. Implementing Machine Learning to the extraction (OCR) and the classification of unstructured data with WorkFusion. I took part in the delivery of the project where we were able to improve the efficiency and effectiveness of their operations by a 70%.
  - **UiPath Trainer**: UiPath trainer to the international Deloitte's team of Belgium and Netherlands.
  - **Naoqi Developer**: Developing functions for the robots of SoftBank Pepper and Nao for commercial and marketing purposes. We collaborated in events and workshops where Pepper was in charge of 'chairing' and introducing the speakers.
  - **Uipath Developer**: Developer for a leading national banking company where we automated two processes. In the delivery of the project, we obtained automation of 100% and 80% respectively, allowing workers to focus on more creative tasks.
- **AEMET** — Madrid, Spain
  Agencia Estatal de Meteorología — *Sept 2017 - Dec 2017*
  - **Internship**: at the Prediction and Modelling Department with a University collaboration where I modeled and predicted the atmospheric pollution of the northern side of Italy.

Education

- **University of Southampton** — Southampton, UK
  Phd on Model Monitoring and AI Ethics — *Feb. 2021 – current*
- **Universitat Autónoma de Barcelona** — Barcelona, Spain
  Master in Mathematical Modelling: Data Science — *Sept. 2018 – Jul 2019*
- **Universidad Complutense de Madrid** — Madrid, Spain
  Degree in Physics — *Sept. 2013 – Feb 2018*
  - **Specialized in Atmospheric Physics and Astrophysics** : Dissertation in 'Atmospheric Contamination Modelling'.
- **Saint Paul Preparatory High School** — Minnesota, US
  International baccalaureate
  - **Honor International Student at Saint Paul Preparatory**: Received a Scholarship for international students with an excellent curriculum (€10,000).

Other Relevant Information

- **Data Science Stack Exchante**: Top 2 user in 2020 and top 20 all time Stack Exchange Profile
- **Python package developer**: Main contributor of category encoders , Main developer of skshift and explanationspace
- **Machine Learning Competitions**: :*1st BCG Gamma Datathon, 1st at Novartis 2021, 3rd Accenture Health Datathon, 2nd at SpainML Telco-churn and Bronze Medal at Kaggle IEEE-CIS Fraud Detection.*
- **Volunteering with Refugees(Athens) - April 2017**: Volunteering experience working alongside war refugees trying to improve their environment
- **Elite Swimming Athlete 2013 -2017**: National recognition for athletes with an outstanding performance. Achievements: Finalist of the Spanish Olympic Trials 2016, 3^{rd} Team in Spanish Honor Division, 2012 Minnesota State Champions as Individual & Team. Record holder of several state events.

References

[1] Carlos Mougan, Richard Plant, Clare Teng, Marya Bazzi, Alvaro Cabregas Ejea, Ryan Sze-Yin Chan, David Salvador Jasin, Martin Stoffel, Kirstie Jane Whitaker, and Jules Manser. How to data in datathons. *Advances in Neural Information Processing Systems*, 2023.

[2] Carlos Mougan and Dan Saattrup Nielsen. Monitoring model deterioration with explainable uncertainty estimation via non-parametric bootstrap. In *AAAI Conference on Artificial Intelligence*, 2023.

[3] Carlos Mougan, Jose Alvarez, Salvatore Ruggieri, and Steffen Staab. Fairness implications of encoding protected categorical attributes. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '23, page 956–966, Montreal, Canada, 2023. Association for Computing Machinery.

[4] Carlos Mougan, Laura State, Antonio Ferrara, Salvatore Ruggieri, and Steffen Staab. Beyond demographic parity: Redefining equal treatment. In *First Workshop on AI meets Moral Philosophy and Moral Psychology. Neural Information Processing Systems*, 2023.

[5] Carlos Mougan, Klaus Broelemann, Gjergji Kasneci, Thanassis Tiropanis, and Steffen Staab. Explanation shift: Detecting distribution shifts on tabular data via the explanation space. In *NeurIPS 2022 Workshop on Distribution Shifts: Connecting Methods and Applications*, 2022.

[6] Carlos Mougan, Georgios Kanellos, and Thomas Gottron. Desiderata for explainable AI in statistical production systems of the european central bank. In *Machine Learning and Principles and Practice of Knowledge Discovery in Databases - International Workshops of ECML PKDD 2021, Virtual Event, September 13-17, 2021, Proceedings, Part I*, volume 1524 of *Communications in Computer and Information Science*, pages 575–590. Springer, 2021.

[7] Carlos Mougan, Thomas Gottron, Georgios Kanellos, Johannes Micheler, and José Martínez. Introducing explainable supervised machine learning into interactive feedback loops for statistical production systems. In Bank for International Settlements, editor, *Data science in central banking: applications and tools*, volume 59. Bank for International Settlements, 2023.

[8] Carlos Mougan, David Masip, Jordi Nin, and Oriol Pujol. Quantile encoder: Tackling high cardinality categorical features in regression problems. In Vicenç Torra and Yasuo Narukawa, editors, *Modeling Decisions for Artificial Intelligence*, pages 168–180, Cham, 2021. Springer International Publishing.

[9] Francisco Castillo-Eslava, Carlos Mougan, Alejandro Romero-Reche, and Steffen Staab. The role of large language models in the recognition of territorial sovereignty: An analysis of the construction of legitimacy. In *European Workshop of Algorithmic Fairness*, 2023.