Research Statement

Carlos Mougan

Professional Background:

- Current Marie Curie Research Fellow and Applied Skills Support Advisor at the Alan Turing Institute.
- Diverse roles: statistician at the European Central Bank, consultant at Deloitte, industry researcher at BBC, BSC, CSIC, Schufa, and AEMET.
- Previous datathons and hackathons competitor.

Research Expertise:

- Specialized in predictive modeling and its societal impact.
- Experience across the different stages of the ML pipeline: data collection [1], data quality [2], preprocessing [3, 4], modeling [5], and monitoring [6, 7].
- PhD on model monitoring (distribution shift, xAI) and AI alignment (ethics, political philosophy) using feature attribution explanations. Title *Model monitoring in the absence of labeled data via feature attribution explanations.*

Publications and Contributions:

- Published in top conferences: NeurIPS'23 [1], AAAI'23 [5], AIES'23 [4], NeurIPS'23 (w) [6], NeurIPS'22 (w) [7], ECMLPKDD'21 (w) [2], IFC [8], MDAI 21 [3], EWAF'23 [9]
- Notable paper on the topic, working with lawyers: "*Necessity of Processing Sensitive Data for Bias Detection and Monitoring: A Techno-Legal Exploration.*"

Technical Skills:

- Developed and contributed to large open-source Python packages: `category_encoders, skshift.`
- Ranked Top 2 of 2020 on DataScienceStackExchange.

Alignment with Position:

- Enthusiastic about contributing to open source projects, privacy-preserving ML, with expertise in model monitoring.
- Keen interest in generative AI.
- Master thesis in GANs. Previous work in data processing under EU law. Phd in feature attributions explanations and fairness.

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, 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.

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

[4] 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.

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

[6] Carlos Mougan, Laura State, Antonio Ferrara, Salvatore Ruggieri, and Steffen Staab. Beyond demographic parity: Redefining equal treatment, 2023.

[7] 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.

[8] Carlos Mougan Mougan, George Kanellos, Johannes Micheler, Jose Martinez, and Thomas Gottron. Introducing explainable supervised machine learning into interactive feedback loops for statistical production system. In *Irving Fisher Committee on Central Bank Statistics,2021*, 2021.

[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.

[10] Ioanna Papageorgiou and Carlos Mougan. Necessity of processing sensitive data for bias detection and monitoring: A techno-legal exploration. In *NeurIPS 2023 Workshop on Regulatable ML*, 2023.