

Jose Cribeiro

+34 617325922 | ✉ jose@cribeiro.net | in linkedin.com/in/cribeiro-ramallo-datascience | github.com/jcribeiro98 | cv.cribeiro.net

SUMMARY

Ph.D. in Data Science & Machine Learning @ KIT tackling problems in Data Generation, Data Mining, and Theoretical Machine Learning. My expertise includes generative methods for self-supervised outlier detection, high-dimensional (NLP & CV) Outlier Detection, kernel-based machine learning, and subspace selection. My publication record can back-up my expertise in deep learning, mathematical inference and probability theory. Take a look at my research in my Google Scholar page!

SKILLS

Technical: Deep Learning, Statistics, Probability Theory, Computer Vision, Anomaly Detection, Experimental Design, NLP, RKHS.

Programming Languages: Python, SQL, MATLAB, R.

Deep Learning Frameworks: TensorFlow, PyTorch, HF Transformers.

Libraries & Tools: NumPy, Pandas, Scikit-learn, OpenCV, Git, Docker.

EXPERIENCE

PhD. Researcher

Karlsruhe Institute of Technology (KIT)

May 2022 – May 2027

Karlsruhe, Germany

- Focused on generative models for self-supervised outlier detection and theoretical machine learning.
- Published multiple articles in the topics of outlier detection, generative models, and estimators in RKHSs.
- Taught courses (24874) – *Praktikum Data Science* and (242424) – *Data Science für die Wissenschaften* (SS 2024-2026).

AI Research Intern

Gradiant

Sep 2021 – March 2022

Vigo, Spain

- Researched on data efficiency algorithms for computer vision models.
- Developed and released an efficiency strategy that significantly reduced runtime compared to competitors, while minimizing performance degradation in the downstream model.

EDUCATION

Ph.D.: Karlsruhe Institute of Technology

PhD. at the Chair of Information Systems

Karlsruhe, Germany

May 2022 – Late 2026 (expected)

Master of Science: Universidad de Santiago de Compostela

MSc. in Statistics

Santiago, Spain

Sep. 2020 – Feb. 2022

Bachelor of Science: Universidad de Santiago de Compostela

BSc. in Mathematics

Santiago, Spain

Sep. 2016 – June 2020

FEATURED PUBLICATIONS

The Minimax Lower Bound of Kernel Stein Discrepancy Estimation.

AIStats

May 2026

Minimax opt., Stein method, RKHS

- Proved that the KSD estimation cannot be performed at a faster rate than $n^{-1/2}$, in $\mathbb{R}^d$ and in general topological spaces.
- This settles the discussion on the consistency rates for KSD (heavily used in Bayesian and Variational models).

Adversarial Subspace Generation for Outlier Detection in High-Dimensional Data.

Transactions on Machine Learning Research (TMLR)

Aug 2025

RKHS, Generative Models, Subspace Selection

- Developed a mathematical framework for analyzing multi-manifold-entangled data (Myopic Subspace Theory; MST).
- Using MST, we created a Generative Model in an operator space for finding each manifold.
- SOTA performance in tabular Outlier Detection, and promising results in images and text.