Erik Sarriegui

Madrid, Spain
linkedin.com/in/eriksarriegui
+34 679 75 85 68
erik@sarriegui.es
eriksarriegui.github.io

ABOUT ME

Aspiring Data Scientist with a strong self-taught foundation in Artificial Intelligence and software development. My fascination with AI models' ability to interpret complex information has driven me to specialize in this field through personal projects and continuous learning. I am currently pursuing a Master's in Data Science, seeking to combine my passion for learning with a professional opportunity that allows me to apply and expand my knowledge in data analysis and machine learning. My goal is to contribute to the development of innovative AI solutions while continuing to grow professionally.

EXPERIENCE

Vicomtech
Business Developer Intern

Donostia and Bilbao, Spain
May 2024 – Oct 2024

As part of the Digital Health and Biomedical Technologies department, I researched the application of Large Language Models in Mental Health within a clinical setting. In this context:

- I fine-tuned language models using Transformers and PyTorch. One of these models, whose code is publicly accessible, is MediFlow (Model, Article), which performs automatic triage with 90% accuracy.
- I built a demo to test the interest of potential clients and collaborators in the proposed solution.

International Experience
Study on Digitalization Trends in the South Korean Healthcare Sector

Seoul, South Korea
Jan 2023 – Apr 2023

It was a study supported by two hospitals in the Basque Country, which analyzed both the digital trends in the South Korean healthcare sector and specific technological tools and projects.

EDUCATION

Master in Data Science
Universidad Complutense de Madrid

Madrid, Spain
2025 - 2026

Bachelor's Degree in Leadership, Entrepreneurship, and Innovation
Universidad de Mondragón

Irún, Spain
2020 – 2024

PROJECTS & MILESTONES

DeepWind

Inspired by projects involving autonomous drones used for inspecting wind turbine blades in motion, I developed DeepWind, a Computer Vision system based on Deep Learning techniques capable of detecting defects in wind turbines. I have published the entire project on GitHub (link) to make it accessible to everyone.

GPT-Bi

It is an open-language model based on Transformers (GPT-2 architecture) pre-trained from scratch to experiment with text generation in Basque. The project's main goal is to advance the field of natural language processing (NLP) for minority languages, such as Basque, which have historically had fewer resources and models. The model is accessible through Hugging Face (link).

Contradictory My Dear Watson

I achieved first place in Kaggle's "Contradictory, My Dear Watson" competition, which focused on Natural Language Inference (NLI). The challenge involved developing a model capable of determining the semantic relationship between pairs of sentences in fifteen different languages, adding a high level of complexity to the problem. To achieve this result, I implemented and adapted a version of the RoBERTa model using the Hugging Face Transformers library. I optimized the training process to achieve an accuracy above 89%, demonstrating strong expertise in LLMs.

LANGUAGES

- *Spanish*: Native
- *English*: B2. I am fully capable of working in English, as demonstrated by a market study conducted in Korea, developed entirely in this language. Additionally, I regularly read scientific articles in English.
- *Basque*: B2

SKILLS

- *Autonomous Learning*: Self-taught development in Data Science demonstrated through competitive projects and contributions on GitHub, with the ability to acquire and apply new knowledge independently.
- *Adaptability*: High ability to adapt to new environments and work methodologies, demonstrated in both international projects and collaborations with different companies and sectors.
- *Results-oriented*: Proven track record of success in technical competitions and business solution development, maintaining high standards of quality and meeting objectives.
- *Technical Communication*: Experience presenting complex concepts to diverse audiences, from technical demonstrations to clients to project documentation.

VOLUNTEERING

Order of the Missionaries of Charity | Mother Teresa of Kolkata　　　　　　　　　　　　Kolkata, India
Daya Dan Orphanage, for Children with Disabilities　　　　　　　　　　　　　　　　　*Apr 2023*

Fundación Parque La Libertad　　　　　　　　　　　　　　　　　　　San José, Costa Rica
Instructor　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　*Feb 2022*