



Erik Sarriegui

Data Science

EMAIL  
erik@sarriegui.es

TELEFÓNICO  
679 75 85 68

CUMPLEAÑOS  
21 de Marzo, 2001

Ubicación  
Madrid, España

Curriculum Vitae  
[Click para descargar](#)

Web Personal  
[Click para descargar](#)

LinkedIn  
[Click para descargar](#)

GitHub  
[Click para descargar](#)

# Acerca de mi

Científico de Datos en formación con una sólida base autodidacta en Inteligencia Artificial y desarrollo de software. Mi fascinación por la capacidad de los modelos de IA para interpretar información compleja me ha llevado a especializarme en este campo a través de proyectos personales y formación continua. Actualmente estoy cursando un Máster en Ciencia de Datos, buscando combinar mi entusiasmo por el aprendizaje con una oportunidad profesional que me permita aplicar y expandir mis conocimientos en el análisis de datos y machine learning. Mi objetivo es contribuir al desarrollo de soluciones innovadoras en el campo de la IA mientras continúo creciendo profesionalmente.

## Skills



**Hugging Face**  
LLM Framework

Transfer Learning y Fine Tuning de LLMs.



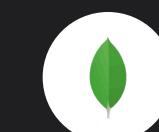
**PyTorch**  
DL Framework

Creación de Redes Neuronales a bajo nivel.



**Python**  
Programación

Lenguaje principal para el desarrollo.



**MongoDB**  
Base de Datos

Base de Datos principal, aunque puedo usar SQL.

## Educación



**Grado en Liderazgo Emprendimiento e Innovación**  
2020 - 2024  
Universidad de Mondragón



**Máster Big Data, Data Science & Inteligencia Artificial**  
2025 - 2026  
Universidad Complutense de Madrid

Acerca de mi

Portfolio



Erik Sarriegui

Data Science

EMAIL  
erik@sarriegui.es

TELEFÓNICO  
679 75 85 68

CUMPLEAÑOS  
21 de Marzo, 2001

Ubicación  
Madrid, España

Curriculum Vitae  
[Click para descargar](#)

Web Personal  
[Click para descargar](#)

LinkedIn  
[Click para descargar](#)

GitHub  
[Click para descargar](#)

# Portfolio

Acerca de mi

Portfolio

## GPT-Bi

Se trata de un modelo de lenguaje abierto basado en Transformers (arquitectura de GPT-2) preentrenado desde cero para experimentar con la generación de texto en euskera. El proyecto tiene como objetivo principal avanzar en el campo del procesamiento del lenguaje natural (PLN) para lenguas minoritarias, como el euskera, que históricamente han tenido menos recursos y modelos. El modelo está accesible a través de Hugging Face.



[Más información](#)



**GPT-Bi**

## DeepWind

Se trata de un modelo de lenguaje abierto basado en Transformers (arquitectura de GPT-2) preentrenado desde cero para experimentar con la generación de texto en euskera. El proyecto tiene como objetivo principal avanzar en el campo del procesamiento del lenguaje natural (PLN) para lenguas minoritarias, como el euskera, que históricamente han tenido menos recursos y modelos. El modelo está accesible a través de Hugging Face (link).



[Más información](#)





Erik Sarriegui

Data Science

EMAIL  
erik@sarriegui.es

TELEFÓNICO  
679 75 85 68

CUMPLEAÑOS  
21 de Marzo, 2001

Ubicación  
Madrid, España

Curriculum Vitae  
[Click para descargar](#)

Web Personal  
[Click para descargar](#)

LinkedIn  
[Click para descargar](#)

GitHub  
[Click para descargar](#)

# Portfolio

Acerca de mi

Portfolio

## MediFlow

Este proyecto se desarrolló mientras formaba parte de Vicomtech (por eso no está en mi web) y consiste en un sistema de triaje que asigna especialidades según el texto proporcionado por un paciente. El reto principal fue el dataset, que no estaba diseñado para clasificación de texto, sino que era conversacional. Se tuvieron que extraer las etiquetas con expresiones regulares y validarlas con otro modelo. Finalmente, se ajustó un XLNet sobre este conjunto de datos, logrando una precisión del ~90%.



[Más información](#)



## MediFlow



## OrangeNet

Es uno de los primeros proyectos de Deep Learning que desarrollé. Tiene como objetivo la clasificación de imágenes de naranjas para identificar la enfermedad que las afecta, si es que la hay. En caso de que no se detecte ningún tipo de enfermedad, la imagen se clasificará como "sana". Aunque posteriormente he realizado otros proyectos de Deep Learning más complejos le tengo un cariño especial a este proyecto por haber sido el primero, funcionar tan bien y estar correctamente presentado y documentado.



[Más información](#)



## OrangeNet



Erik Sarriegui

Data Science

EMAIL  
erik@sarriegui.es

TELEFÓNICO  
679 75 85 68

CUMPLEAÑOS  
21 de Marzo, 2001

Ubicación  
Madrid, España

Curriculum Vitae  
[Click para descargar](#)

Web Personal  
[Click para descargar](#)

LinkedIn  
[Click para descargar](#)

GitHub  
[Click para descargar](#)

# Otros proyectos

Acerca de mi

Portfolio

Todos los trabajos mencionados anteriormente son los más recientes y están relacionados con mi formación. Sin embargo, también he realizado otros proyectos en los que, aunque el aspecto técnico no sea tan prominente o incluso esté ausente, demuestran mi orientación a objetivos. En todos ellos he colaborado con terceros: con bares y discotecas en el caso de Uri, con dos hospitales en el caso de OpenPsy, y con todo tipo de entidades en Olympia, desde ayuntamientos hasta empresas privadas y fundaciones.



## Uri | Más información

Se trata de una app para que personas en un pub puedan votar la siguiente canción que quieren que suene. Esta aplicación tenía conexión con Spotify para gestionar la música de manera autónoma y llegó a ser utilizada en varias ocasiones en diferentes discotecas.



## OpenPsy | Más información

Es un proyecto que comencé y que se convirtió en mi Proyecto de Final de Grado. Se realizó una profunda investigación sobre la aplicación clínica de modelos del lenguaje y se desarrollaron diferentes modelos. Uno de estos fue MediFlow descrito en este documento.



## Olympia | Más información

Estuve trabajando como promotor y organizador de eventos de Magia durante 3 años. Durante este tiempo organicé más de 15 eventos, facturé más de 25 mil euros y coordiné equipos de un máximo de 8 personas. [Este es](#) un vídeo de una gala con música en directo que organicé.