Fernández Edreira **Diego**

BIOLOGIST · BIOINFORMATICIAN · DATA SCIENTIST

A Coruña, Spain

981-916-857 // 661-122-869 | diego.fedreira@udc.es | DiegoFE94 | diego-fernandez-edreira | @diego_edreira

Profile

Bioinformatician and data scientist focusing on the area of biomedicine, I am interested in all the existing omic fields. I prefer to work in a team but I have no problem with working individually. Mainly experienced as bioinformatic and data analyst in different projects related to biomedical field. Currently focused microbiome and metagenomics profiles data integration and analysis with Machine Learning frameworks. Experience working with different Machine Learning frameworks (SVM, RF, Glmnet, etc.) and dimensionality reduction techniques (FS wrapped, filter and embedded) applied in different problems (classification, clustering and regression). Daily use of Python and R. Experience in R Shiny web applications for project data visualization, through basics Shinys or complex Dashboards for dynamic interaction with projects (e.g. https://diegofedreira.shinyapps.io/shinyapp/). Workflow experience using GitHub, currently working on my personal repo (https://github.com/DiegoFE94) and the repo of my PhD lab, Machine Learning in Live Sciences (MALL).

Skills

Programming	Use Python on machine learning and image analysis. Use R on machine learning, statistics, computational biology and data visualization. Advance use of LaTeX. Advance use of High Performance Computing (slurm resource management). Basic notions of hadoop and spark.
Web	R Shiny with R. Flask with python.
Others	Git and Office.
Languages	Mother tongue Spanish, Galician (CELGA 4) and fluent English.

Work Experience

Bioinformatician at Universidade da Coruña

A Coruña. Spain

6 MONTHS

May 2023-October 2023

- During this time I was working as a bioinformatician at the University of A Coruña for the Centre for Research in Information and Communication Technologies (CITIC). During this period I was working on an application for the detection of polyps in colonoscopy using Deep Learning. As well as the integration of the models generated in web applications and mobile applications. Throughout the 6 months I was working with python for image processing and model training (opencv, tensorflow, keras, scikit-learn...). For the web development I used flask and for the Android application I used TensorFlow Lite and Dart.

Bioinformatician at Fundación Profesor Novoa Santos

A Coruña. Spain

12 MONTHS

March 2022-March 2023

- During these 12 months I was working as a bioinformatics researcher in a project called Practicum Direct (IN845D-2020/03), focused on analysis, validation and testing of ad hoc simulation models designed with open data and predictive models based on real data. I had several duties: review of the state of the art in terms of availability of open omics and clinical data, implementation of predictive models on real open data in R/Python and create dashboards with R and Shiny for the manipulation and presentation of results in an interactive way.

Bioinformatician at Fundación Profesor Novoa Santos

A Coruña. Spain

6 MONTHS

April 2021-October 2021

- During these 6 months I was working as a bioinformatics researcher in a colorectal cancer project. During this period I gained experience in metagenomic data processing, as well as, developing Machine Learning and Artificial Intelligence pipelines. Also I gained experience in the presentation and visualization of complex data.

Education

Ph.D. in Information and Communication Technologies

A Coruña. Spain

University of A Coruña *2021-Currently*

- The proposal of this thesis is the development of a research based on different fields such as Artificial Intelligence, Medicine and Bioinformatics. In this research, based on Machine Learning algorithms, a new analysis methodology capable of integrating and extracting the maximum amount of information possible from multi-omics data related to different clinical conditions will be developed. This thesis is being conducted by making use of open access data made available to the research public by a multitude of projects and institutions.

Fundamentals of Deep Learning

A Coruña. Spain

NVIDIA *2022*

- This workshop focuses on teaching the fundamental techniques and tools required to train a deep learning model, enhance datasets through data augmentation to improve models accuracy and leverage transfer learning between models to achieve efficient results with less data and computation In summary, through this course I gained experience with common deep learning data types and model architectures.

IX International Workshop on Medical Imaging, Capture and Integration of Clinical Data

A Coruña. Spain

University of A Coruña - RNASA-IMEDIR *2021*

- This whorkshop focuses on teaching the basics of applied medical imaging (creation of models for image segmentation, new technologies in radiological practice, deep learning for medical image classification, etc.). It also teaches methodologies for the correct capture of medical data (omic data, clinical data, etc.) in order to follow good practice.

Course for the use of Supercomputer Centre

Santiago de Compostela. Spain

CESGA *2020*

- This course focuses on teaching the organisation of the Finisterrae II Supercomputer. It points out the resources available, the libraries and compilers necessary to command processes. It also explains in detail how to make efficient use of the queuing system available.

Bioinformatics for Health Sciences Master's degree

A Coruña. Spain

University of A Coruña *2018-2020*

- Learning and reinforcement of the use of bioinformatic tools, image analysis, biostatistics, Machine Leraning methodology and computational intelligence for high dimensional data. Thesis master directed by Dr. Carlos Fernández Lozano with the title "Machine learning based analysis of metagenomic profiles for the stratification of patients affected by Type I Diabetes" and developed in R and partially executed in in High Performance Computers. This thesis obtained an honors mention and the final results have been published in the high impact JCR journal Expert Systems with Applications.

Biology degree

A Coruña. Spain

University of A Coruña *2012-2018*

- Main interest in biochemistry, genetics and microbiology fields. Thesis degree directed by Dr. María Isabel González Siso with title "Isolation and characterization of clones with hydrolase activity from a metagenomic library of thermal water". Thesis on the bioprospection of new enzymes in metagenomic libraries with the aim of commercializing new lipases, through species isolation, lipolytic activity assays and High Throughput Sequencing (HTS) techniques.

Publications

2024 Fernández-Edreira, D., Liñares-Blanco, J., Patricia, V., & Fernandez-Lozano, C. (2024). **VIBES: A consensus subtyping of the vaginal microbiota reveals novel classification criteria.** In *Computational and Structural Biotechnology Journal* (Vol 23, p.148).

2024 GOLOB, Jonathan L., et al. (2024). **Microbiome preterm birth DREAM challenge: Crowdsourcing machine learning approaches to advance preterm birth research.** In *Cell Reports Medicine* (Vol 5, No.1, p.101350).

2021 Fernández-Edreira, D., Liñares-Blanco, J., & Fernandez-Lozano, C. (2021). **Machine Learning analysis of the human infant gut microbiome identifies influential species in type 1 diabetes.** In *Expert Systems with Applications* (Vol 185, p. 115648).

2021 Fernández-Edreira, D., Liñares-Blanco,J.,& Fernández-Lozano, C. (2021). **Machine Learning-based analysis of metagenomic profiles for the stratification of patients affected by type I Diabetes.** In *Proceedings of the MOL2NET'21, Conference on Molecular, Biomedical & Computational Sciences and Engineering, 7th ed., 25 January 2021–30 January 2022, Multidisciplinary Digital Publishing Institute Proceedings.*

2021 Edreira, D. F., Blanco, J. L., Galdo, B. C., García, V. M. M., & Pazos, A. (2021). **La Inteligencia Artificial como herramienta para la gestión y explotación de datos, informaciones y conocimientos biomédicos en entornos "Big Data" en la nube.** In *I+ S: Revista de la Sociedad Española de Informática y Salud* (No. 143, p. 30-39)

2020 Fernández-Edreira, D., Liñares-Blanco,J.,& Fernández-Lozano, C. (2020) **Identification of Pevotella, Anaerotruncus and Eubacterium Genera by Machine Learning Analysis of Metagenomic Profiles for Stratificaction of Patients Affected by Type I Diabetes.** In *Multidisciplinary Digital Publishing Institute Proceedings* (Vol. 54, No. 1, p. 50).

International Conferences

MOL2NET'21. Conference on Molecular, Biomedical & Computational Sciences and Engineering, 7th ed: AIMEDIC-08. *A Coruña, Spain*

Poster speaker and paper publication *November 2021*

- "Machine Learning-based analysis of metagenomic profiles for the stratification of patients affected by type I Diabetes"

III XoveTIC Congress *A Coruña, Spain*

Poster speaker and paper publication *September 2020*

- "Identification of Prevotella, Anaerotruncus and Eubacterium Genera by Machine Learning Analysis of Metagenomic Profiles for Stratification of Patients Affected by Type I Diabetes"

Cancer and microbiota: evidences *A Coruña, Spain*

Assistance *October 2020*

- Conference on the role of microbiota in human health. More specifically in the types of cancer that have more impact on society nowadays. Artificial intelligence as an effective method in the diagnosis and prognosis of some of these types of cancer was also discussed.

Microbiota: What's cooking? *A Coruña, Spain*

Assistance *October 2019*

- Pioneering multidisciplinary meeting point that serves as a focus of inspiration for research actions that generate new protocols of care action, including the creation of patents on diagnostic devices, prognosis, therapy, control and care monitoring.

Organisation of Activities

2023 Teacher of **Introduction to Applied Artificial Intelligence with Python (2023)** In *University of A Coruña, A Coruña. Spain. 9 to 30 January*

2021 Organising Committee of **IX Workshop Internacional en Imagen Médica Captura e Integración de Datos Clínicos - Principios básicos de Inteligencia Artificial y computación avanzada en el ámbito Biomédico.** In *University of A Coruña, A Coruña. Spain. 28 June to 8 July*