Rosario **Medina-Rodríguez**
PhD. in Computer Science | Data Scientist

🔗 rosario1316@gmail.com 📱 +51 951520726 📍 Lima, Perú
in linkedin.com/in/ramedinar ⬡ github.com/rmedinar 🐦 @rmedinar
👥 Member of:
ⓘ ↗ Perú - RENACYT Researcher
ⓘ ↗ IA-PUCP, Artificial Intelligence Group
ⓘ ↗ *Scientific Association for Research, Development and Innovation* - PERU I+D+I

🎓 EDUCATION

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
PHD. IN ENGINEERING
August 2022 | Lima, Perú

UNIVERSITY OF SÃO PAULO
MASTER IN COMPUTER SCIENCE
July 2012 | São Paulo, Brazil

BS IN SYSTEMS ENGINEERING
December 2009 | Arequipa, Perú

☰ SKILLS

Programming:
Python • C/C++ • LaTeX • Matlab
R • Java • Perl • HTML

Operative Systems:
Windows • Linux • Mac OS X

Databases:
PostgreSQL • PostGIS

Image Libraries:
Scikit-image, • PIL • OpenCV •
ImageMagick • ImageJ • Pov-Ray

Multispectral Image Processing:
QGis • Rasterio • GDAL • Earthpy •
Fiona • Libtiff

Deep Learning Frameworks:
Pytorch • Tensorflow

🖥 COURSEWORK

Deep Learning Specialization by
DeepLearning.AI on Coursera
**International Summer School
on Deep Learning** - GDANSK
University (July 5 - 9, 2021).

🗛 LANGUAGES

Spanish	●	●	●	●	●
English	●	●	●	●	○
Portuguese	●	●	●	○	○
French	●	○	○	○	○

👤 OTHER ACTIVITIES

SWIMMING
VOLLEYBALL, RUN, 🚲
🎧 80'S MUSIC
WATCH MOVIES, DANCE

💼 EXPERIENCE

FARMACIAS PERUANAS | DATA SCIENTIST
April 2022 – Today | Lima, Perú

Python | SQL | PowerBi | BigQuery | Google Cloud Platform | Google Data Studio

WORLD BANK | SHORT-TERM CONSULTANT
June 2022 | Lima, Perú

Python | OpenCV | Tesseract | Pytesseract | PyGrok

**NATIONAL AEROSPACE RESEARCH AND DEVELOPMENT COMISSION
CONIDA** | REMOTE SENSING AND DEEP LEARNING RESEARCHER
(July 2020 - Dec 2020) (June 2021 - Dec 2021) | Lima, Perú

Python | LaTeX | QGis | PostGis | GDAL | rasterio | earthpy | fiona | pytorch | google colab
wandb | geopandas | gradio | shapely

RESEARCH CENTER OF UNIVERSIDAD DEL PACÍFICO
↗ **CIUP** | RESEARCH ASSISTANT
April 2020 – Nov 2021 | Lima, Perú

Python | Plot.ly | LaTeX | scikit-learn | geopandas | PostgreSQL | QGis | KarateClub | Weka

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
↗ **PUCP** | TEACHING ASSISTANT (DATA VISUALIZATION)
Sep 2019 - June 2021 | Lima, Perú

Python | matplotlib | geopandas | dash | D3 | Plot.ly | jupyter notebook

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ | MASTER THESIS
CO-SUPERVISOR
Aug 2020 - Dec 2020 | Lima, Perú

Deep Learning | Python | LaTeX

UNIVERSIDAD DEL PACÍFICO | RESEARCH ASSISTANT
May 2019 - Sep 2019 | Lima, Perú

Python | Plot.ly | LaTeX | scikit-learn | Weka | keras | tensorflow

PERÚ I+D+I | RESEARCH ASSISTANT (IMAGE PROCESSING - MACHINE
LEARNING)
July 2016 - July 2017 | Lima, Perú

Python | OpenCV | Raspberry Pi | tkinter

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ | RESEARCH ASSISTANT
(PROJECT LEADER)
Feb 2014 - March 2016 | Lima, Perú

C++ | Weka | OpenCV | Netbeans | QT | LaTeX

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ | TEACHING ASSISTANT
(COMPUTER SCIENCE APPLICATIONS)
Mar 2014 - Dec 2015 | Lima, Perú

C++ | Python

INSTITUTO DE BIOCIÊNCIAS, UNIVERSIDAD DE SÃO PAULO | RESEARCH
ASSISTANT (SEARCH FOR PATTERNS IN DNA AND VISUALIZATION)
Sep 2012 - Aug 2013 | São Paulo, Brazil

Java | Perl | R | MySQL

GLOBAL SYSTEM & CONSULTING | RESEARCHER IN DOCUMENT IMAGE
PROCESSING AND RECOGNITION
Feb 2009 - July 2009 | Arequipa, Perú

Java | Netbeans | JRE | ImageJ | ImageMagick | Tesseract

📖 PUBLICATIONS (☑ GOOGLE SCHOLAR PROFILE)

November 2021 **Medina-Rodríguez R**, Beltrán-Castañón C and Hashimoto RF. *An Approach to Growth Delimitation of Straight Line Segment Classifiers Based on a Minimum Bounding Box.* Entropy. 2021 Nov 19;23(11):1541.

October 2021 **Medina-Rodriguez, R.**, Fuentes Navarro, E. L., Beltrán-Castañón, C., Nunez-del-Prado, M., and Alatrista-Salas, H. *Mobiloscope: A Technological Solution for Early Mastitis Detection in Dairy Cattle.* IEEE Latin America Transactions, 20(1), 117-125. 2021.

July 2021 Roy Melendez, César Beltrán Castañón and **Rosario Medina-Rodríguez**. *Sperm Cell Segmentation in Digital Micrographs based on Convolutional Neural Networks using U-Net Architecture.* 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS). Aveiro, Portugal. 7-9 June, 2021

April 2020 **R. Medina-Rodríguez**, Alvaro Talavera, Martín Hernani-Merino, Juan Lazo-Lazo and Mazzon, Jose Alfonso. *Global Brand Perception based on Social Prestige, Credibility and Social Responsibility: A Clustering Approach.* 6th International Conference on Information Management and Big Data. SIMBig 2019. Lima,Perú. August 21-23, 2019

November 2018 **R. Medina-Rodríguez**, R. F. Hashimoto and C. Beltrán. *Evaluation of the Impact of Initial Positions obtained by Clustering Algorithms on the Straight Line Segments Classifier.* 5th IEEE Latin American Conference on Computational Intelligence, LACCI 2018, Guadalajara, Mexico, November 7-9, 2018.

June 2017 Jorge Vargas Florez, **R. Medina-Rodríguez** and Rafael Alva-Cabrera. *A Clustering Optimization Approach for Disaster Relief Delivery: A Case Study in Lima-Peru. (Revised Selected Papers).* Information Management and Big Data - Second Annual International Symposium, SIMBig 2015, Cusco, Peru, September 2-4, 2015, and Third Annual International Symposium, SIMBig 2016, Cusco, Peru, September 1-3, 2016, Revised Selected Papers. Springer

November 2016 Jorge Vargas Florez, **R. Medina-Rodríguez** and Rafael Alva-Cabrera. *A Clustering Optimization Approach for Disaster Relief Delivery: A Case Study in Lima-Peru.* Proceedings of the 3rd Annual International Symposium on Information Management and Big Data - SIMBig 2016, Cusco, Peru, September 1-3, 2016.

October 2015 D.Gárate, **R. Medina-Rodríguez**, F. Incahuanaco and C. Beltrán. *A heuristic model for determining the sperm motility grade by video tracking.* XLI Conferencia Latinoamericana en Informática (CLEI 2015). IEEE Xplorer

September 2015 **R. Medina-Rodríguez**, H. Alatrista, L. Guzmán and C. Beltrán. *Sperm cells segmentation in micrographic images through Lambertian reflectance model.* 16th International Conference on Computer Analysis of Images and Patterns (CAIP 2015). Springer LNCS.

May 2014 G. Cruz, C. Metcalfe, N. de Setta, E. Cruz, A. Vieira, **R. Medina-Rodríguez**, and M. Van Sluys. *Virus-Like Attachment Sites and Plastic CpG Islands: Landmarks of Diversity in Plant Del Retrotransposons.* PloS one, 9 (5), pp. 97099-97099.

November 2013 **R. Medina-Rodríguez**, J. P. Mena-Chalco, and H. Bernedo Cordova. *DNA Symphony: A new method to Represent Genomic Sequences.* International Conference of the AB3C (X-Meeting) & Brazilian Symposium on Bioinformatics (BSB).

August 2013 **R. Medina-Rodríguez** and R. F. Hashimoto. *Evolutive Algorithms Applied To The Straight Line Segment Classifier.* XXVI SIBGRAPI 2013 – Workshop of Theses and Dissertations - Conference on Graphics, Patterns and Images.

August 2012 **R. Medina-Rodríguez** and R. F. Hashimoto. *Algoritmos evolutivos aplicados ao classificador baseado em segmentos de reta.* Master Dissertation – University of Sao Paulo.

August 2011 **R. Medina-Rodríguez** and R. F. Hashimoto. *Combining Dialectical Optimization and Gradient Descent Methods for Improving the Accuracy of Straight Line Segment Classifiers.* XXIV SIBGRAPI 2011 – Conference on Graphics, Patterns and Images.

December 2008 **R. Medina-Rodríguez** and J. P. Mena-Chalco. *Firmas Genéticas en Secuencias de ADN: Un análisis en regiones codificantes y no codificantes de proteínas.* I Simposio Peruano de Computación Gráfica y Procesamiento de Imágenes (SCGI).

🖥 RESEARCH ACTIVITY DESCRIPTION

SHORT-TERM CONSULTANT AT WORLD BANK

Objective: The aim of this project was to extract text information from Whatsapp's history and screenshots. This information was part of post-training activities for the reopening of schools in Perú.

Responsibilities:
- Preprocess images from whatsapp screenshots using OpenCV.
- Extract text information from images using Tesseract OCR.
- Extract users information and answers from Whatsapp's history chats.
- Create a database with phone numbers or names, time the message was seen and their answers.

Python | OpenCV | Pytesseract | Pygrok

PILOT STUDY FOR THE ESTIMATION OF COCA-LEAF PRODUCTION REGIONS IN PICHARI (2020) AND HUALLAGA-PERÚ (2021) BASED ON DEEP LEARNING AND USING SPOT6/7 MULTISPECTRAL IMAGES.

Objective: The aim of these projects was the estimation of coca-leaf growing regions in Pichari and Huallaga-Perú based on deep learning and using SPOT6/7 images .

Responsibilities:
- Project management and development, including meetings with the specialists and organizing tasks with one team member.
- Literature revision for multispectral image analysis and prioritized crops detection.
- Design and propose a methodology to achieve the stated goal.
- Multi-spectral image processing: vegetation indices calculation (NDVI, GNDVI), polygons to rasters conversion (obtain ground truth), cloud removal, and cutting the images into tiles to create an image dataset.
- Define a deep learning architecture to train a model to segment the coca-leaf production regions.
- Perform experiments with the proposed dataset and tune the hyperparameters.
- Reconstruct the predicted tiles into a geo-referenced image to visualize the results.
- Generate polygons from the geo-referenced predictions tiles. These polygons could be opened in QGis to get a shapefile.
- Implement a GUI interface in a jupyter notebook with Gradio to show the results.
- Writing technical and scientific documentation.

Python | LaTeX | QGis | GDAL | rasterio | earthpy | fiona | Pytorch | Wandb | Gradio | geopandas | albumentations

FINANCIAL INCLUSION

Objective: The aim of this project is the generation of bank users' profiles which allows the characterization of new potential clients coming from mobile users.

Responsibilities:
- Extract information from a PostgreSQL database with bank transactions.
- Extract information from telephone users stored in plain files and extract all the geo-positions in several comprised files based on a user id.
- Generate the mobility models based on Mobility Markov Chains for bank and telephone users and compute a matrix distance between both users' models.
- Apply the Co-clustering technique on the distance matrix to group telephone users with bank users with the aim to assign potential telephone users to a bank group.
- Define metrics with entropy and density to determine the best number of clusters to be used.
- Generate a visualization and extract statistics from the groups obtained.
- Obtain a vector representation using *Graph2Vec* to obtain a vector representation based on the mobility models for bank and telephone users and the associated social class for the bank users.
- Train conventional machine learning classifiers with the generated vectors as features and their respective social classes as a label. The aim is to predict the telephone users' social class based on their mobility models.
- Writing technical and scientific documentation.

Python | scikit-learn | PostGis | Plot.ly | SQL | Weka | LaTeX

DEVELOPMENT OF A DECISION SUPPORT SYSTEM WITH THE USE OF DRONES FOR THE MANAGEMENT, OPTIMIZATION AND CONTROL OF HIGH ANDEAN AGRICULTURAL SYSTEMS IN THE FACE OF CLIMATE CHANGE.

Objective: The aim of this project is the implementation of a decision-making support system based on the exploitation of information captured by aerial vehicles.

Responsibilities:
- Data extraction from non-relational databases (Mongo DB).
- Application of clustering techniques on spatial and transactional data.
- Writing technical and scientific documentation.

Python | OpenCV | PostGis | LaTeX

LOW-COST ``SMART" MICROSCOPE FOR SMARTPHONES: TECHNOLOGICAL SOLUTION FOR THE DETECTION OF DISEASES IN CATTLE.

Objective: The aim of this project is to build a portable microscope with the help of a raspberry pi and a computer application to identify and count somatic cells in milk samples. In this way, it is possible to quickly and early identify possible mastitis in cattle, helping farmers and companies in the milk management sector.

Responsibilities:
> Installation of a Linux-based distribution in the raspberry Pi.
> Implement an algorithm to segment somatic cells from cattle milk based on image processing techniques.
> Implement a GUI using Tkinter to capture an image with the microscope attached to the raspberry pi then process the image to segment the cells and count them to diagnose mastitis based on the number of cells found in the sample.
> Writing technical and scientific documentation.

`Python` `scikit-image` `OpenCV` `Raspberry Pi` `tkinter`

COMPUTATIONAL PLATFORM FOR AUTOMATIC ANALYSIS OF MICROGRAPHIC IMAGES FOR THE EVALUATION OF MALE FERTILITY.

Objective: Develop a computational platform conformed of three modules which allows the interaction with a microscope to an automatic analysis of the concentration, morphology, and motility of sperm cells so that male fertility can be determined.

Responsibilities:
> Project management and task assignation to members of the team and coordinate meetings with the reproduction center involved in the project.
> Implementation of the morphology module, which consists of image acquisition at the laboratory, image preprocessing, and segmentation of the sperm cell from the background using color spaces. Implement an algorithm to segment each part of the sperm cell based on image processing techniques and mathematical morphology.
> Implementation of an algorithm to extract features from each sperm part following WHO requirements for morphology analysis.
> Train three conventional machine learning classifiers to determine which sperm cell has a normal or abnormal morphology in an image captured by a microscope.
> This project concluded with a CASA software registration at the (National Institute for the Defense of Competition and the Protection of Intellectual Property (INDECOPI).
> Writing technical and scientific documentation.

`OpenCV` `LaTeX` `Weka` `C++` `Qt`

AUTOMATIC ANALYSIS BY DIGITAL IMAGE PROCESSING TO DETERMINE THE SEVERITY DEGREE OF ``YELLOW RUST" ON COFFEE LEAVES.

Objective: The aim of this project is to develop an intelligent component to determine the severity of coffee leaf rust through automatic analysis of leaf images captured by a cell phone camera.

Responsibilities:
> Perform experiments to classify images between healthy and diseased coffee leaves applying conventional machine learning classifiers.

`Weka` `LaTeX` `C++`

AUTOMATIC ANALYSIS BY DIGITAL IMAGE PROCESSING TO DETERMINE THE SEVERITY DEGREE OF ``YELLOW RUST" ON COFFEE LEAVES.

Objective: The aim of this project is to develop an intelligent component to determine the severity of coffee leaf rust through automatic analysis of leaf images captured by a cell phone camera.

Responsibilities:
> Perform experiments to classify images between healthy and diseased coffee leaves applying conventional machine learning classifiers.

`Weka` `LaTeX` `C++`

RECOVERY OF DOCUMENTARY INFORMATION AND ADMINISTRATION.

Objective: Processing scanned documents in white background.

Responsibilities:
> Implement an algorithm to remove scanned blank pages.
> Implement an applet in Java to edit scanned documents by applying rotations, zoom, skew, and removing stains by size and selected regions.
> Apply an OCR to a previously selected area.

`Java` `ImageMagick` `ImageJ` `Tesseract`

❝ REFERENCES

Ronaldo Fumio Hashimoto
Associate Professor, IME-USP (Brasil)
@ ronaldo@ime.usp.br

César Beltrán Castañón
Principal Professor, PUCP (Perú)
@ cbeltran@pucp.pe

Miguel Nuñez del Prado Cortez
Researcher - Data Scientist, (Perú)
@ miguel.nunez@peruidi.com