GRISEL ELIANA JAEL QUISPE ARAMAYO

+33 7 45 49 81 20 ⋄ Saint Etienne, France

griseleliana79@gmail.com ⋄ linkedin.com/in/griselquispe ⋄ griselquispe.github.io

EDUCATION

Master in Machine Learning and Data Mining, Université Jean Monnet ⸻ Expected 2025
Relevant Coursework: Advanced Machine Learning, Deep Learning, Data Mining for Big Data, Probabilistic Graphical Models, Advanced algorithms and programming.

Engineering Degree in Mechatronics, Bolivian Catholic University ⸻ 2014 - 2021

SKILLS

Technical Skills: Python, Data preprocessing, Machine Learning algorithms, Deep Learning frameworks (TensorFlow, PyTorch), Statistical analysis, Natural Language Processing, Computer Vision, Model evaluation and validation, Big Data processing (Hadoop, Spark).

Soft Skills: Analytical Thinking and Problem Solving, Curiosity and Continuous Learning, Collaboration and Teamwork, Time Management and Prioritization, Communication and Presentation Skills.

Languages: English (C1), French (A1), Spanish (Native)

EXPERIENCE

Machine Learning Research Intern ⸻ Apr 2024 - Aug 2024
HidroSciences Montpellier, TETIS Lab ⸻ *Montpellier, France* In this project

- Designed and implemented a data preprocessing pipeline to extract relevant segments from scientific documents, reducing data cleaning time by 50%. Utilized the pre-trained GROBID model (Hugging Face) and regular expressions (Regex).
- Developed and deployed a supervised machine learning model for text classification, achieving 98% precision and 94% recall. Leveraged techniques such as BERT, SVM with Bag of Words, and cross-validation, implemented using PyTorch.
- Transformed unstructured data into structured data using Named Entity Recognition (NER) to extract and categorize text entities, employing Regex and EntityRuler.
- Engineered extracted entity data with a focus on explainability through bar charts, box plots, and correlation analysis, utilizing NumPy, scikit-learn, Matplotlib, and Pandas.

Co-Researcher Darwin Team ⸻ Dec 2022 - March 2023
CIDIMEC (Center of research develoment and innovative in Mechatronics Engineering) ⸻ *La Paz, Bolivia*

- Preprocessed image datasets using techniques such as resizing, normalization, grayscale conversion, dataset augmentation (150% increase), denoising with Gaussian filters, and histogram equalization. Tools: OpenCV, Matplotlib, NumPy.
- Researched and implemented state-of-the-art deep learning techniques, developing a convolutional neural network with an encoder-decoder structure. Achieved an F1 score of 94%. Tools: TensorFlow, scikit-learn.

PROJECTS

SUPERSTORE Dataset - KAGGLE Studying this data from a supermarket, I hope to find ways to improve their business, like figuring out which products are most popular or which areas need more attention.

- Performed feature engineering, including outlier removal, correlation analysis, one-hot encoding, and standardization using the min-max scaling technique.
- Implemented K-Means clustering to segment customers into three groups based on profit and sales data.

- Applied the Apriori algorithm to discover transactional patterns, identifying the most demanded item categories, regions, and customer segments.

- Developed a decision tree model to predict item profitability, achieving an accuracy of 95.66%.

Credit Card Approval Classification using SVM This project aims to automate the credit card approval process by employing machine learning algorithms.

- Conducted feature engineering, including data distribution analysis, one-hot encoding, feature importance determination using Random Forest, and dataset standardization.

- Trained and optimized a Support Vector Machine (SVM) model using cross-validation for hyperparameter tuning. Utilized PCA for visualization, achieving a precision of 97%. Tools: scikit-learn.

- Designed and implemented a data pipeline following Object-Oriented Programming (OOP) principles.

Poverty prediction Model This project was my thesis in Engineering School, focus in predict the poverty level in Bolivia based in analysis of Satellite Imagery

- Preprocessed satelite images using techniques such as resizing, normalization, grayscale conversion, dataset augmentation (30% increase), denoising with Gaussian filters. Tools: OpenCV, Matplotlib Numpy.

- Implemented deep learning techniques, comparing between state of the art architectures (VGG, AlexNet, ResNet). Achieved a precision of 97%. Tools: TensorFlow, Scikit-learn Keras.

EXTRA-CURRICULAR ACTIVITIES

- Co-founded the first Student Energy chapter in Bolivia a connected with other international chapters such as Canada, Trinidad & Tobago. As vice president, recruited volunteers to organize activities such as debates or discussions about youth empowerment in the energy sector. I had the privilege of representing Bolivia at the SES 2019 climate summit, giving a talk on "The incidence of young people in the formulation of energy policies in Bolivia".

- Musician in the women group **Pankaritas del Viento.**

REFERENCES

- Sarah VALENTIN.

  TETIS LAB/Researcher

  sarah.valentin@cirad.fr