Jeanfranco David Farfan Escobedo

jeanfrancodfe@gmail.com github.com/jeanfranc0 personal web

Education

2021-2025 **PhD in Computer Science**, UNICAMP, Brazil. Thesis: Deep Learning.

2019-2021 **M.Sc. in Computer Science**, UNICAMP, Brazil. Thesis: Conversational Systems.

2012-2017 **B.Sc./Eng. in Computer and Systems Engineering**, UNSAAC, Peru. Ranked in the top fifth percentile throughout academic years. Thesis: Computer Vision.

Awards and Recognitions

2021 **Shell Oil Company**, PhD Scholarship.

2019 **Sinch Latin America**, Master's Scholarship.

2018 **First Place** in the AgroHack hackathon. Developed a mobile app for recognition, monitoring, and support of plant diseases using IBM Cloud services. Cusco, Peru.

2014 **UNSAAC**, Winner of the UNSAAC Technology Fair with a project using PIC 16F84A.

Languages and Technologies

Machine/Deep Learning (ML/DL), Natural Language Processing (NLP), Computer Vision (CV), Text-To-Speech Processing (TTS), Content-Based Information Retrieval (CBIR), OpenCV, Scikit-Learn, Matplotlib, Pandas, Keras, PyTorch, TensorFlow, HuggingFace, LLM.

Projects, Experience, and Publications

2026-Present **Machine Learning Researcher, Involves LATAM**

- Architected and deployed a scalable machine learning framework to automate the labeling of hundreds of highly imbalanced SKU datasets, bridging academic research concepts into a production-ready environment.
- Engineered an end-to-end ML pipeline that processes hundreds of thousands of SKUs daily, maintaining high model performance despite severe data sparsity.

2025-Present **Postdoctoral Researcher, UNICAMP (Sponsored by Shell Oil). Principal Investigator: Alessandra Davolio**

- Developed ML-based frameworks for automated interpretation of large-scale 3D/4D seismic data.Also, Applied supervised/unsupervised learning to detect reservoir dynamics using dRMS and dAmp attributes.
- Designed QC and confidence-estimation methods to improve reliability of time-lapse seismic signals. Bridged machine learning outputs with reservoir engineering decision-making.

2024-2026 **Software Developer in TTS, Samsung Instituto De Desenvolvimento (SiDi), Brazil.**

- Worked with global teams to integrate **Bixby Talkback** and **Samsung Interpreter** into Samsung's ecosystem, ensuring support for Brazilian Portuguese, European Portuguese, and Mexican Spanish across devices in over 200 countries.
- Enhanced the TTS model for Bixby and Samsung Interpreter by curating, creating, and reviewing audio data, adjusting training hyperparameters, and validating outputs to reduce prosody errors, mispronunciations, and a robotic voice.

2021-2025 **DL Researcher, UNICAMP (Sponsored by Shell Oil). Advisor: Anderson Rocha**

- Developed an efficient end-to-end proxy to replace numerical simulators during data assimilation. Addressed high dimensionality, limited sample sizes, and accurate reconstruction after dimensionality reduction: **Paper1**, **Paper2**, **Paper3**, **Paper4**, **Paper5**.
- Implemented a 3D CNN for dimensionality reduction and simulator response replication. Augmented dataset using GANs; applied VAE for reconstructing 3D geological uncertainties post-reduction.
- Achieved significant improvements: reduced simulation times by 52.45% while maintaining high accuracy. Replication error was only 3.4%, and reconstruction showed 99% correlation with original 3D uncertainties. Enhanced data assimilation and geological modeling efficiency.

2021 **ML/DL Instructor, University of Engineering and Technology (UTEC), Lima, Peru.**

- Delivered a comprehensive ML/DL course integrating theory and applied projects.

2019-2021 **NLP Researcher, UNICAMP (Sponsored by Sinch). Advisor: Julio Cesar dos Reis**

- Worked on improving ML and NLP models at Sinch Latin America, reducing labeled data requirements and optimizing context management in dialogue flows: **Paper6**, **Paper7**.
- Proposed active learning technique to minimize labeled data using CNNs; developed window-based approach for context management, focusing on prior utterances.
- Conducted comparative analysis between pretrained BERT in Portuguese and proposed methodology.
- Reduced time and cost of data labeling, achieving 2.5% improvement in intent classification.

2017-2018 **CV Researcher, UNSAAC, Cusco, Peru. Advisor: Lauro Enciso Rodas**

- Worked on image recognition of historical buildings in Cusco, comparing bag-of-words and CNN methods: **Paper8**, **Paper9**.
- Implemented image recognition using OpenCV, Scikit-Learn, and TensorFlow.
- Improved accuracy and efficiency via CNN application vs traditional methods.

2016-2018 **Software Engineer, Money Express Center S.A.C., Cusco, Peru.**

- Developed a new monetary transaction system to improve efficiency for desktop and web applications.
- Built system using Java and Node.js with real-time transaction visualization via Socket technologies.
- Achieved 15% reduction in customer service time with instant transaction updates.

Community Service

2023-Present Volunteer, IEEE INTERCON Reviewer.

2021 Volunteer, Data Science Woman Peru.

2012-2016 Volunteer, ACM chapter Cusco.

Languages

Spanish **Mother tongue**.

Portuguese **Fluent**.

English **Advanced**.