Federico Ferreyra

Argentinian & French (EU) citizen
E-mail: federicoferreyra65@gmail.com
Phone number: +54 9 11 5570 5924

Porfolio: fede6590.github.io
LinkedIn: federico-ferreyra
GitHub: fede6590

Machine Learning Engineer - Data Scientist

Audio Engineering with +4 years of experience and specializing in Machine Learning (+1 year). I've completed multiple projects using **Python**, **Numpy**, **AWS**, **Keras**, **Tensorflow**, **Pandas**, **Scikit-Learn**, **Docker** and the basics in **SQL** among others. I also have +3 years with **Python** applied to Digital Signal Processing (DSP) and Digital Audio. I had a chance to mentor students in applied mathematics, algebra and physics, in addition to my experience in commercial audio consulting. Currently working on AI projects at Anyone AI and as a Data Science Tutor at Coderhouse.

Work Experience

Data Science Tutor at Coderhouse

Feb 2023 to date

- Respond to students' queries.
- Correct challenges and intermediate deliverables of the final project.
- Follow up with the students: accompany them in their learning process throughout the course, motivate them to meet the challenges and offer support for any difficulties they may encounter.

Machine Learning Engineer at Anyone AI

Apr 2022 to date

- Object detection for in-store inventory management: Scalable microservices architecture using Docker to implement Flask, Redis, and a fine-tuned model (based on yolov5) trained on AWS EC2 Server with a custom dataset prepared on S3 storage. Ready to deploy with a login security layer to access the API and run inferences. Detects missing object in store shelves and return the image with a Gaussian filter for visualization.
- Products review classification: Analyze sentiment in product reviews for a movie streaming service. Manipulated data that is not in a traditional format, pre-processed it, and vectorized text data using BoW and TF IDF. Trained a word embedding and used it as a vectorization source for the data. Trained a sentiment analysis model to detect positive and negative opinions for movie reviews.
- Image Classification for E-Commerce: Predict vehicle make and model from unstructured e-commerce images. Trained on a pre-built dataset of 196 classes. Visualized and cleaned the dataset, pre-processed and augmented data, and trained a fine-grained classification model using convolutional neural networks achieving 82% accuracy in the prediction of make and model combined. Deployed in AWS instances using Docker, using an API-based web-service application.
- Home Credit Risk Analysis: Predicted whether a person applying for a home credit will be able to repay their debt or not. Manipulated and visualized data, and performed data pre-processing for a large dataset of +350,000 transactions. Trained many supervised models were achieving +0.72 ROC AUC. Models used were DecisionTree, XGBoost, and LightGBM.
- Salary Prediction Model: The goal was to predict salary levels based on historical data for sports players. Collected and analyzed data via an API using Python and Pandas. The original data was unbalanced. Cleaned up data, generated additional fields, stored and created a base dataset. Manipulated and visualized data. Performed feature engineering and standardization. Selected evaluation metrics and baseline models. Trained a linear regression model, achieving an F1 score of 76%.

Tech stack: Python, Numpy, SciPy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow, Keras, PyTorch, Docker, Flak, Redis, AWS (EC2 & S3), Bash, Git, OpenCV, Postman

Math & Physics Teacher 2014 to 2022

- Prepared +50 students for college admission providing personalized mentoring according to their career goals. From mathematics, algebra, or physics to DELF French preparation.

Audio Engineering Consultant 2014 to 2022

- Designed and deployed audio technology in entertainment and food industries providing consultancy on client's commercial premises and their needs.
- Each project was planned following best practices and recommendations to achieve an optimal acoustic result based on the client's requirements and physical characteristics.

Projects

- **Speaker diarization with Siamese Neural Networks (SNN)** 2022
  Simple implementation using TensorFlow and Keras to build a twin architecture and calculate dissimilarities between speakers as final project of a Seminar (Neural networks applied to Musical Information Retrieval).

Tech stack: Python, Numpy, SciPy, Librosa, TensorFlow, Keras

- **Adaptive filter for signal prediction with noise** 2019
  Python implementation of an adaptive Kalman filter, aiming at predicting the sound pressure level SPL (dBSPL) at a certain distance from the source and in which the measurement may be subject to high noise levels.

Tech stack: Python, Numpy, Digital Signal Processing (DSP)

- **Impulse Response measurement software** 2018
  MATLAB development to apply convolutional operations using sine sweeps to obtain the Impulse Response of an enclosure and extract acoustic parameters. Obtaining EDT, T20, T30, C50.

Tech stack: MATLAB, REW, Reaper/Audacity

Languages

Fluent in **English** (B2+) - **Spanish** native - **French** native

Education

National University of Tres de Febrero (Buenos Aires) 2012 to 2023
Audio - Sound Engineering

Certifications

Anyone AI 2022
Machine Learning Engineer

UNTreF 2022
Neural networks applied to Musical Information Retrieval (MIR)