Ivan Ledesma

Buenos Aires, Argentina
Phone number: +54 2964 510583 E-mail: ivanarielledesma@gmail.com
Linkedin: https://www.linkedin.com/in/ivanledesma-/
GitHub: https://github.com/ledesma-ivan
Web: https://ledesma-ivan.github.io

Data Scientist |Machine Learning Engineer

Machine Learning Engineer with a diverse portfolio of successful projects with technologies such as **Python, Pandas, Numpy, Scikit-learn, Keras, and TensorFlow**. My experience spans a range of domains, including automated product categorization for e-commerce, salary prediction models, product review classification, and Home Credit Risk Analysis, all executed with excellence during my tenure at Anyone AI. My approach to professional growth is marked by self-driven learning and a strong commitment to achieving tangible results. I thrive as a collaborative team player, harnessing my passion for the field to inspire and innovate alongside my colleagues. Currently, I am actively seeking my next opportunity within the machine learning industry. My aim is to tackle intricate business challenges, make a meaningful impact, and continue my journey of personal and career development. I am eager to contribute my expertise and enthusiasm to a dynamic organization in pursuit of groundbreaking solutions.

Work Experience

Tutor Data Science & Data Engineering at Coderhouse Jan 2023 to date

- Accompanying students in their learning process, motivating them to meet the challenges, and offering support in the difficulties they may encounter, evaluating the performance of practical work.

Machine Learning Engineer as Freelancer Jan 2023 to Apr 2023

- Running ML applications in Docker environments: Design, implementation, and deployment of machine learning applications in Docker containers, additionally: Refactoring, cleanup, unit testing, and documentation of an existing Python machine learning pipeline.

Main technologies: Python, Docker, Podman.

Machine Learning Engineer at Anyone AI Apr 2022 to Nov 2022

- **Automated product categorization for e-commerce:** Creating relevant tags and categories on products allows e-commerce companies super apps and marketplaces to automatically categorize products, whether they are new products uploaded by a user or products from the seller that need automatic categorization on a large scale. Automation in this topic not only saves manual time and effort, but creates a taxonomy system that improves the process of a customer finding what they are looking for, improving their conversion metrics, and triggering frictionless and engaging shopping

experiences. In this project, natural language processing was applied to create a multi-label system capable of classifying e-commerce products automatically.

- **Product review classification:** Analyze sentiment in product reviews for a movie streaming service. Manipulated data that was 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:** Predict 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 achieving +0.72 ROC AUC. The 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 then 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%.

Main technologies: Python, Numpy, Pandas, Scikit-learn, Keras, PyTorch, Docker, AWS, TensorFlow.

Projects

Docker PostgreSQL Data Pipeline 2023

- This project consists of the implementation of a data pipeline using Docker and PostgreSQL. The goal is to create an efficient and scalable solution for data processing and storage using containerization and database technologies. The pipeline handles data manipulation and processing in a container environment using Docker, and data persistence in a PostgreSQL database.

Main Technologies: Python, Docker, PostgreSQL.

Student Dropout Project 2021

- Predicts whether a student studying at the university will be able to drop out of the university or not Data was manipulated and visualized, and data preprocessing was performed and trained in a supervised model achieving **+0.80 ROC AUC**. Model used **XGBoost**

Main Technologies: Python, Numpy, Scikit-learn, Xgboost.

Skills

Tech Skills: Python, R, SQL, PostgreSQL, Numpy, Pandas, Keras, Scikit-learn, Redis, Pytorch, Docker, AWS, Fast API, Flask, Terraform, ETL.
Agile Methodologies: Scrum and Kanban.
Other Tools: Git, GitHub.
Languages: Intermediate English, Spanish Native

Education

Guillermo Brown National University 2020 to date
Bachelor's Degree in Data Science

Certifications

Specialized program: Machine learning - **Coursera**	Currently
Computer Vision certificate - **Kaggle**	2022
Intro to Deep Learning certificate - **Kaggle**	2022
Feature Engineering certificate - **Kaggle**	2021
Data Visualization certificate - **Kaggle**	2021
Intermediate Machine Learning certificate - **Kaggle**	2021
Pandas certificate - **Kaggle**	2021
Python certificate - **Kaggle**	2021