

Phil Salazar

+1 (707) 853-3757 • phil.salazar64@gmail.com • linkedin.com/in/phil-salazar • phil-salazar.github.io

Experienced financial services engineer with a robust full-stack background, including cloud services, Python, SQL, API development, and React. Recently expanded expertise into LLMs, ML, and NLP, with foundational knowledge in corporate finance and financial services at Harvard. Proven ability to bridge technology and finance to drive innovative solutions.

WORK EXPERIENCE

Grupago

09/2024 - Present

Fintech startup offering credit products to entrepreneurs in Latin America

Data Scientist • Contractor

- **Developed and deployed a comprehensive machine learning pipeline in Python**, including a robust credit risk model utilizing diverse alternative data sources, advanced machine learning algorithms, and Natural Language Processing (NLP) techniques for enhanced accuracy.
- **Engineered an Agentic AI system using Large Language Models (LLMs)**, automating key credit analysis processes and providing intelligent decision capabilities.
- **Authored extensive documentation** for machine learning pipelines and models, ensuring clarity, maintainability, and knowledge transfer across the team.

Bankaya

05/2023 - 08/2023

Fintech startup offering BNPL financial products to bottom-of-the-pyramid customer base

Credit Risk Analyst • Internship

- **Implemented a new credit risk strategy** specifically for returning customers, leveraging machine learning techniques within AWS, Python, and SQL environments, resulting in a significant improvement in overall credit margin by more than \$1M annually.
- **Managed key relationships and negotiations** with alternative data vendors in Spanish, securing favorable pricing terms relative to usage volume in production models.

Indigov

01/2022 - 08/2022

Data Engineer

- **Designed and implemented efficient Python and bash scripts, along with robust ETL pipelines**, integrating diverse data from external vendor APIs, ensuring high data quality, seamless data flow, and ultimately accelerating client onboarding processes by 25%.
- **Managed and optimized the execution of complex Python data pipelines**, simultaneously provisioning cloud resources utilizing Prefect and Terraform, significantly reducing overhead resources by 18%.

GoodLeap

11/2019 - 11/2021

Data Engineer

- **Led an Agile business intelligence team**, consistently delivering over 95% of portfolio reporting and interactive functionality on time to diverse business stakeholders and C-suite executives.
- **Engineered a full-stack business intelligence solution**, developing SQL queries, maintaining data models, building Python ETL & ELT pipelines, and designing visualization dashboards in React for over 100 bespoke financial dashboards.

- **Ensured software quality and maintainability** by writing meticulous tests for all developed components and maintaining comprehensive documentation for pipelines, data models, and dashboards.

Amazon

06/2017 - 09/2018

Retail Vendor Manager

- Grew \$20M vendor portfolio by 23% year over year using SQL to analyze sales trends.
- Served as vendor lead time subject-matter expert, collaborating with supply chain analysts to reduce inventory lead time by 12%, resulting in over \$7M additional revenue.

EDUCATION

Master in Public Administration in International Economics

Harvard University

08/2022 - 05/2024

Advanced Statistics, Machine Learning Operations, Deep Learning, **Corporate Finance (MIT Sloan), Financial Services (HBS)**

Course Assistant: Advanced Microeconomic Analysis (Prof. Luis Armona, Prof. Jie Bai)

Research Assistant: Randomized Control Trial (RCT): Economic impact of trade fairs in Chile (Prof. Jie Bai)

Bachelor of Science in Engineering

Stanford University

08/2012 - 06/2016

PROJECTS

Prescribe.ai

Developed app to support accurate medication prescriptions for under-resourced medical professionals in Latin America. Managed Docker container virtual environments, deployed in GCP, trained LLM using Hugging Face and Weights & Biases, and built frontend in React.

Predicting Disease Burden in Brazil

Developed machine learning model in R to predict municipalities with highest disease burden to prioritize investment in health infrastructure.

SKILLS

Cloud Platforms & Infrastructure

- AWS
- Databricks
- Docker
- GCP
- Prefect
- Terraform

Machine Learning

- Deep Learning
- Keras
- NLP
- Tensorflow

Programming

- Github
- NumPy
- Pandas
- PySpark
- Python
- R
- React
- sklearn
- SQL
- Unix