Etienne Doidic

310-863-8322 | etiennedoidic@gmail.com | San Diego, CA

Profile

Data Scientist & Machine Learning Engineer with 8+ years of experience solving complex data challenges and **4+ years in the identity verification (IDV) industry.** Proven track record of building automated fraud detection systems, document detection and classification pipelines, and delivering privacy-first ML models adopted by global clients. Recognized as a subject matter expert in **fraud detection, document liveness, and IDV risk modeling**.

Experience

Machine Learning Engineer I Trulioo *August 2021 - present*

Patented ML Fraud Detection: Designed and deployed an encrypted-input risk model that generated 5,000+ features to detect onboarding fraud without exposing sensitive data. The solution replaced 3rd-party IDV systems and saved the company **millions annually** across millions of transactions.
Custom ML for Clients: Partnered with enterprise clients (PayPal, Uber, Ekata) to rapidly deliver **privacy-compliant fraud models** tailored to unique datasets (spanning up to 1 year of onboarding activity). Reduced fraud rates while cutting model turnaround time to under one week per client.
Document Fraud & Liveness Detection: Developed a synthetic image training dataset to address scarce spoofed identity document examples (photo-of-photo, composite, and generative fakes). Engineered an **automated image capture rig** and multi-stage augmentation pipeline to simulate real-world spoofing scenarios. This enabled quick model iteration and improved detection of common spoofing vectors that represented ~80% of observed fraud, increasing detection performance by ~25-40%.
ML Infrastructure (Linux): Built and maintained **in-house supercomputer** for large-scale ML training and secure data storage.
Labeling tools: Built several **real-time fraud labeling dashboards** capable of processing 100k+ transactions per day. Enabled analysts to efficiently flag suspicious activity, identifying up to 200 fraudulent onboarding attempts daily and track model uplift.

Education

B.S. Data Science I UC San Diego *September 2017 - June 2021*

Relevant Coursework: Statistics, Data Structures & Algorithms, Distributed Comupting, Scalable Analytics (Amazon EC2), Spatial analysis (GIS), Web-scraping, Predictive-modeling, Natural language processing (NLP), Database Management, Data visualization, Web Design

Skills

Programming Languages : Python, JavaScript, SQL, C++, CSS, HTML, MATLAB, R
Libraries & Frameworks: tensorflow, torch, React.js, Svelte, jquery, Esri/ArcGIS, Flask, Raspberry Pi,
DevOps: CI/CD, Docker, Git
Databases & Cloud: SQL, MongoDB, Amazon EC2, Redshift, Lambda
Languages : English - Native, French - Native, Spanish - Proficient