John Paul Feliciano

johnpaulfeliciano98@gmail.com | (951) 315-6763 | github.com/johnpaulfeliciano98 | linkedin.com/in/johnp-feliciano

Computer Science Graduate with Real-World ML, NLP, and Analytics Experience

Computer Science graduate with practical experience in machine learning, natural language processing, and data analytics. Developed predictive models and automated tools using Python, SQL, and visualization software to process structured and unstructured data, support dynamic data reporting, and extract insights from large datasets.

Education

Oregon State University – Bachelor of Science in Computer Science March 2025
Relevant Coursework: Machine Learning (Capstone), Introduction to Databases, Data Structures, Analysis of Algorithms, Discrete Structures in CS

California State University, Fullerton – Bachelor of Science in Health Science May 2020
Relevant Coursework: Introduction to Probability & Statistics, Measurement & Statistics in Health Science, Public Health Research & Evaluation

Projects

Real-Time EMS Call Prediction – Capstone github.com/johnpaulfeliciano98/ems-call-forecasting
- Developed and implemented machine learning models (XGBoost, K-Means) to forecast EMS call volumes
- Engineered geospatial and temporal features; evaluated LSTM and GNN models for time series forecasting
- Built REST APIs (Django) and React/Mapbox interface to visualize predictions in dynamic reports
- Delivered insights through data interpretation, visualizations, and model comparisons

Contract Clause Analyzer – Internship github.com/johnpaulfeliciano98/contract-clause-analyzer
- Created a natural language processing tool to analyze PDF contracts using OpenAI Assistants API
- Used regex, PyPDF2, and API orchestration (Agency Swarm) to detect overpayment clauses
- Delivered a GUI (Tkinter) for report generation and clause flagging

Invoice Anomaly Detector – Internship github.com/johnpaulfeliciano98/Risk-Assignment
- Developed unsupervised learning pipeline (DBSCAN) to detect anomalous invoice descriptions
- Engineered NLP features from unstructured transaction data to improve model clarity and fraud detection
- Generated clustering visualizations and model summaries for audit analysts

NPI License Verifier – Internship
- Automated license verification using Python and NPPES API, analyzing over 100K healthcare provider records
- Processed large structured datasets to ensure compliance, accuracy, and eliminate manual data entry

Skills

Languages & Tools: Python, SQL, Git, Jupyter, VS Code, REST APIs
Libraries & Frameworks: Pandas, NumPy, Scikit-learn, XGBoost, Seaborn, Mapbox
Core Competencies: Data wrangling, Data mining, Statistical analysis, Data modeling, EDA, NLP, Machine learning, Model evaluation, Reporting, Visualization of structured and unstructured data

Experience

Data Science Intern, Inland Empire Health Plan (IEHP) – Rancho Cucamonga, CA June 2024 – December 2024
- Designed and implemented a prototype to extract contract risk language using NLP and OpenAI APIs
- Built DBSCAN-based anomaly detection for procurement transactions using unstructured data inputs
- Automated physician license validation process using structured datasets and NPPES API
- Created ERDs and data flow documentation for machine learning initiatives in fraud detection
- Generated dynamic visualizations and summary reports to inform strategic purchasing decisions

Certifications

IBM Data Science Professional Certificate – In Progress (Coursera)