Christopher Daigle

pcjdaigle@gmail.com | +1 (520) 203-6290
linkedin.com/in/christopherdaigle

quantchris.com | github.com/christopherdaigle/
pypi.org/user/christopherdaigle/

Data scientist specializing in machine learning and software engineering • Army Veteran • Open source contributor • Leadership ranging from small team project management, running a startup, and lead pilot in Afghanistan • Comfortable with ambiguity and driving results with little oversight

Key Skills

Technology: Python • R • Oracle SQL • Git & GitHub • Bash / Unix / Shell • Stata • MATLAB / Octave • Pandas • NumPy • SciKit-Learn • Plotly & Seaborn • PyTorch • Keras & Tensorflow • Markdown & LaTeX

Quantitative: Machine Learning • Statistical Analysis & Predictive Modeling • Supervised Learning • Unsupervised Learning • Natural Language Processing (NLP) • Dimensionality Reduction • Hypothesis Testing

Professional Experience

The Hartford Financial Services Group, Inc. – *Senior Data Scientist* Jul 20 – Pres.

- Invented generalized anomaly detection algorithm to identify problem claims for workers compensation
- Created procedures for feature selection pipeline for the Claims Data Science organization
- Helped standardize data science practices with peer review, version control, and modularized testable code
- Introduced novel methods, like Variational Autoencoders (VAE, deep learning), for high cardinality problems

Pratt & Whitney – *Manager, Data Scientist* Jan 19 – Jul 20

- Helped reduce cost by $29 million by identifying the optimal allocation of 75,000+ parts sold by 5,000+ vendors by creating an algorithm and engineering software
- Reduced analysis time from 1.5 months to < 1 minute by creating an algorithm, engineering an API, and a deploying a Flask app resulting in continued global supply chain operations
- Performed data science, project management, data engineer, software engineer, and devops duties

Boise Analytics – *Partner, Data Scientist* Dec 17 – Jan 19

- Assisted 43 non-profits and small businesses solve data problems through data science solutions
- Increased company talent by interviewing and mentoring 38 data analysts

University of Connecticut – *Economics Instructor* Aug 16 – Aug 18

- Instructed microeconomics and economic research methods to undergraduates

Boise State University – *Economic Researcher* Jan 14 – May 16

- Produced economic research in partnership with Yale University to measure GDP from satellite imagery
- Measured returns to investment in education for students in Idaho, partnered with Boise State University's Economics Department and Idaho Voices for Children

Veterans Affairs – *Work Study* Apr 13 – May 14

AAI Corporation – *Lead Pilot, F-227* Oct 10 – Apr 13

US Army – *Sergeant, Drone Pilot* Sep 04 – Oct 10

Projects / Products

Clinical Referral *Data Science* (proprietary model and methodology)

Purpose: assign to nurses claims for workers compensation that are most likely to be complex or costly to manage

Outcome: currently in development

Machine Learning: VAE, KMeans, DBSCAN, FAMD, MCA, Decision Tree, Anomaly Detection, Hypothesis Testing

Technology: Python, Linux, Oracle SQL, PyTorch, Keras, Tensorlfow, Sklearn, Matplotlib, Seaborn, VSCode, Jupyter, Git, GitHub

Rebate Optimization *Software Engineering* (proprietary software)

Purpose: increase rebates from suppliers, reduce spending, and reduce overall cost

Outcome: application to determine the optimal allocation of spending at the part level for 5,000+ vendors over 75,000+ jet engine components

Technology: Python, NumPy, Pandas, Oracle SQL, PyInstaller

Award: Special Award for Innovation at Pratt & Whitney awarded to ~5 people out of more than 300k people a year

Alternative Vendor Identification *Software Engineering* (proprietary software)

Purpose: mitigate impact of COVID-19 on global flight operations

Outcome: application that identifies vendors having shared capability or sole source for repairs – performs for entire supply base in <1 minute what used to take 5 senior sourcing professional 1.5 months to analyze for a single vendor

Technology: Python, NumPy, Pandas, Oracle SQL, Flask

Commodity Classification Innovation *Natural Language Processing, Classification* (proprietary software)

Purpose: identify jet engine commodities from purchase orders executed by global supply buyers

Outcome: model that classifies 90%, up from 60%, of $16 billion worth of purchase orders

Technology: Python, SQL, Pandas, NumPy, NLTK, SciKit-Learn (sklearn), Tensorflow and Keras

Machine Learning: Multinomial Naïve Bayes, AdaBoost, Bagging, Random Forest, TF-IDF

Performance: 94% F-1 Score, 96% Recall, 93% Precision

Find Donors for Charity *Supervised Learning* quantchris.com/project/Donor-Classification/

Purpose: maximize the likelihood of receiving donations by predicting if a person receives income exceeding 50k/year

Technology: Python, Scikit-Learn (sklearn), Pandas, NumPy, Seaborn, Plotly, PyCharm, Jupyter Notebook

Machine Learning: Ensemble Methods (ADABoost, Random Forest, Gradient Boosting), Logistic Regression, KNN, Naïve Bayes, Grid Search, Feature Scaling (Standardization, Normalization, Logarithmic Transform), One-Hot-Encoding (OHE)

Performance: 87.26% Accuracy, 76.05% F-0.5 Score

Predicting Movements in Social Security Filings *Supervised Learning,* quantchris.com/project/sup-ss-move

Purpose: determine if movements in social security filings can be predicted from economic and financial indicators

Technology: R, R-Studio, Python, beautifulsoup, Pandas

Machine Learning: Logistic Regression, Limited Dependent Variable (LDV), Greedy Selection Methods (Backward, Forward, Sequential Replacement), Hypothesis Testing (Augmented Dickey-Fuller, Likelihood-Ratio Test)

Performance: 93% Accuracy; 92% F1-Score

Identify Customer Segments *Unsupervised Learning, Clustering* quantchris.com/project/unsup-cust/

Purpose: determine Bertelsmann Arvato Analytics' customer segments to optimize outreach through mailings

Technology: Python, NumPy, Pandas, Seaborn, SciKit-Learn

Machine Learning: K-Means, Dimensionality Reduction (Principle Component Analysis - PCA), Feature Scaling (Standardization, Normalization), Imputation

Education

MS, Quantitative Economics (STEM), University of Connecticut, CT (Maj. GPA 3.95)

Certifications

Nanodegree, Data Scientist, Udacity (In Progress)

Certificate, SAFe Scrum Master (SSM - 92474883-9992), Scaled Agile

Nanodegree, Machine Learning – Introduction, Udacity

Certificate, Natural Language Processing with Python, Udemy