

# P. Christopher J. Daigle

[pcjdaigle@gmail.com](mailto:pcjdaigle@gmail.com) | +1 (520) 203-6290

[quantchris.com](http://quantchris.com) | [github.com/christopherdaigle/](https://github.com/christopherdaigle/)

[linkedin.com/in/christopherdaigle](https://www.linkedin.com/in/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

**Pratt & Whitney** – *Manager, Data Scientist*

Jan 19 – Pres.

- Reduced cost by \$29 million by identifying the optimal allocation of 75,000+ distinct part numbers sold by 5,000+ vendors by creating an algorithm and engineering software
- Increased analysis speed critical to mitigating COVID-19 from 1.5 months to less than 1 minute by creating an algorithm, engineering an API, and deploying a web app resulting in continued global aviation operations
- Classified 90% of \$16 billion worth of misclassified jet engine components with NLP
- Classified 60% of \$10 billion of misclassified jet engine components by mining data and developing ETL
- Identified opportunities for and drove data science projects through the development lifecycle
- Improved development operations by implementing version control, server management, and general development operations standards resulting in reduced re-work and increased output of developers

**Boise Analytics** – *Data Scientist*

Dec 17 – Jan 19

- Assisted 43 non-profits and small businesses in solving data problems by applying data science solutions
- Increased company talent by interviewing 38 data analysts
- Increased capability by mentoring teams of 20 data analysts and scientists on data science methods

**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 research for education and development in partnership with non-profit organizations and Yale University

**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

**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 a maximum of 5 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/](https://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](https://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/](https://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

## **Predict Clothing Items** *Deep Learning, Classification* [tinyurl.com/DLCloth](https://tinyurl.com/DLCloth)

*Purpose:* create an application that can be trained on any set of labeled images to predict the contents of an image

*Technology:* Python, PyTorch, argparse, PIL, Scikit-Learn, Pandas, NumPy, Seaborn

*Machine Learning:* Artificial Neural Networks (ANN), Transfer Learning (VGG11), Dropout, Rectified Linear Unit (ReLU)

## **Education**

**MS**, Quantitative Economics (STEM), University of Connecticut, CT (Maj. GPA 3.95) Dec 18

*PhD Course Work:* Microeconomic Theory I & II, Macroeconomic Theory I & II, Econometrics I & II, Industrial Organization I, Advanced Mathematical Economics

*MS Course Work:* Applied Econometrics II, Python Programming, R Programming, Big Data (Machine Learning), Convex Optimization with Python, Panel (Longitudinal) Data

**BA**, Honors degree, Economics with a Quantitative Emphasis; Minors: Mathematics & Applied Mathematics, Boise State University, ID (Cum. GPA 3.691) May 16

Calculus I, II, & III, Differential Equations, Linear Algebra, Numerical Analysis, Probability & Statistics, Computational Mathematics

*Honors:* Econometrics, Money & Banking, Quantitative Methods, Senior Thesis

## **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