Daniel K. Griffin

85 Rio Robles East, #2405, San Jose, CA | (513) 646-2454 | dcompgriff@gmail.com | https://dcompgriff.github.io/

Summary of Education and Experience

Senior R&D Data Scientist and Machine Learning Engineer — **7+ Years**
Companies, Universities, Government Research Institutions — Over 50 Data Science Projects

Masters in Computer Science – Machine Learning Focus — **GPA: 3.85/4.0**
UNIVERSITY OF WISCONSIN-Madison, WI

Bachelors of Science in Computer Engineering — **GPA: 4.0/4.0**
UNIVERSITY OF CINCINNATI-Cincinnati, OH

Relevant Skills

Data Science and R&D

- Advanced Statistical Modeling
- Semi-Supervised Learning
- Advanced NLP/NLG
- Causal Analysis
- Bayesian Statistics
- Advanced Unsupervised Learning
- Reinforcement Learning
- Deep Neural Networks
- Advanced Supervised Learning

Data Science Frameworks

- Scikit-Learn, Pandas, Numpy, Scipy
- Auto Sklearn, Imbalanced-Learn
- Tensorflow, Keras, PyTorch
- Spark, Spark MLLib, Spark SQL
- DoWhy, CI-Toolkit, CausalML
- Stable-Baselines RL, AI Gym
- PyMC3, Statsmodels, Tensorflow Probability, Uber Pyro
- Jupyter, Kubeflow, MLFlow

Software Engineering

- Python (8 years)
- Java & Java Web Apps (9 years)
- AWS Cloud Services
- GCP Cloud and AI Services
- C++, Embedded C
- SQL & NoSQL Databases

Management & Communication

- Cisco CX Technical Staff Data Scientist
- WPAFB AI Project Management
- Technical Publications Author
- Teacher of Undergraduate AI at UW, Graduate ML TA
- Technical Manager of 8 Data Scientists at Cisco
- AI/ML Blog Creation
- Data Science Conference Lecturer
- Technical Manager of 6 Major Data Science Projects at Cisco

Professional Experience

Cisco Staff Data Scientist (May 2018 – Present; ~2+ years)

I led technical projects for a group of eight data scientists developing a wide range of machine learning models and systems to directly optimize customer experiences at Cisco. I leveraged terabyte quantities of data and the latest AI/ML techniques to build powerful data science products for a variety of Cisco business problems. This includes performing R&D to develop practical methods to discover and identify cause-effect relationships in complex business problems, and proposing actions mathematically derived to cause business growth; optimizing content, timing, and sequences of decisions for email engagements with customers; classifying requests and automating responses to customers through social media; forecasting product sales and opportunities; and optimizing customer journeys to automatically improve Cisco's customer experience. My methods spanned advanced statistical analysis (causal analysis, statistical decision systems, reinforcement learning, bayesian statistics, etc.), deep learning systems (deep NLG/NLP, deep reinforcement learning, deep statistical models etc.), semi-supervised methods (transfer learning, weak supervision, few shot learning, etc.), and many others. I developed and deployed my models using Spark/Hadoop, as well as the data science infrastructure on Google Cloud including Dataproc, ML Engine, the AI Platform, and GPUs/TPUs. I also planned and vetted technical designs for all data science projects in our group, wrote technical papers, and presented at data science conferences. Within my first two years at Cisco, I was promoted from data scientist to senior data scientist, and most recently to staff data scientist.

Graduate Researcher, Teacher, & TA: University of Wisconsin (Aug. 2016 – May 2018; ~2 years)

I performed research in the David Page Machine Learning and Biostatistics Group at the University of Wisconsin, developing new machine learning methods for discrete event, continuous, time series data sets (a valuable but under-explored area of statistical modeling). I investigated and developed various multi-task and transfer machine learning methods based on traditional machine learning, as well as methods that use temporal-point process models. In addition, I developed new sensitivity analysis benchmarks for deep object recognition models, and a time series database system targeted specifically at advanced machine learning modeling applications (beyond simple aggregation streaming statistics). I was also a teacher for the Artificial Intelligence Course with over 400 students, as well as a Teaching Assistant for the graduate level machine learning course. I taught data science, data mining, advanced classification and clustering, intelligent search, and optimization.

Data Scientist: University of Cincinnati & Wright Patterson AFB (May 2012-Aug 2016; ~4 years)

I performed AI and ML R&D, developed software applications, and managed data science projects in the Human Performance Wing at Wright Patterson. I performed data science in the areas of automated health analysis, performance readiness prediction, injury risk, and chemical assay quantification. I used a wide variety of both standard and state-of-the-art machine learning models and data-driven analysis. I also invented a flexible, disposable, passive sweat-sensing embedded Band-aid system, and full-stack supporting software system for pro-athlete health monitoring.

Machine Learning Projects (https://dcompgriff.github.io/portfolio/)

Prescriptive Actions Engine Product 'BRIAIN' (Cisco)

- **Description:** Performed R&D, Development, and Deployment of an AI service that learns to prescribe optimal decision-making and action policies to optimize any metric. Applied to maximizing customer engagements with email campaigns. I was a finalist for the 'Best CX Product' award at Cisco.
- **Methods:** Advanced statistical decision theory, game theory, agent-based AI, causal analysis, and advanced reinforcement learning (Bandits, Model free ML based, deep learning based, policy-gradient based, along with extensions to the classical reinforcement learning framework).

Causal Inference Product 'Lever' (Cisco)

- **Description:** Performed R&D, Development, and deployment of a causal inference pipeline to answer 'what-if' and 'root-cause' questions about Cisco's CX strategy. I won an award for best 'Top Line Business Growth' at Cisco for estimated reclamation of over 12% of lost contract dollars.
- **Methods:** Regression discontinuity design, Difference-in-Differences, Machine Learning Model Based Uplift Modeling, Potential Outcome Modeling, Causal Graphical Models and Do-Calculus. DoWhy, causalml, EconML, the Causal Discovery Toolbox, and custom code were heavily used.

Social AI Project (Cisco)

- **Description:** Developed a layered suite of six automated AI classification, regression, and response models for Cisco's social media listening team that responds to over one million posts a week and 8,200% efficiency improvement. I was a finalist for the 'Best CX Product' award at Cisco.
- **Methods:** Advanced deep recurrent and attention neural networks, pre-trained and custom deep embeddings, classical models like SVMs, logistic regression, models with n/skip-grams, PCA+tf-idf.

ConTEXT AI Project (Cisco)

- **Description:** Developed a suite of general text processing services to analyze Cisco's support cases, email engagements, community forums, product documents, and Cisco.com web pages to provide hundreds of ML driven insights and automated actions, and grew it into a program with multiple data scientists. I won 2nd place at a competition and a $15k prize for my work.
- **Methods**: Custom-invented semantic clustering method that far exceeded the capabilities of traditional methods, deep neural networks, deep unsupervised methods, advanced semi-supervised topic labeling

Sales Analysis and Forecasting (Cisco)

- **Description:** Developed an advanced quarterly sales forecasting model that is actively used to forecast sales and make multi-year strategic business decisions. This project was used to provide reasonable sales projection estimates, eliminating previously overly-zealous expectations of sales and sales growth figures.
- **Methods:** Custom analysis method using robust Huber linear regression, Gaussian Process Regression, ARIMA models.

Sales Opportunity Scoring Model (Cisco)

- **Description:** Developed a system for predicting the likelihood of a sales opportunity being booked, given that it has been opened by a seller.
- **Methods:** Spark ETL and ML pipeline on top of Hadoop+Hive along with Jupyter notebooks

Colorimetric Assay Quantification (Wright Patterson AFB)

- **Description:** Developed a mobile phone app and custom modeling method to quantify colorimetric assay values for remote medical scenarios.
- **Methods:** Color theory, linear algebra, non-gridded multivariate interpolation algorithms.

Time Series Transfer Learning (UW Research)

- **Description:** Developed a novel method for performing general transfer learning over any time-series data set without traditional assumptions by using a generative probability model.
- **Methods**: Bayes nets, generative frequentists statistical models, transfer learning theory.

'Normal Noise Sensitivity' Deep Model Benchmark (UW Research)

- **Description:** Developed a novel 'natural sensitivity analysis' benchmark for object recognition models for performing sensitivity analysis to natural variations in camera images; applied the analysis method to over 10 state-of-the-art object recognition models; and developed eight best practices to developing robust neural networks.
- **Methods:** Mask R-CNN, RetinaNet, Faster R-CNN , RPN , Fast R-CNN , R-FCN, YOLO, distributed over nine GPUs with 750 Million object candidates per sensitivity test.