

Sean Conroy: Data Science Professional

address: **Dallas, TX 75228**

[Github Profile](#)

phone: **412-855-7246**

[StackOverflow Profile](#)

email: seantconroy@gmail.com

[Website / Portfolio](#)

overview

Strong analytics, data science, engineering & manufacturing background and 10 years of experience writing code in production environments. Known for capacity to quickly gain mastery of a new field.

specific interests	languages	Platforms	ml workflows	<u>processing</u>
really fast R code, time-series classification, unsupervised learning, parallelization, shiny apps	Python, R, SQL, VBA, HTML, Matlab	Google Cloud Platform (BigQuery, GCS, ML), Prefect, Apache Beam / Dataflow, DVC, Shiny, Docker, CircleCI, Jira	xgboost, dbscan, TensorFlow / Keras, scikit-learn	data.table, doparallel, pandas / dask,

experience

ActivTrak

June 2021 – July 2023

Senior Software Engineer (Data Science)

- **Location:** Primary developer for ActivTrak’s Location Insights program, which “continues momentum influencing over \$1M of new deals and upgrades”
 - Custom ML deployment with multiple models that uses wifi and ip data to predict employee location (remote vs. office) for ~ 2000 clients every hour with > 95% accuracy.
 - Feature engineering for up to 70M rows per client in BigQuery SQL
 - XGBoost / dbscan / Louvain community detection with Dask + Prefect Python workflow wrapper
- **Identity Model:** designed end-to-end workflow in R for clustering multiple user records into a “person-record” using custom named entity recognition, custom string distance metric, Louvain community detection, and other custom clustering algorithms.
- **Google Calendar Integration:** Python POC using Google / Outlook calendar API’s to detect Out of Office events
- **Experimental Agent:** developed and deployed monitoring software using Python that tracks and records a variety of user activity useful to predict productivity for iOS and Windows

Comstock Resources (Oil & Gas E&P)

January 2018 – June 2021

Data Analyst III, Business Development Team

Principal developer for a team that created analytical tools for the Business Development, Operations and Finance departments using R, Python, SQL and Spotfire:

- Deployed machine learning models to predict wells reaching line pressure / undergoing “roll-over”
 - Ensemble models using Tensor Flow Neural Nets + XGBoost trained on custom-built train set
 - Generate nightly predictions being used by reservoir engineers
- Built machine learning model in R / Python to predict Haynesville gas well performance
 - Included over 100 engineered features, (geology, well design, spacing and completion parameters)
 - Automated data preprocessing, encoding, and model hyperparameter tuning
 - Created ML models for Clustering, Classification & Regression using XGBoost + DBSCAN
- Automated Time Series Forecasting of Gas Wells (**Decline Curve Analysis**):
 - **Deliverable:** Spotfire template for automated forecasting of oil & gas wells
 - Custom algorithms for noise removal, data segmentation, and forecasting
 - Non-linear regression, parallel processing, ~ 5000 lines of code, 7 libraries built
- Pressure Normalized Rate (PNR) Forecasting with Aries Comparison:
 - **Deliverable:** Spotfire template that is now widely used by organization’s engineers and analysts

- Iterative calculation of Bottom Hole Pressure using complex petroleum engineering models
- Functionality to load, translate, plot, and adjust “Aries”-style forecasts
- (Pason Drilling Analytics) Automated Analysis of 10-second Drilling Data:
 - Multi-step workflow for cleaning data, automated joins, detecting rig-states, and analyzing data for automated KPI analysis

Diab Americas (Aerospace Foam Manufacturing & Analytics)

June 2016 – January 2018

Process Engineer

- Built and deployed several interactive databases + dashboards using (MS Access / SQL / VBA) for Production Management, Laboratory Management, Quality and Inventory Control
- Built numerous analytical tools w/ “dashboards” for automated data processing / report generation
- Process Engineer for 4 product lines (provided in-depth engineering analysis of extrusion data)

Poly-America (Blown Film Manufacturing Analytics)

February 2013 – May 2016

Assistant Laboratory Manager

- Regularly perform extensive statistical analysis for internal corporate sales & marketing clients
- Develop numerous VBA automation tools / “dashboards” / lab software
- Completed Projects:
 - (R) Process Variation analysis tool: used to analyze years of 24/7 production data
 - (R) Production/QC/Lab Data System Aggregation: product-based datasets for analysis
 - (VBA) Materials Traceability Solution: provided complete traceability for two mfg. locations consuming over 1 billion lbs. of polyethylene per year
 - (VBA): Automated Film Statistics Toolbox; ANOVA-style analysis for arbitrary datasets
 - (VBA): Lab Information Management System; enable sample tracing and data analysis

Natural Composites (Material Science Startup)

August 2011 – February 2013

Research and Development Engineer

Antioch Ministries International, Kurdistan, Iraq

September 2007 – August 2008

Missionary / English Professor, Koya University

Sandia National Laboratories, Livermore, CA

May 2005-August 2007

Research Intern, DOE Q (Top Secret) Clearance

(Education)

– **Masters of Engineering**

- GPA: 3.66

Baylor University

Graduated 2011

– **Bachelor of Science, Physics**

- GPA 3.22

Carnegie Mellon University

Graduated with honors, 2007

machine learning / data science:

▪ **Coursera:**

- [Neural Networks and Deep Learning with Andrew Ng, deeplearning.ai](#)
- [Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization](#)
- [Structuring Machine Learning Projects](#)
- [Data Science Specialization with R, Johns Hopkins](#)
- [Bayesian Statistics: From Concept to Data Analysis](#)
- [Getting Started with AWS Machine Learning](#)
- [Deep Learning Bootcamp](#) (week-long intensive on neural networks using TensorFlow / Keras)
- [Statistical Process Control](#) (week-long training at Univ. Tennessee Statistics Dept.)
- ONLC Training Classes: [Python for Data Science](#) (Dec, 2018), [Introduction to SQL Server](#), (Feb, 2020)