

Ryan Johnson

Data Scientist/Machine Learning Engineer *Seattle, WA*

+1 (253) 315-0389 | riosanjuan314@gmail.com | rtjohnson12

Skills

Programming Languages	Python (<i>pandas, sklearn, pytorch, pyspark, optuna</i>), R, Scala, Go, SQL, Bash
Machine Learning	Supervised Learning, Clustering, A/B Testing, Ensemble Learning, Dimensionality Reduction, Time-Series Analysis, Feature Engineering & Selection, Deep Learning, Generative AI, Graph-Based Learning
MLOps	AWS, MLflow, Docker, Airflow, Spark, Kubernetes, Terraform, Jenkins, Linux, Git

Work Experience

Data Scientist

Chewy

Jan 2022 - Nov 2024

Bellevue, WA

- Designed and deployed an end-to-end ML pipeline for forecasting Autoship orders. Combining subscription-level features with PySpark and AWS Glue, training and inference with AWS SageMaker, and orchestration through Apache Airflow, the pipeline includes automated image building, ETL, tuning, and inference, outperforming the existing heuristic-based approach by 10% MAPE.
- Designed and deployed a Hierarchical Reconciliation framework with AWS SageMaker, allowing for seamless integration with existing forecasting systems and providing a robust and scalable solution for reconciling forecasts across multiple granularities. This framework is now used by a majority of our production ML pipelines.
- Designed and deployed a forecast evaluation framework, allowing for automated comparison of model performance across our product catalogue, enabling us to rank model variants for specific use-cases like holiday planning, short-range labor planning, and long lead time buying decisions.
- Led weekly white-paper reviews covering the latest research in Data Science/Machine Learning.

Data Scientist

The Energy Authority

Sep 2019 - Jan 2022

Bellevue, WA

- Overhauled Hydro-Power Forecasting pipeline by integrating new data feeds from NW river system and plant-level generation. The system ingests and cleans data from thousands of river meters and achieved a 15% reduction in system-level forecast MAPE.
- Designed and lead transition to new CI/CD system for running containerized models with Docker and GitHub Actions, enabling streamlined deployments and establishing new team standards for project version control and production deployment.
- Worked with Software Engineering team to develop model failsafe system. In the case of critical data feed failures, production models revert to failsafe mode and utilize a reduced set of features, ensuring that our real-time traders will always have actionable forecasts 24/7.

Education

MS in Applied Mathematics

University of Washington

Dec 2019

Seattle, WA

- Professional Excellence Award for distinguished academic merit and professionalism

BS in Mathematics and Computer Science

Santa Clara University

Jun 2018

Santa Clara, CA

- Paul R. Halmos Award for Outstanding Academic Record
- President of Pi Mu Epsilon, Mathematics Honor Society