Ryan Johnson

Data Scientist *Seattle, WA*

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

Summary

I am an experienced Data Scientist with 5 years of expertise in desinging, developing, and deploying data-driven solutions. I have a proven track record of delivering actionable insights and scalable solutions, and a destinguished aptitude for working with cross-functional teams and stakeholders. I am passionate about researching cutting-edge ML frameworks and getting ideas off the whiteboard and into production environments.

Work Experience

Chewy

Jan 2022 - Present

Data Scientist II - Supply Chain Forecasting

Bellevue, WA

- Architected and deployed an end-to-end ML pipeline for forecasting Chewy Autoship orders. Combining PySpark for ETL processes via AWS Glue, model training and inference via AWS Sagemaker, and orchestration through Apache Airflow dags, the pipeline outperformed existing heuristic-based model by 3-5% MAPE.
- Architected and deployed a Hierarchical Reconciliation framework, 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.
- Enhanced the core SKU demand model by integrating promotion features, fine-tuning model accuracy and responsiveness to market trends and promotional impacts.
- Redesigned unconstrained demand model, which is central to our core SKU demand model. The updated framework includes an ensemble of conversion-rate heuristic models and tree-based models, reducing out-of-stock MAPE by up to 10%.
- Led effort into topic demand forecasting, employing a temporal knowledge graph approach to capture seasonal and thematic trends effectively.

The Energy Authority

Sep 2019 - Dec 2022

Data Scientist - Analytics

Bellevue, WA

- Overhauled Hydro Forecasting pipeline by integrating new API data feeds from NW river system and plant-level generation, achieving a 5% reduction in system-level MAPE.
- Lead transition to new CI/CD system for running containerized applications, enabling streamlined deployments and builds and establishing new team standards for project version control and production deployment.
- Worked with Software Engineering team to develop model failsafe system providing forecasts to real-time traders during critical data feed failures.

Education

University of Washington

Dec 2019

MS in Applied Mathematics

Seattle, WA

- Professional Excellence Award for distinguished academic merit and professionalism

Santa Clara University

Jun 2018

BSc in Mathematics and Computer Science, Emphasis in Data Science

Santa Clara, CA

- Paul R. Halmos Award for Outstanding Academic Record.

Skills

Programming Languages	Python (*pandas, pytorch, sklearn, optuna*), R (*tidyverse*), SQL, Scala, Bash
Machine Learning	Supervised Learning (*linear regression, logistic regression, SVMs, random forests, transformers*), Ensemble Learning (*bagging, boosting, stacking*), Dimensionality Reduction (*PCA, t-SNE*), Feature Engineering & Selection, Deep Learning
MLOps	AWS (*Sagemaker, EC2, Glue*), Docker, Apache Airflow, Jenkins, Git