Ryan Johnson

Machine Learning Engineer *Seattle, WA*

+1 (253) 315-0389 | rtjohnson1206@gmail.com | rtjohnson12 | rtjohnson12

Summary

I'm an experienced Data Scientist with over 4 years of experience in the field, eagerly transitioning into my new role as a Machine Learning Engineer. My expertise lies in unraveling complex datasets and leveraging cutting-edge machine learning technologies. I focus on transforming theoretical models into scalable, production-ready solutions. Working with subject matter experts and business leaders, I aim to foster collaborative environments to develop data-driven strategies and propel innovation and informed decision-making.

Skills

Programming Languages	Python (*pandas, pytorch, sklearn*), R (*tidyverse*), SQL
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 (*one-hot encoding, normalization, regularization*), Deep Learning (*traditional neural nets, encoders, transformers*)
MLOps	AWS (*Sagemaker, Athena, Glue, s3*), Spark, Docker, Apache Airflow, Jenkins, Git
Visualization	ggplot2, Matplotlib, Plotly, Shiny, Streamlit, Tableau

Work Experience

The Energy Authority
Bellevue, WA

Machine Learning Engineer III - Supply Chain Forecasting
Dec 2023 - Present

- In my expanded role, I've spearheaded the enhancement of many of our production pipelines, taking a proactive approach to improve our team's flagship products through meticulous testing and scalability improvements.
- Led effort into topic demand forecasting, employing a temporal knowledge graph approach to capture both seasonal and thematic trends effectively.
- Led effort into customer behavior analytics, leverage big data technologies like Spark and Hadoop to draw conclusions from our massive user clickstream dataset that aid in our model development.

Data Scientist II - Supply Chain Forecasting
Jan 2022 - Dec 2023

- Architected and implemented a full-stack, end-to-end ML pipeline for improved forecast accuracy of 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 project not only outperforms the existing heuristic-based model by 3-5% MAPE, but also provided a scalable framework with automated backtesting and dynamic performance reporting functionality.
- Engineered and deployed a cutting-edge Hierarchical Reconciliation framework as a callable API, 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.
- Redesigned unconstrained demand model, which is central to our core SKU demand model. The updated framework includes an ensemble of clickstream heuristic models and self-attention-based transformer models allowing for more accurate and insightful demand forecasting and reducing out-of-stock MAPE by up to 10%.

The Energy Authority
Bellevue, WA

Data Scientist - Analytics
Sep 2019 - Dec 2022

- 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
Seattle, WA

MS in Applied Mathematics

- Professional Excellence Award for distinguished academic merit and professionalism

Santa Clara University
Santa Clara, CA

BSc in Mathematics and Computer Science, Emphasis in Data Science

- Paul R. Halmos Award for Outstanding Academic Record.