

Rick Arko

7750 Westside Dr unit 210 San Diego, CA 92108 | (815) 355-5604 | rickarko@pm.me | <https://github.com/RickArko> | rickarko.com

Accomplished Sr Data Scientist and Machine Learning Engineer with a demonstrated history of transforming data into valuable insights and production software products at the billion dollar enterprise scale.

- **Languages/Tech:** Python, SQL Server, Git, Azure, Bash, Jupyter, Databricks, Docker, HTML, CSS, R, Claude Code, Codex, MongoDB, PowerBI, Tableau, TypeScript
- **Frameworks/Libraries:** Pandas, NumPy, Scikit-learn, FastAPI, Flask, Dash, Plotly, Streamlit, Matplotlib, Seaborn, PyTorch, Pyomo, Polars, PySpark, Nixtla, Optuna, Autogluon, MLFlow, vue.js
- **Algorithms/Skills:** Gradient boosting, Random Forest, Quantile Regression, Cross-validation, Classification, Regression, Conformal Calibration, Statistical Inference, Outlier Detection, Clustering, Tree Methods, Time Series, Shapley, Prophet, SARIMA(X), Exponential Smoothing, Hierarchical Reconciliation

Experience

Shipbob – Chicago, IL

Senior Data Scientist

(Feb 2022 - Present)

- Tech lead/owner of ML Platform (Databricks / AzureML) - the predictive engine for: forecasting, transit time prediction, and inventory optimization on a billion dollar (>100m order) e-commerce platform
 - Built the predictive software behind the labor/transportation forecast, zone skipping, inventory optimization/replenishment, eta prediction and more
 - Built >100 data pipelines (polars/pandas/spark etc.) to augment models with relevant features
- Lead delivery estimation ML efforts resulting in distributed training/inference backtesting/simulation engine processing >50m orders/year
- Built and tested enterprise grade APIs w/ FastAPI/Locust swarms, created reporting dashboards with admin UI's for updating approved overrides by zipcode, carrier, service-level, etc.
- Create rich interactive web applications and model user interfaces and admin panels (Streamlit/replit/vue)
- Built a Streamlit template with laaC, azure app services with entra-auth, and secure Python SDKs for company data feeds) this enabled driven scientist to go from concept to POC in a matter of hours enabling a rich suite of internal tooling for decision making up to the executive level.
- Establish scientific engineering standards across >50 scientific repositories (git, configuration files, linting, unit/integration testing, etc.), mentored and directly managed junior scientists.

Coyote Logistics – Chicago, IL

Senior Data Scientist/ML Engineer

(June 2021 - Feb 2022)

- Lead development for automated load tender acceptance system to allow for automated bidding on Customer freight as well as automated acceptance for incoming load tenders.
- Interview, hire, and mentor Data Science applicants helping to grow from a department of 2 to 10 or more.

Data Scientist

(July 2015 - June 2021)

- Develop and deploy various pricing models from concept to production using ensemble/boosting techniques LightGBM, XGBoost, Catboost, Random Forest Quantile Regression, KNN-regression, etc.
- Implement shapely API to give detailed intuitive reasons for production load recommendations and pricing models.
- Present and communicate key results and ideas to executive leadership.
- Develop and automate forecast pipelines with R, SPSS, and Shiny utilizing SARIMAX and ETS hybrid models
- Optimize supply chain by creating algorithms to match empty assets to customer freight networks.
- Automate data collection and reporting procedures to deliver consistent insights to executive leadership.

RAR-Enabled – (Data Analyst)

(January 2015 – July 2015)

Education

Lake Forest College – Lake Forest, IL

Graduated December 2014

Bachelor of Arts in Economics and Mathematics