ARYA AMARNATH

Looking For Junior Data Scientist / Machine Learning Engineer Roles
+1-970-420-8630 • realyx.arya@gmail.com • www.linkedin.com/in/arya-amarnath6215 • github.com/aamarnath1
• Fort Collins, CO - Can Relocate

Education

Colorado State University — Fort Collins, CO
Bachelor of Science in Data Science with Economics Concentration + Statistics Minor — 08/2020 - 05/2024

- 4th Paradigm, Economics Student Leadership Club, National Society of Black Engineers

American International School of Riyadh — Riyadh, Saudi Arabia
H.S. Degree with International Baccalaureate — 08/2015 - 05/2020

- HL: Physics, English and ITGS | SL: Econ, Math and Spanish AB Initio

Experience

Realyx LLC — Colorado
Real Estate Manager / Portfolio Analyst — 08/2021 - Present

- Wrangle, clean, classify and visualizing data across multiple real estate assets producing monthly analytic reports and new insights on assets with a focus on future success
- Managing and overseeing the daily operations of properties across Colorado coordinating maintenance, lease creation, and maintaining responsive customer service to tenants and owners utilizing property management tool, Stessa

Angi — Denver, CO
Data Science Intern — 06/2023 - 08/2023

- Undertook the development of a predictive NLP text classification model using AWS SageMaker and XGBoost
- Implemented advanced preprocessing / tokenization techniques: tfidf, BERT, and imputation techniques to enhance model accuracy by over 15%
- Integrated model into SageMaker Pipeline, setting the stage for future endpoint deployment
- Shadowed members of the Applied Data Science team, gaining insights into SEO/M, Sales, Customer-Matching with Subscription Clients (contractors) and End-Clients (homeowners)
- Designed a query using SQL and Snowflake to enhance customer-to-sales rep interactions

interviewIA — Denver, CO
Data Analyst Intern — 06/2021 - 01/2022

- Learned essentials of Data Wrangling and Analysis via BA/BI tools: Pendo, Apache Superset, and Tableau
- Created weekly + monthly KPI reports for 1000+ users across 30 companies providing insight into business decisions for 4 months
- Tasked with understanding the basics of back-end development using AWS EC2 servers with Docker and Jenkins for support

DataNow Solutions — Hosur, Tamil Nadu
Full-Stack / DevOps Trainee — 01/2019 - 03/2019

- Introduced to essentials of front and backend development for an open-source data portal that helped calculate tuition fees for families in Riyadh using Angular, NodeJS and Elasticsearch
- Worked closely with software engineers and DevOps team to get familiar with intermediate JavaScript knowledge (OOPs and Inheritance)
- Tasked with gathering, sorting and cleaning raw data into useable data for another data portal that presented data analytics on oil production across wells in the US

Skills

Data Science:
Python • R • SQL • Big Data • Git • Supervised/Unsupervised Machine Learning • EDA + Visualization • Deep Learning • NLP • Computer Vision • Classification + Prediction • MLOps • A/B Testing • Multivariate Statistical Analysis • Time-Series Analysis • HTML & CSS • Angular + Node JS • Numpy + Pandas • Sci-kit + Keras + TF • Gretl/STATA • PowerBI • Tableau • Snowflake • Jenkins • AWS SageMaker • Agile Tools

Academia:
Calc. 3 & Linear Algebra • Intermediate Econometrics + Data Forecasting • Optimization Methods in Data Science • Statistical + Topological Data Analysis • Data Wrangling + Discrete Structures

Certifications: Practical Data Science on the AWS Cloud • IBM Machine Learning Professional - IP

Notable Projects

Palestine Israel Conflict Time-Series Prediction — 2024

Utilizing machine learning and econometric approaches to predict time-series conflict events from ACLED, enabling a deeper understanding of the dynamics within the Israel-Palestine conflict through data-driven analysis.

- Temporal Feature Creation, Log + Lagged Variables, Feature Selection (MI + RF), PCA, Label/One-Hot Encoding and LSTM with Grid Search
- LSTM achieved 85% Accuracy on Test Data Set Confirming High-Level Predictions of Past Data when Predicting Conflict 'Sub-Event Types'

Predicting Breast Cancer Type via Classification Models — 2024

Predicting Benign or Malignant Cancer Cells via General Linear Models, Random Forest, K-Nearest Neighbors and Support Vector Machine

- Using L1/L2 Regularization and PCA techniques, GLM achieved an 96.7% Test Accuracy