Dagart Allison

Los Angeles, United States   dagartga@gmail.com   in/dagart-allison   https://github.com/dagartga

SKILLS

Technical Skills: Python, SQL, Git, Linux, AWS, Databricks, PySpark, SparkSQL, Pandas, Numpy, Statsmodels, SciPy, Seaborn, Matplotlib, TensorFlow, Deep Learning, Artificial Intelligence, A/B Testing, Linear Regression, Logisitic Regression, Classification, Random Forest, XGBoost, Clustering, PCA, Forecasting, Prophet, ARIMA, Holt-Winters, Anomaly Detection, Interpretability, SHAP, Statistics, Chi-Squared, T test, Confidence Intervals, Hypothesis Testing, Model Deployment, Docker, Flask, Streamlit, Deployment.

Soft Skills: Research, Communication, Accountability, Initiative, Collaboration, Critical Thinking, Passion, Presentation, Project Delivery, Idea Generation.

PROJECTS

Business Data Science Portfolio Webpage | https://dagartga.github.io/
- Classification of customer signup based on mail marketing campaign, comparing Logistic Regression, Decision Tree, Random Forest, and KNN.
- Customer Loyalty Score prediction based on customer data, comparing Linear Regression, Decision Tree, and Random Forest.
- A/B Testing for Delivery Club marketing campaigns using Chi-Squared Goodness of Fit.
- Tableau Dashboard for Geo tracking of Earthquake activity.

Boosted Models for Parkinsons Prediction Web App | https://boosted-models-for-parkinsons-prediction.streamlit.app/
- https://github.com/dagartga/Boosted-Models-for-Parkinsons-Prediction
- Machine learning classification of patients future severity based on protein and peptide mass spectrometry data. Utilized automated hyperparameter tuning with hyperopt. Addressed class imbalance with SMOTE. Prediction probability comparing XGBoost, CatBoost, and LightGBM. Optimized based on AUC for each UPDRS 1-3 (0.796, 0.881, 0.721).

Salifort Motors Employee Retention Model Web App | https://salifort-motors-employee-retention.streamlit.app/
- https://github.com/dagartga/Salifort_Motors_Employee_Retention_Project
- Machine learning classification of employees' likelihood to leave the company. EDA discovered 3 main cohorts of employees with high exit rates. Feature engineering used to find feature interaction. Random Forest and XGBoost fine-tuned and evaluated on Recall (0.93) and F1 (0.95). Feature Importance shows average monthly hours, over worked high performer, last evaluation, and satisfaction level are the main predictors.

CERTIFICATIONS

AWS Machine Learning Specialty | 2024

AWS Certified Cloud Practitioner | 2023

Google Advanced Data Analytics | 2023

Springboard Machine Learning Engineer Career Path | 2022

EXPERIENCE

Data Scientist | Tripalink | Los Angeles, CA | August 2022 - December 2023
- Worked with the Databricks framework to analyze website traffic KPIs and Google Ads performance.
- Created a data pipeline for Google Analytics data using Big Query and Databricks to parse data and store as Delta tables. Built a Superset dashboard to deliver insights into website traffic based on Daily Active Users, Daily Sessions, Conversion Rate, and Bounce Rate resulting in a 7.5% increase in website conversions.
- Built an anomaly detection model for website KPIs using Holt Winters, ARIMA, or SARIMA depending on the metric characteristics.
- Developed a ranking algorithm for rental property listings that resulted in 20% more available units being displayed and improved Click Through Rate by 13%.
- Built a financial forecasting pipeline using a Holt-Winters or SARIMA model for larger historical data and SMA for newer financial data. Budget forecast model improved expense budget MAPE from 70% to 10% and reduced the need for managers to do tedious budgeting work.

Property Manager | Realty Services Management Group | Tustin, CA | May 2016 - February 2022
- Working independently as the main property manager for 8 rental properties in a family run business.
- Streamlined operations through systematic batching of tasks to reduce time spent on remedial tasks by 50%. Implemented cloud accounting for remote work.
- Responsibilities and Skills: Basic accounting skills with Quicken. Coordinating repairs with contract workers and tenants. Manage business licenses, utility and mortgage payments, and tax preparation. New tenant acquisition and tenant move out accountability.

Associate Scientist | Dart Neuroscience LLC | San Diego, CA | November 2014 - February 2018
- Took an automated robotics system from non-operational to fully functioning for high-throughput crystallization experiments. The robotics went from a capacity of 30 trays per day to 100 trays per day.

- Utilized and maintained the robotics for the crystallization group. Communicated effectively with the group to ensure experiments were properly setup and delivered results. Produced proteins structures for a new target in multiple forms.
- Worked closely with the Rigaku engineering team to optimize high throughput performance and customization of the CrystalMation robotics. Designed and performed biologic experiments on therapeutic proteins.
- Developed my own experimental pipeline for protein crystallization. Strong skills in searching through scientific journals for impactful information for experimental design.

Research Assistant II | The Scripps Research Institute | La Jolla, CA | April 2013 - October 2014
- Ran the Rigaku automated robotics for protein crystallization experiments.
- Improved work capacity from 50 trays per day to 100 trays per day. Improved accuracy of automation from 80% to 95%.
- Harvested crystals for the Joint Center for Structural Genomics.

Medical Laboratory Technician | PacifiCord | Irvine, CA | February 2012 - February 2013
- Worked with focus and strong attention to ensure the irreplaceable product, cord blood, was processed and stored effectively for life saving therapies.
- Critical decision making on processing protocol to ensure the highest level of stem cells were saved and stored.
- Detailed note taking and following regulatory guidelines for stem cell therapies.

Research Associate | Eli Lilly & Company | San Diego, CA | October 2008 - October 2009
- Maintained the lab robotics, participated in the evaluation and purchasing of new robotics. Designed and scripted protocols for the robotics.
- Developed co-crystallization experiments for proteins which led to novel structures and produced relevant data for lead compound development. Extensive experience with crystal looping, cryoprotectants, and cryocooling. Analyzed beamline screen shots to create a consistent high resolution freezing protocol.
- Successfully crystallized a novel membrane protein to 1.2 Angstrom resolution. Performed the following membrane specific techniques: lipidic cubic phase and detergent screening. Optimized membrane protein crystallization for screening of 1,000 mixture fragments requiring soaking of 300+ crystals with high resolution requirements of 2.5 Angstrom data

Research Associate, Biologist | SGX Pharmaceuticals | San Diego, CA | May 2004 - October 2008
- Crystallized soluble proteins using vapor diffusion techniques in hanging drop and sitting drop experiments. Developed crystallization strategies for various proteins, including kinases, GTPase, and histone methyltransferase resulting in high resolution crystal structures.
- Delivered approximately 30 crystallization trays scored daily for 3 years resulting in a discerning eye for protein behavior in various conditions.
- Developed co-crystallization experiments for proteins which led to novel structures and produced relevant small molecule data for drug development. Used experimental data and molar excess ratios to find optimal concentration of compound.

EDUCATION

Bachelor of Science, Biology | University of California, San Diego | La Jolla, CA | 2003
- Men's Swimming Team Captain and Division II All-American