Zachary Raup

Data Scientist
Zachary.Raup@mail.com • linkedin.com/in/zachary-raup • zraup.github.io/Zachary-Raup

Summary

Physics graduate and certified data scientist with experience applying machine learning in healthcare, retail, and astrophysics. Proficient in Python, SQL, and deep learning. Skilled in data cleaning, model development, evaluation, and communicating insights through interpretable visualizations and metrics.

Technical Skills

Python, SQL, PyTorch, scikit-learn, Pandas, Deep Learning, Computer Vision, Tableau, Power BI, Jupyter, Git, AWS

Projects (Available on GitHub Portfolio)

Chest X-Ray Pneumonia Detection Jun 2025
Classifiy Chest X-Ray images with an ensemble pipeline using ResNet18, DenseNet121, and EfficientNet-B0 with transfer learning, 3-fold cross-validation, and **Grad-CAM** interpretability. **Achieved 91% accuracy**
Skills: Deep Learning, PyTorch, CNNs, Cross-Validation (CV), Grad-CAM, Medical Imaging

Walmart Sales Forecasting with Machine Learning Mar 2025
Developed and evaluated regression models to forecast weekly retail sales using economic and calendar data. Achieved $\mathbf{R^2 = 0.988}$ and **RMSE** $\approx$ **\$61K** using XGBoost, outperforming other models. Included feature importance, residual diagnostics and 5-fold CV.
Skills: Regression Analysis, Hyperparameter Tuning, Feature Engineering, XGBoost, LightGBM, Residual Diagnostics

Predictive Maintenance in Manufacturing Jul 2025
Trained XGBoost, Random Forest, and Logistic Regression models to detect machine failures from sensor data. XGBoost achieved **98.6% accuracy** and perfect AUC. Used **SHAP** to explain predictions and rank feature importance.
Skills: Classification, XGBoost, SHAP, SMOTE, Model Interpretability, Confusion Matrix, ROC Analysis

Experience

Senior Manufacturing Tech | **Manufacturing Tech II** | **Manufacturing Tech I** — DSM-firmenich Mar 2023–Present
Programmed CNC machines and applied **data-driven approaches** to optimize Class III medical device production under GMP standards. Analyzed machine logs, process data, and operational metrics to identify inefficiencies, reduce downtime, and enhance process reliability through **statistical insights** and **visualization tools**.

Undergraduate Researcher — Kutztown University Oct 2021–Mar 2023
Developed **Python-based models** to analyze transit and radial velocity datasets for exoplanet and binary star systems. Applied **signal processing** and **curve-fitting techniques** to extract and estimate key parameters, contributing to data-driven discoveries.

Undergraduate Astronomy Researcher — Univ. of Southern Queensland (Australia) May–Aug 2022
Processed and analyzed photometric data from the TESS mission and Mt Kent Observatory using **Python**. Performed **data cleaning**, **light curve modeling**, and **statistical analysis** to refine transit timing estimates, aiding in the validation of exoplanet candidates.

Education

DataCamp — Data Scientist with Python Career Track Sep 2024
Topics: Python, ML, SQL, Git, Data Visualization

Kutztown University of Pennsylvania — B.S. Physics, *Summa Cum Laude* Dec 2022
GPA: 3.92
Awards: Chambliss Student Academic Achievement Award, Roy W. Hamme Memorial Award, KURF Grant, NSF IRES Grant

Certifications

Data Scientist Associate, Data Analyst Associate, Python Data Associate, SQL Associate — DataCamp

Publications & Presentations

Schulte, J., Raup, Z., et al. (2024). *Migration and Evolution of Giant ExoPlanets (MEEP) I.* arXiv:2401.05923
Presented exoplanet transit modeling research at the 241st AAS Meeting (Seattle, WA) and the 42nd CPC Astronomers' Meeting (Gettysburg, PA)