Oran Looney
Data Scientist

608-234-1066
olooney@gmail.com
linkedin.com/in/oran-looney
oranlooney.com
github.com/olooney

Objective

Seeking a senior individual contributor role where I can focus on developing cutting-edge models that solve complex real-world problems and deliver meaningful value.

Experience

Company	Job Title	Time Period
QBE	Lead Machine Learning Engineer	2024-Present
Frazier Healthcare Partners	Director, Data Science and AI	2022-2024
Healthgrades	Senior Data Scientist	2014-2021
Healthgrades	Senior Software Architect	2013-2014
CPM Marketing Group Inc.	Director of Software Development	2012-2013
CPM Marketing Group Inc.	Software Engineer	2007-2012
Epic Systems	Interface Analyst	2005-2007

Education

Degree Earned	Institution	Graduated
M.S. Physics	Indiana University	2004
M.A. Mathematics	Indiana University	2004
B.S. in Physics & Mathematics	University of Oregon	2001

Skills

Core Skills: Machine Learning • Statistics • Software Development • Data Visualization

Tools: Python • NumPy • Pandas • matplotlib • Keras • Hyperopt • scikit-learn • SciPy • statsmodels • SQL • PostgreSQL • Linux • Docker • Databricks • Spark • AWS • Azure • R • Tidyverse • ggplot2

Modeling Expertise: Neural Networks • Generalized Linear Models • Regularized Regression • Hypothesis Testing • ANOVA • Random Forest • Boosted Trees • GAM • ARIMA • k-means • PCA • GMM • MARS • LLMs • Vector Embeddings

Soft Skills: Machine Learning Best Practices • Statistical Methodology • Software Architecture • Scientific Communication • Mentoring and Training • Design of Experiments • Project Management

Notable Projects

QBE

• Led development of a REST API for over a dozen insurance pricing models. Architected a framework to reduce code duplication across models, conducted fuzz testing, and introduced logging, monitoring, semantic versioning, deployment pipelines, and automated documentation generation.

• Built pipelines to extract structured information from documents using OCR and LLMs. Trained junior data scientists to build similar pipelines. Developed a benchmark framework to evaluate such pipelines.

• Optimized a RAG chatbot by introducing a repeatable benchmark and conducting a series of structured experiments to identify the optimal chunking strategy, vector database hyperparameters, model version, and system prompt.

Frazier

• Developed a classification model to identify 'at-risk' students given a longitudinal dataset of clinical notes and demographics.

• Conducted a statistical analysis of the impact of operational quality metrics (quality of treatment, clinician experience) on third-party outcome measures (Vineland, VB-MAPP) for treating children with autism with ABA therapy which found statistically significant effects for key explanatory variables.

• Advised on a project to predict medication adherence from clinical interviews conducted by pharmacists. The focus was on using interpretable ML techniques to derive insight into which questions had the greatest effect on medication adherence for different classes of drugs.

• Developed a neural net model to estimate the probability of a physician responding to a survey based on the physician's profile, the timing and length of the survey, and the incentive offered.

• Consulted on a project to use an LLM to identify and flag sections of free-text medical notes relevant to certain conditions such as COPD or Diabetes. This project was designed to save time when reviewing medical records for correctness.

Healthgrades

• Worked on a large suite of neural net models used to estimate a patient's future healthcare utilization. Migrated model training from R to Python and Keras. Introduced multi-label training to decrease training time and benefit from transfer learning. Implemented minority oversampling and scaled the training set size to ten million medical records. Performed hyperparameter optimization, augmented the data with census data, evaluated new data features, and optimized inference time.

• Developed an LDA topic model to automatically group providers into sub-specialties using only historical claims data.

• Developed a time series model to predict website traffic by geography and category. This model was used to estimate feasible contract sizes and set delivery guarantees for promoted search campaigns.

Selected Writing

Large Language Models: Multimodal LLMs, LLM Benchmarks

Machine Learning: PCA, Backpropagation

Statistics: Cross Validated Answers, Logistic Regression

Numerical Programming: Playfair Ciphers, Fast Fibonacci

Recreational Mathematics: Angle Addition, Kaprekar, History of Computing