Robert **Brown**

www.rbrown.org • elreb • rbrtbrwn
robert@rbrown.org • 510.xxx.xxxx

SKILLS

Programming and ML Frameworks: Python (PyTorch, TensorFlow, scikit-learn, Pandas, NumPy), SQL, R, PySpark
LLM and AI Tools: LLMs (Fine-tuning, RAG, prompt engineering, LangChain, LangGraph, MCP, Hugging Face Transformers
Cloud and Infrastructure: AWS (SageMaker, Lambda, EC2, S3, OpenSearch), GCP, MLflow, Airflow, Docker, CI/CD
Data Engineering: ETL pipelines, real-time data processing, API integration, database design
ML Specializations: NLP, time series forecasting, survival analysis, A/B testing, ensemble methods, neural networks

EXPERIENCE

MACHINE LEARNING ENGINEER | Atropos Health — December 2022 - Present • San Francisco, CA

- Led development of "ChatRWD", an agentic AI application using LLMs and RAG to accelerate Real-World Evidence generation from days to minutes, eliminating human bottlenecks in healthcare research workflows
- Architected and deployed production LLM systems for healthcare data analysis, implementing fine-tuning strategies and prompt engineering techniques to achieve domain-specific performance improvements
- Optimized core ML production codebase achieving 6x performance gains through parallel processing and 100x memory efficiency improvements, handling large-scale healthcare datasets
- Built end-to-end ML pipelines integrating traditional ML and LLM components for automated evidence generation in pharmaceutical and healthcare decision-making

DATA ENGINEER | Berlin Brands Group — October 2021 - December 2022 • San Francisco, CA

- Architected scalable data infrastructure using AWS services (S3, Lambda, EC2) and Snowflake, building automated ETL pipelines to process multi-source data streams and provide actionable insights for executive decision-making
- Developed predictive revenue forecasting system using time series models (ARIMA, LSTM, Prophet) integrated into custom Django application, enabling data-driven brand acquisition decisions
- Deployed machine learning models to production environment with monitoring and automated retraining capabilities

DATA SCIENTIST | Los Angeles Co. Public Health Dept. — May - October 2021 • Los Angeles, CA

- Created production-ready dashboards using Dash and Plotly for real-time epidemiological surveillance and Airflow-based ETL systems for whole genome sequencing data processing, serving 10+ million LA County residents
- Implemented statistical modeling for disease trend analysis and outbreak detection, informing public health policy decisions

HEALTH DATA SCIENCE FELLOW | Insight Fellows — September - November 2020 • San Francisco, CA

- Built predictive churn model achieving 78% ROC AUC using advanced feature engineering and survival analysis techniques for healthcare client retention
- Delivered data-driven recommendations to healthcare stakeholders, demonstrating ROI of ML-powered customer segmentation strategies

EPIDEMIOLOGIST II | Alameda Co. Public Health Dept. — August 2018 - August 2021 • Oakland, CA

- Led COVID-19 data operations as Data Chief for Incident Command System, managing 25+ data professionals and building county's first agent-based modeling platform for disease forecasting
- Designed production ETL pipelines and SQL Server databases for surveillance of 70+ communicable diseases, ensuring real-time data availability for public health response and resource allocation optimization, producing quarterly disease reports and comprehensive annual surveillance reports

RESIDENT DATA SCIENTIST | Method Data Science — November 2018 - April 2019 • Remote Work

- Developed ensemble ML algorithm (77% AUC) for hip replacement prediction and built customer segmentation dashboard using unsupervised learning, resulting in 13% customer base increase

EDUCATION

UC BERKELEY | MPH in Biostatistics & Epidemiology — May 2017 | Berkeley, CA

UC DAVIS | BS in Neurobiology, Physiology, & Behavior, BA in Psychology — May 2011 | Davis, CA

RESEARCH & IP

Patent

- "Systems and Methods for Automated Evidence Generation" - US Patent Application 20250078969 A1 (Filed 8/29/2023)

Selected Publications

- Low, Y. S., Jackson, M. L., Hyde, R. J., Brown, R. E., et al. (2025). Answering real-world clinical questions using large language model, retrieval-augmented generation, and agentic systems. DIGITAL HEALTH, 11.
- Brown, R. E., Turner, C., Hern, J., & Santos, G.M. (2017). Partner-level substance use associated with increased sexual risk behaviors among men who have sex with men in San Francisco, CA. Drug and Alcohol Dependence, 176–180.
- **Additional Publications:** 4 peer-reviewed publications in clinical informatics, tuberculosis epidemiology, clinical microbiology, and HIV surveillance across Journal of Public Health Management & Practice, Journal of Clinical Microbiology, and Observational Health Data Sciences and Informatics.