SILVIA MIRAMONTES, PhD(c)

San Francisco Bay Area, CA ◇ Silvia.Miramontes@ucsf.edu ◇ (619)-862-0895

SUMMARY AND QUALIFICATIONS

Versatile data scientist trained in clinical informatics, machine learning, and epidemiology, looking to transition into industry to work on larger studies of real world evidence. Technical expertise in big data management (e.g. EHR), machine learning, causal inference, NLP, and analysis of complex clinical data, resulting in 10 peer-reviewed publications and 1 NIH-funded grant. Adept at interdisciplinary science, leading and contributing to projects across clinicians, epidemiologists and research teams.

RESEARCH AND PROFESSIONAL EXPERIENCE

Graduate Student Researcher January 2021 - Present
UCSF Bakar Computational Health Sciences Institute

Project 1: Early Detection of Disease leveraging EHR and Clinical Notes

· Deployed machine learning models (Random Forests, SVMs, Logistic Regression) on longitudinal EHR and clinical notes to predict cognitive impairment, improving risk stratification to enable early clinical intervention.

· Engineered patient-level longitudinal datasets by reconstructing clinical visit timelines, integrating ICD-coded diagnoses, demographics and unstructured text data to enhance predictive modeling of disease trajectories.

Project 2: Modeling Disease Progression Using Longitudinal EHR Data

· Led real-world evidence studies modeling disease progression using longitudinal EHR data, applying Cox proportional hazards models for time-to-event risk estimation.

· Applied causal mediation analysis to evaluate direct and idirect effects of social and clinical risk factors on disease trajectories.

Project 3: Clinical Diagnostic Trajectories in EHR Data

· Designed and built longitudinal diagnostic trajectory datasets to model disease progression patterns, including diagnostic reversion of various disease stages.

· Led study design and cohort definition, feature selection, and model validation to ensure clinically meaningful results.

· Collaborated closely with clinicians to audit neuropsychological assessments and ensure clinical relevance of data-driven models.

Project 4: Extraction of Social Determinants of Health from Clinical Notes

· Optimized Named Entity Recognition (NER) pipelines using cTAKES to extract housing and food insecurity indicators from clinical notes in EHR data by conducting data preprocessing and text analysis to improve information extraction accuracy.

Project 5: Longitudinal Analysis of Biomarkers and Disease Progression

· Evaluated the relationship between longitudinal biomarker trends and cognitive decline over 14 years in a national cohort, modeling trajectories of memory change over time.

Project 6: Classification of Organ Abnormalities from Radiology Reports

· Built and optimized multi-class NLP pipelines to classify organ-specific abnormalities from radiology reports, developing models (Naive Bayes, SVMs, fine-tuned BERT) and integrating clinical input from oncologists to improve diagnostic accuracy and interpretability.

Staff Research Associate April 2021 - September 2021
UCSF Memory and Aging Center

· Developed a CNN-based tool for automated cell segmentation in brain tissue images, deploying the pipeline via AWS EC2/S3 and JupyterHub to support scalable, real-time image analysis.

Research Assistant and Computer Systems Engineer January 2019 - January 2021
Lawrence Berkeley National Laboratory

- Designed and optimized deep learning-based image analysis frameworks for material science applications, leveraging CNNs and high-performance computing (HPC) techniques (multi-core CPUs and GPUs) to improve computational efficiency and scalability.
- Developed CNN models for automated material phase classification, streamlining workflows and significantly reducing manual inspection time in large-scale imaging datasets.

EDUCATION

University of California, San Francisco
Doctorate in Biological and Medical Informatics 2021 - June 2025
University of California, Berkeley
Master of Information and Data Science 2019 - 2020
Bachelor of Arts in Applied Mathematics 2016 - 2018

TECHNICAL SKILLS

Programming: Python, R, SQL, PySpark, Bash, TensorFlow, PyTorch.

Modeling: Predictive modeling (Random Forests, SVMs, BERT, U-Net), regression modeling (linear, logistic, survival), causal inference, longitudinal mixed models, multiple imputation for missing data, natural language processing (NER, cTAKES, text classification, summarization, bias mitigation).

Cloud Computing: AWS (EC2, EMR, S3, Lambda, SageMaker), Google Cloud Platform (multi-core CPUs), Docker, JupyterHub, DynamoDB, Serverless Architecture.

SELECTED PUBLICATIONS

Miramontes S., Khan U., Ferguson E.L., Sirota M., Glymour M.M. The association of cholesterol levels with memory and memory change over a 14-year period in a US national cohort. Alzheimer's and Dementia: Translational Research & Clinical Interventions. doi: 10.1002/trc2.70021.

Miramontes S., Pereda Serras C., Woldemariam S., Khan U., Li Y., Tang A.S., Tsoy E., Oskotsky T., Sirota M. Alzheimers Disease as a Womens Health Challenge: A Call for Action on Integrative Precision Medicine Approaches. NPJ Womens Health. doi: 10.1038/s44294-024-00021-3

Tang A, Rankin KP, Cerono G, **Miramontes S.**, et al. Leveraging Electronic Health Records and Knowledge Networks for Alzheimer's Disease Prediction and Sex-Specific Biological Insights. Nat Aging 4, 379-395 (2024). doi: 10.1038/s43587-024-00573-8.

Tang A.S., Woldemariam S.R., **Miramontes S**., Norgeot B., Oskotsky T.T., Sirota M. Harnessing EHR data for health research. Nature Medicine. doi: 10.1038/s41591-024-03074-8

Kim M.H., **Miramontes S.**, Mehta S., Schwartz G.L., Kim Y.J., Yang Y., Hill-Jarret T.G., Cevallos N., Chen R., Glymour M.M., Ferguson E.L., Zimmerman S.C., Choi M., Sims K.D. Extracting Housing and Food Insecuruity Information from Clinical Notes Using cTAKES. Health Services Research. doi.org/10.1111/1475-6773.14440.

Ferguson E.L., Mehta S., **Miramontes S.**, Choi M., Kim Y.J., Hill-Jarret T.G., Cevallos N., Yang Y., Zimmerman S.C., Chen R., Kim M.H., Sims K.D., Schwartz G.L. Identifying and addressing housing insecurity in older patients: trends, referrals, and inequities in a California health system. The Gereantologist. In Press.