Nasreen Buhn

103 Windsor Ave,
Narberth, PA 19072
nbuhn@calpoly.edu • (484) 433-6379 • https://www.linkedin.com/in/nasreenbuhn

EDUCATION

California Polytechnic State University, San Luis Obispo, CA

Bachelor of Science in Biological Sciences, Concentration in Cellular and Molecular Biology, Minors in Computer Science and Bioinformatics, *Cum Laude*   June 2025

Relevant Coursework: Bioinformatics Algorithms, Bioinformatics Capstone I & II, Deep Learning, Introduction to Data Science, Data Structures, Applied Experimental Design and Regression Models, Project Based Object Oriented Programming and Design

RESEARCH EXPERIENCE

California Polytechnic State University, Bioinformatics Research Group, Biological Sciences and Computer Science and Software Engineering Departments
Undergraduate Research Assistant, January 2024–June 2025

Advisors: Jean Davidson, Paul Anderson
Collaborated on the design and development of a zero-shot entity resolution methodology assessing the impact of context in NIH-funded chronic lower back pain (CLBP) knowledge graph research. Evaluated entity pairs to build a ground truth dataset and helped craft LLM prompts for entity resolution across varying context levels.

Advisors: Paul Anderson, Javin Oza
Assessed the performance of protein clustering and alignment algorithms for classifying carbonic anhydrase proteins. Applied SVD and K-means clustering, evaluating the impact of sequence length and clustering conditions using silhouette scores. Compared multiple sequence alignment (Clustal Omega) and multiple structure alignment (mTM-align, FoldMason) methods with phylogenetic tree analysis.

California Polytechnic State University, Computer Science and Software Engineering Department
Summer Undergraduate Research Program, College of Engineering, June 2024–September 2024
Undergraduate Research Assistant, September 2024 – Present

Advisor: Jonathan Ventura
Trained, tested, and evaluated supervised deep learning models for fluorescence microscopy image denoising as part of NIH funded research. Developed a stitching algorithm to reconstruct full-resolution images from overlapping, denoised segments with varying light intensities. Assessed denoising performance using PSNR and SSIM image quality metrics.

QuantumCyte, California Polytechnic State University, Bioinformatics Capstone *(DATA 441 & 442)*
Undergraduate Bioinformatics Consultant (Internship), April 2025–June 2025

Advisors: Jean Davidson, Paul Anderson
Co-developed a modular, interactive RNA-seq analysis pipeline in R using Google Colab. Built core modules for data cleaning, differential gene expression analysis (DESeq2), PCA, gene set enrichment (fgsea), and GO analysis, enabling real-time adjustment of parameters such as p-value thresholds, sample groupings, and sample size assumptions. Created integrated visualizations illustrating gene expression patterns, PCA results, and sample size estimates.

RELATED PROFESSIONAL EXPERIENCE

Pheast Therapeutics
Bioinformatics & Data Science Consultant, August 2025–Present

Jazz Pharmaceuticals
Regulatory Strategy Intern, June 2023–August 2023
Conducted a comprehensive review of Real-World Evidence (RWE) use in FDA New Drug Applications, Biologics License Applications, and supplemental submissions, examining purposes of use, primary endpoints, and cases where submissions were unsuccessful due to RWE control arm limitations.

PUBLICATIONS

Lin, D., Koenig, C., Kaplan, S., Bittner, M., Paraiso, M., **Buhn, N**., et al. (2025). Investigating the Impact of Context on Zero-Shot Entity Resolution: Applications in Chronic Lower Back Pain. 2025 8th International Conference on Information and Computer Technologies (ICICT), IEEE. https://doi.org/10.1109/ICICT64582.2025.00067

Anderson, P., Lin, D., Davidson, J., Migler, T., Ho, I., Koenig, C., Bittner, M., Kaplan, S., Paraiso, M., **Buhn, N**., et al. (2024). Bridging Domains in Chronic Lower Back Pain: Large Language Models and Ontology-Driven Strategies for Knowledge Graph Construction. IWBBIO 2024. Lecture Notes in Computer Science, Springer. https://doi.org/10.1007/978-3-031-64636-2_2

CONFERENCE PRESENTATIONS

ORAL PRESENTATIONS

Bittner, M., **Buhn, N.,** et al. (2025, June) Bioinformatics Capstone 2024: Solving real world genomics problems with biotechnology collaborators: Developing a Comparative Transcriptomics Pipeline to Identify Key Biomarkers for Pre-metastatic Tumor Buds. Bailey College Student Research Conference, California Polytechnic State University, San Luis Obispo, CA.

POSTER PRESENTATIONS

Adunur, S., **Buhn, N.,** Hagen, G., Ventura, J. (2024, September) Deep Learning for Microscope Image Denoising. Poster session at Summer Undergraduate Research Program Symposium, California Polytechnic State University, San Luis Obispo, CA.

Koenig, C., Bittner, B., Kaplan, S., Paraiso, M., **Buhn, N**., et. al. (2024, June) Novel Approaches to Literature Analysis and Entity Resolution for Chronic Lower Back Pain; Knowledge Graphs and Large Language Models. Poster session at Bailey College Student Research Conference, California Polytechnic State University, San Luis Obispo, CA.

Buhn, N. (2023, October). Real World Evidence in FDA Drug Approvals. Poster session displayed at The Organization for Professionals in Regulatory Affairs (TOPRA) Annual Symposium, Lisbon, Portugal.

WORKS IN PROGRESS

Buhn, N., Adunur, S., Hamilton, J., Levis, S., Hagen, G., Ventura, J. Comparison of Deep Learning Approaches for Extreme Low-SNR Image Restoration. Research paper in preparation to be submitted summer / fall 2025.

RELEVANT PROJECTS

VIDA: Visual Intelligence for Detecting Anemia, California Polytechnic State University, Deep Learning (CSC 487)
Co-developed an ensemble model approach for anemia detection from peripheral blood smear images and CBC data. Fine-tuned VGG-16 architecture on images, achieving an F1-score of 80.18%, which increased to 85.02% with the addition of CBC data in the ensemble model. Assessed model performance using Grad-CAM, identifying prediction drivers and visual cues that influenced model decisions.

Predicting Cancer Incidence Rates from Toxic Chemical Spills, California Polytechnic State University, Introduction to Data Science (DATA 301)
Built a machine learning pipeline to predict county-level cancer incidence rates from chemical spills reported in the EPA's Toxic Release Inventory. Used PubChem's API to extract carcinogen classifications for each release. Conducted extensive data cleaning and feature engineering, including encoding, scaling, and transformations. Trained multiple regression models (Random Forest, KNN, SVR) to evaluate predictive relationships.

SERVICE

Alzheimer's Association
Walk to End Alzheimer's Volunteer, 2021 – 2024
Assisted with team fundraising events, contributing to over $60,000 raised annually (2022–2024), and supported annual Walk to End Alzheimer's operations.