Suchit Bhayani

suchit.bhayani@gmail.com | (925) 875-8737 | linkedin.com/in/suchit-bhayani | suchitbhayani.github.io

EDUCATION

University of California San Diego San Diego, CA
Bachelor of Science in Data Science, Minor in Mathematics Expected: June 2027
- GPA: 3.9/4.0
- Relevant Coursework: Deep Learning, Computer Vision, Data Mining, Data Visualization, Probability, Statistics

EXPERIENCE

Nike June 2025 - Aug. 2025
Data and Machine Learning Engineer Intern
- Designed governance frameworks for enterprise ML, generative AI, and BI infrastructure, enforcing compliance
- Engineered Databricks workflow leveraging MLflow to automatically flag ethics violations in deployed models
- Scaled a recommender system using PySpark, enabling product similarity recommendations for consumers

UC San Diego | Data Science Department Jan. 2025 - Present
Teaching Assistant
- Tutor for DSC 20, 30, and 40B, covering data structures and algorithms, object-oriented design, graph theory, and more
- Apply understanding of Python and Java via office hours and online question-answering platform

UC San Diego Health | Li Lab Oct. 2024 - Present
Machine Learning Researcher
- Engineer and clean datasets with pandas to build regression models to identify and analyze drivers of stem cell renewal
- Perform time series differential gene expression analysis in R to identify biologically significant gene clusters

WorldQuant June 2024 - Dec. 2024
Quantitative Research Consultant
- Research, implement, and backtest over 500 trading strategies with FastExpression for potential portfolio integration
- Present research of high performing alpha strategies (2.83 Sharpe) to portfolio managers and executives

Digital Prudentia June 2024 - Sep. 2024
Data Science and Engineer Intern
- Utilized retrieval-augmented generation (RAG) with Azure OpenAI to develop a skin cancer detection model
- Linked documents with text and image data to the model using Azure AI Search, enabling contextual insights
- Applied prompt engineering and hyperparameter tuning to fine-tune the model and optimize performance

PROJECTS

Music Recommender System
Full-Stack Development, Software Engineering, Agile Development, REST API, Containerization, MERN Stack
- Built music recommender and analyzer with React, Express/Node.js, and FastAPI, containerized using Docker
- Integrated MongoDB, Spotify OAuth 2.0, and managed API communication between frontend, backend, and ML service
- Developed CI/CD pipeline with GitHub Actions to automate testing, building, and deployment across all services

Personalized AI Health Insights
Big Data, Scalable Systems, Natural Language Processing (NLP), Healthcare AI, Data Engineering
- Developed a data pipeline in dask to efficiently process and analyze millions of Apple Watch health data points
- Engineered health insight generation pipelines with fine-tuned HuggingFace LLMs, enabling health-specific inferences
- Built an interactive dashboard using plotly to visualize trends, insights, and actionable recommendations

Accelerating ML with Automated Feature Engineering
AutoML, Large Language Models, Statistical Feature Selection
- Automated an ETL pipeline with OpenRouter API for integrating LLM domain knowledge into the AutoML paradigm
- Validated performance of generated features using XGBoost and RandomForest models across 3 benchmark datasets

SKILLS

- **Languages:** Python, Java, JavaScript, SQL (PostgreSQL, NoSQL), HTML/CSS
- **Libraries:** pandas, PySpark, dask, scikit-learn, NumPy, scipy, plotly, seaborn, Matplotlib, BeautifulSoup
- **Frameworks:** PyTorch, TensorFlow, Keras, React, Express, Node.js, FastAPI
- **Tools:** AWS (S3, EC2), Azure (DevOps), Docker, Git, GitHub, MLflow, Jupyter, Linux/Unix
- **Interests:** Gym, nature, music, basketball, football