KEVIN BUI

📞 (818)-823-1376 ✉ kevinbui806@gmail.com 💼 https://www.linkedin.com/in/kevin-bui 🏠 kevinbui2.github.io

Education

University of California San Diego
Bachelors of Science in Data Science *La Jolla, CA*

Experience

Medtronic **June 2024 – Present**
Data Software Engineer I *Remote*
- Developed RAG Chatbot by webscraping lesson documents for users to search and understand training material.
- Integrated LLM and ML models to analyze and extract key information from contracts, as well as predicting resource and financial assignments.
- Implemented Snowflake data warehouse and Apache Airflow ETL pipeline for PMO team's data mart.
- Designed Power BI reports for Project Managers and Finance Teams across the company by aggregating multiple data sources for easy analysis of their resources and finances.
- Streamlined collection and presentation of project KPIs by automating process with Power Apps and Power BI.

Medtronic **June 2023 – Aug 2023**
Data Engineer Intern *Minneapolis, MN*
- Created ETL tool for PMO team to automate workflow for their resource data into suitable format for finance team.
- Designed Power BI reports for stakeholders and PMO team to identify important metrics within their data.
- Developed dbt data pipeline for Medtronic's DRM training courses to automate workflow.
- Migrated Medtronic's DRM project database to new platform using Azure Data Factory.

SeeByte **Aug 2022 – Dec 2022**
Data Engineer Intern *San Diego, CA*
- Worked alongside the SeeByte ML Team to facilitate data preparation using Computer Vision Annotation Tool.
- Assisted the team on developing automated tools in data processing for future machine learning projects.
- Trained ML model to improve Automatic Target Recognition of mine-like objects for ROV Autonomy.
- Developed dashboard using Flask that allows user to control Mobile Gateway Buoy Vehicle.

Acoustical Engineering Services **Jun 2021 – Aug 2022**
Data Analyst / Acoustical Engineer *Los Angeles, CA*
- Worked at Acoustical Engineering Services to collect, clean, and analyze data for clients, verifying acoustics were in compliance with regulation for each client's proposal.
- Created visualizations for clients' data, clarifying the necessary changes on designs of buildings for compliance.
- Assisted in modeling buildings from clients' design proposals using AutoCAD, testing acoustics.

Projects

SDG&E Overhead Structure Detection | *Docker, PostgreSQL, DETR* **Jan 2024 – March 2024**
- Worked with SDG&E to help create data pipeline for validation of assets using Google Street View API.
- Developed automated workflow from initial data scraping of images to validating assets with SDG&E database.
- Trained DETR object detection model to detect various SDG&E assets from Google Street View images.
- Created Docker environment to spin up container for accessing SDG&E database and local database.

Extreme 2018 Electoral Swings Analysis | *Redis, MongoDB, PostgreSQL, LlamaIndex* **May 2023 – June 2023**
- Utilized Psycopg2 to query news articles from SDSC database to clean and store into Redis and MongoDB.
- Trained LlamaIndex model by ingesting queried articles to create an LLM app to answer possible reasons for extreme electoral swings by use of input prompts.

Technical Skills

Languages: Python, SQL, Scala, R, JavaScript, TypeScript, Java, DAX, M, Power Fx, HTML, CSS
Technologies: PostgreSQL, SQL Server, SQLite, Redis, Cassandra, MongoDB, Neo4J, Solr, Snowflake, Kubernetes, Spark, Kafka, Airflow, AWS, Azure, GCP, dbt, Glue, Data Factory, Docker, Tableau, Power BI, Power Apps, Excel
Other: NumPy, Pandas, Dask, PySpark, PyTorch, TensorFlow, Scikit-learn, Matplotlib, D3.js, Git, Bash
Certifications: AWS Certified Cloud Practitioner