

Joshua Payapulli

Los Angeles, CA | joshua.payapulli@gmail.com | payapull@usc.edu | <https://payapulli.github.io/>

EDUCATION

University of Southern California

M.S. Applied Data Science

Relevant classes: Data Mining, Machine Learning, Databases, Optimization, UX design and strategy, Fairness in AI

Los Angeles, CA

Aug 2023 - May 2025

GPA: 3.88

The University of Manchester

B.S. Computer Science.

Relevant classes: Fundamentals of AI, Algorithms, Machine Learning and Optimization, etc

Manchester, United Kingdom

Sept 2018 - June 2021

Grade: First Class Honors

EXPERIENCE

Averlon

United States (Remote)

Software Engineer Intern

July 2025 - Present

- Built ETL pipeline in Python to parse Terraform configs into graph-structured inputs for LLM-based risk summarization, leveraging retrieval-augmented generation (RAG) and in-context learning (ICL) to generate interpretable blast radius forecasts and automatically tag high-risk infrastructure pull requests.
- Modeled attack-chain graphs as Weighted Constraint Satisfaction Problems (WCSPs) and implemented min-cut optimization algorithms to identify lowest-cost mitigation strategies for software vulnerabilities; built an adjacency matrix based reachability algorithm to expose reachable assets within constrained IP regions.
- Designed agentic workflow to analyze IAM relationships across AWS/Azure multi-cloud systems; integrated LLMs to reason over ARNs, roles, and policies, producing structured outputs to support cloud infrastructure security decision-making.

Align Technology (Invisalign/iTero)

San Jose, CA

Data Analytics Intern

June - Aug 2024

- Implemented a K-means clustering algorithm with cosine similarity to remove duplicates from a policy document corpus and categorize the remaining documents, reducing the document corpus by 70% and streamlining document management.
- Built ETL pipelines using Python and SQL, ensuring data correctness when integrating compliance data into Azure.
- Conducted exploratory data analysis and applied statistical modeling to identify inefficiencies and risks in legal approval workflows, implementing automated solutions that increased workflow efficiency by 30%.
- Presented Power BI dashboards to legal executives, highlighting actionable insights to support data-driven decision making.

AND Digital

Leeds, United Kingdom

Full-stack Software Engineer

Oct 2022 - July 2023

- Built a real-time office attendee tracking feature on mobile app, integrating Firebase's database with React Native. This enhancement provided instant updates on user presence, improving oversight and coordination for 100+ team members.
- Implemented an event management feature by integrating React.js (Redux) with Firebase, enabling real-time event updates, eliminating conflicts, and optimizing resource utilization through data-driven insights/automated availability tracking.
- Designed and implemented RESTful APIs to manage event subscriber notifications, enabling real-time updates, reducing delivery latency, and ensuring seamless integration between the event management system and mobile/desktop clients.

PricewaterhouseCoopers (PwC)/ First Abu Dhabi Bank

Dubai, United Arab Emirates

Consultant/ Business Analyst - Contract

Aug 2021 - July 2022

- Managed compliance data for the Customer Protection Regulation (CPR) program at Emirates NBD, performing regular data cleansing, and performing statistical analysis in Excel to ensure 100% compliance with Central Bank regulations.
- Automated data visualization, statistical analyses, and forecasting logic in Excel, improving compliance communication and reducing manual effort, resulting in a 50% faster average turnaround for Steering Committee (SteerCo) reports.
- Presented Power BI reports to senior stakeholders, facilitating data-driven decision making for regulatory initiatives.

Careys Plc

London, United Kingdom

Data Analyst Intern

June 2019 - Aug 2019

- Engineered a robust ETL pipeline using Python and SQL, integrating data from APIs into an Azure data warehouse.
- Produced analytical reports using Power BI, leveraging real-time weather data to assess safety risk at construction sites.

TECHNICAL SKILLS

Languages: Python (TensorFlow, PyTorch, Sci-Kit Learn), SQL, R, JavaScript, TypeScript, Java, HTML/CSS

Tools/Frameworks: Power BI, AWS, Azure, Terraform, Docker, Databricks, Tableau, Git/GitHub, Kubernetes

Other key competencies: ETL & Data Pipeline Engineering, Cloud Data Infrastructure, Data Modeling, Scalable Data Systems, LLM Integration, Agentic AI, Retrieval-Augmented Generation (RAG), In-Context Learning (ICL), Generative AI, Fine-Tuning LLMs