# YASH VISHE

La Jolla, CA | P: +91 8104652268 | yvishe@ucsd.edu | LinkedIn | website

## EDUCATION

**UNIVERSITY OF CALIFORNIA, SAN DIEGO** — La Jolla, CA
Masters in Data Science — Sep 2024 - Jun 2026

**UNIVERSITY OF MUMBAI** — New York, NY
B.E. in Computer Engineering (CGPA: 9.31/10) — Jul 2016 - Jul 2017

## TECHNICAL SKILLS

**Technical Skills:** Programming (Python, SQL, C++), Machine Learning, Deep Learning, Natural Language Processing, Computer Vision, Generative AI, Large Language Models (LLMs)
**Techniques:** Data Analysis, Data Engineering, Feature Engineering, Model Evaluation, Transfer Learning, Fine-tuning LLMs, Prompt Engineering, Statistical Analysis
**Tools & Software:** TensorFlow, Keras, Scikit-Learn, Transformers, LangChain, NumPy, Pandas, BERT, PySpark, Git, Google Cloud Platform, Gemini API, AWS, Apache

## WORK EXPERIENCE

**Data Analyst** — Mumbai, India
Media.net — Jul 2023 – Aug 2024
- Designed efficient ETL pipelines, migrating 46 alerts from Apache Superset to Apache Airflow to reduce the load on AWS Redshift, resulting in 38% faster data retrieval.
- Led the development of a bid adjustment algorithm in Python, strategically increasing bids with volume growth and reducing it otherwise, achieving a 13% increase in overall traffic volume.
- Created 8 dashboards on Tableau for revenue reporting and tracking CVRs, eliminating manual work and making the reporting process 78% faster.

**Software Development Intern** — New Delhi, India
Neumann Fornaxx — Nov 2022 – Feb 2023
- Led the design of an automated system to delete duplicate data, reducing database size by 62%.
- Automated the scraping of over 100,000 rows of data from Zomato and Swiggy using Python CRONs and BeautifulSoup, enabling detailed market analysis and data-driven decision making.

## PROJECTS

**Community Analysis using Graph Networks —** Python, SpaCy, NetworkX **(Github)** — Apr 2024
- Engineered an advanced Named Entity Recognition system that achieved a 60% improvement in character recognition, parsing interpersonal relationships from the Percy Jackson series across 10 books.
- Created interactive network graphs, identified groups of characters using community detection, and tracked the importance of all characters over the entire series of books.

**Document Query Agent using Multi-Modal RAG —** Langchain, Python, VertexAI, ChromaDB **(Github)** — Mar 2024
- Implemented a document query agent with a multi-modal self-reflective RAG for accurate data retrieval. Integrated Tavily web search agent to mitigate hallucination issues, boosting answering accuracy by 26%.
- Programmed the Gemini-1.5-pro model to generate summaries of unstructured data, including images and tables using LangGraph, improving retrieval accuracy by 23%.

**Book Recommendation Engine** — Python, NLTK, Sklearn (Github) — Jan 2023
- Scrapped GoodReads website to retrieve information on over 1 million books.
- Developed book recommendation system using cosine similarity, TFIdvectorizer and collaborative filtering to analyze book data, delivering five personalized recommendations aligned with user preferences based on book summaries.

## RESEARCH AND PUBLICATIONS

**Enhancing Brain Tumor Detection Using Multimodal Approach** — Coimbatore, India
IEEE · International Conference on Artificial Intelligence and Machine Learning Applications — May 2024
- Researched optimization functions to reduce false negatives in medical image analysis, achieving a recall improvement to 98.2%.
- Achieved an accuracy of 96.33% in brain tumor identification by training ResNet50 and ResUNet deep learning models for classification and segmentation using transfer learning.