PRIYANKA NIDADAVOLU

+1 (213) 760-5184 | pnidadavolu@ucsd.edu | LinkedIn | GitHub | Portfolio

SUMMARY

Data Science MS student at UC San Diego, available for **Summer 2026 Internship**. 2.5+ years at IBM building ML systems and data pipelines. Cut cloud costs by 40%, built high-throughput validation pipelines. Designed a Bayesian recommendation system with socially-aware modeling and rigorous experimental evaluation..

EDUCATION

University of California – San Diego　　　　　　　　　　　　　　　　　　　　Sept 2025 – Present
Master of Science in Data Science (GPA: 3.9/4.0)　　　　　　　　　　　　　　　　　　*San Diego, CA*
- **Coursework:** Recommender Systems, Machine Learning, Statistical Modeling, Big Data Analytics using Dask, Data Management for Data Science, Numerical Linear Algebra

K L University　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　2019 – 2023
Bachelor of Technology in Computer Science (GPA: 3.67/4.0)　　　　　　　　　　　　*Vijayawada, India*
- **Coursework:** Data Science, Database Management Systems, Data Warehousing and Mining, Data Visualization Techniques, Big Data Analysis

EXPERIENCE

UC San Diego, Information Technology Services　　　　　　　　　　　　　　　Oct 2025 – Present
Student Build & Release Engineer (Part-Time)　　　　　　　　　　　　　　　　　　*San Diego, CA*
- Maintain CI/CD infrastructure supporting engineering services, improving deployment reliability and automation across production systems.

IBM　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　Jan 2023 – Jul 2025
Software Developer (promoted from Intern, Jan 2023)　　　　　　　　　　　　　　　*Bangalore, India*
- Designed an LLM-based test routing system that analyzed code diffs to predict affected cloud platforms, eliminating redundant cross-platform test execution and reducing CI/CD costs by **40%**.
- Overhauled the API data validation framework, cutting validation time by **90%**.
- Analyzed multi-cloud usage logs using clustering and time-series decomposition to identify redundant storage patterns, producing optimization insights adopted in enterprise cost strategy.

Samsung PRISM　　　　　　　　　　　　　　　　　　　　　　　　　Jun 2021 – Mar 2022
Research Intern, NLP & Speech　　　　　　　　　　　　　　　　　　　　　　　*Remote*
- Engineered a phoneme dataset from **10,000+ contact names** using CMU Pronouncing Dictionary with recursive word breaking.
- Validated phoneme accuracy using Facebook's Wav2Vec2 speech recognition model and Levenshtein distance, improving recognition accuracy by **10%**.

PROJECTS

Socially-Aware Recommendation System (Production Deployed) | *Python, FastAPI, Docker, NetworkX*　　Dec 2025
- Developed a top-N recommendation system on the Epinions dataset (**664K+ reviews, 49K users**) based on the SBMC framework (McAuley et al., IJCAI 2017), combining collaborative filtering with a trust-based Markov Random Field (MRF) to model social influence.
- Engineered **35+ features**, benchmarked against Jaccard and Bayesian baselines, and improved ranking quality to **AUC 0.68** with gradient-boosted ensembles.
- Deployed production REST API (FastAPI) + interactive UI on Hugging Face Spaces with an A/B testing framework comparing social bayesian, bayesian, and jaccard models via top-K overlap analysis and per-user predict-probability scoring; Dockerized for reproducible deployment.(live demo).

Large-Scale NYC Taxi Data Pipeline & Analysis | *Python, PyArrow, Dask, AWS S3, Folium*　　Feb 2026
- Built a pipeline processing **57 GB of NYC taxi data (3.4B+ rows)** from AWS S3, pivoting trips into hourly aggregations across 443 Parquet files with schema normalization and mismatch detection.
- Ran PCA at scale using Dask, performed tail analysis (power-law alpha estimation), bootstrap eigenvector stability testing (Procrustes distance, subspace affinity), and visualized PC scores on Folium maps.

TECHNICAL SKILLS

Languages:	Python, SQL (PostgreSQL), R
Statistics:	Hypothesis Testing, A/B Testing, Causal Inference, Bayesian Methods, Regression
ML / Deep Learning:	Scikit-learn, XGBoost, PyTorch, Hugging Face Transformers, LangChain
Data & Cloud:	Spark, Dask, PyArrow, Parquet, Neo4j (Graph DB), Qdrant (Vector DB), AWS (S3, EC2), Docker, Git
Visualization:	Matplotlib, Seaborn, Plotly, Tableau

ACHIEVEMENTS & CERTIFICATIONS

- **Star of the Month**, IBM (Nov 2023): Delivered ML-driven optimization reducing cloud costs by 40%.
- **People's Choice Award**, IBM: Peer-nominated for data-driven decision-making project.
- **Codehers Coding Challenge** (2022): **Top 30 in India** in national competitive programming contest.
- **Certifications:** AWS Academy Cloud Foundations | Big Data (UCSD/Coursera) | IBM Data Science Foundations Level 1 & 2