Arpita Mangal

(628) 488-2112 ◆ amangal@ucdavis.edu ◆ San Francisco, CA ◆ LinkedIn ◆ Github ◆ Portfolio

SKILLS

Specialities: Experiment & Causal Inference | Customer Analytics| Marketing Analytics | Data Visualization
Technologies: Python | SQL | NoSQL | R | Excel | Git | Jira| Tableau | Redshift | Google Analytics | GCP | Looker
Certifications: AWS Cloud Practitioner | HackerRank SQL Advanced

WORK EXPERIENCE

Data Scientist, Practicum **Qualcomm** *San Francisco (Remote)* | Sep'22 - Jun'23

Led a 5 member team to analyse branding effectiveness in the compute market as part of the UC Davis MSBA.

- Scraped social media and e-commerce data from platforms like Twitter, Reddit, Amazon, and Target using **Beautiful Soup**, **Selenium** and **Tweepy**, stored the database in **MongoDB**, and pre-processed with **NLTK**.
- Refined **BERT** model from TensorFlow hub to comprehend **brand sentiment**, user needs & competitor landscape.
- Applied **K-means clustering**, to segment 900+ products with 26 feature points to drive product focus strategies.
- Designed a customer preference survey and analysed responses using **Structural Equation Modelling (SEM).**
- Formulated **social media influencer strategy** to identify target audience, relevant influencers, and marketing channels, evaluated influencer reach using UTM Parameters.

Associate Vice President, Strategy & Analytics **Kotak Securities** *Mumbai, IND* | Jul'21 - Jul'22

Increased CLV from $140 to $190 by deploying ML models to improve conversion, retention & enable personalization

- Achieved 30% lower acquisition cost by building a **marketing-mix model** that strategically optimised ad spend.
- Increased engagement to 13% through deep learning-based **customer segmentation** for targeted merchandising.
- Built a **propensity data mode**l to recommend stocks based on factors influencing clients trading behaviour.
- Examined **CRM** data- inbound conversations through **chat bot**, sales representative, **customer service request** and coordinated with cross-functional stakeholders to **enhance customer journey** boosting retention by 40%.
- Explored in-product customer behaviour using **pre-post event analysis**, **cohort analysis** & segmentation analysis.
- Devised **AB / multivariate tests**, defined engagement metrics for marketing campaign, advised optimal strategies.
- Architected **ETL pipelines** to maintain data on **cloud-premise** facilitating reporting, data modelling and analysis.
- Applied **NLTK** capabilities to evaluate impact of news & events on stock traded by estimating volume spikes.

Graduate Analyst **Barclays** *Pune, IND* | Jul'19 - Jul'21

*Constructed a **fraud detection model** to identify business banking mules and suggested prevention plans.*

- Trained an XG-Boost model with high accuracy & specificity using scikit-learn, reducing detection time by 26%.
- Leveraged variable importance to classify **personas**, distinguishing banking mules from legitimate customers.

Developed features for Barclays Mobile Banking (BMB - Mobile application for Internet Banking of Barclays).

- Coded new app. features, accessibility enhancements while maintaining legacy code. Codebase: ~ 20000 lines.
- Implemented app trackers to augment and empower the multi-touch attribution capabilities. (X-code, Jenkins)

EDUCATION

Master of Science, **Business Analytics**	**University of California, Davis**	*San Francisco, CA*	Jun. 2023
BTech & MTech, **Engineering**	**Indian Institute of Technology Kharagpur**	*Kharagpur, IND*	Jun. 2019

COURSEWORK

- Recommender Systems
- Machine Learning
- Natural Language Processing (NLP)
- Statistical Exploration & Reasoning
- Machine Learning on Cloud (GCP)
- Programming & Data Structures
- Big Data (CNN, GANS, SNA)
- Data Design and Mining
- Logistics and Supply Chain

PROJECTS

Customer Spend Behaviour and Social Network Analysis (SNA) over Lifetime (Venmo Data) | Project Link

- Utilised **MLlib** to predicted customer churn from category, recency, frequency of transactions and SNA metrics.
- Calculated Social Network metrics- clustering coefficient (triplets), page rank at each point in lifetime in **PySpark**.

Synthetic Image Generation From Text Prompts (Generative AI) | Project Link

- Fine-tuned a Stable diffusion model from **TensorFlow** hub to generate flower images from text prompts in **Keras**.
- Converted text prompts into embeddings using a **CLIP** model and then mapped into the **U-Net** via attention layer

ADDITIONAL QUALIFICATIONS

Awards: MSBA Fellowship, GSM Grant, 1st Position - Barclays Hackathon, Finalist -The Economist, Young Blood-GYWS

Volunteer (Leadership): Public Relations Officer - GYWS (Non-profit run by students of IIT Kharagpur to uplift rural areas & provide free education to underprivileged children), Governor - TAdS (Adventure Club, IIT Kharagpur)