Naman Vipul Chheda

San Jose, CA | +1 (408)-591-0652 | namanvipul.chheda@sjsu.edu | LinkedIn | Github

EDUCATION

Master's in Applied Data Intelligence, San Jose State University   **May 2027**
- **Coursework:** Data Warehouse and Pipeline, Data Visualization, Mathematical Methods for DI

B.Tech in Computer Science and Engineering (Data Science), DJSCE   **May 2025**
- **Coursework:** Computational Finance (Honors), Big Data Engineering, Machine Learning, Artificial Intelligence

EXPERIENCE

Inventure Growth and Securities Limited | *Data Analyst Intern*   **Mar 2024 - May 2024**
- Developed a Python and SQL pipeline that consolidated disparate market signals from spreadsheets and ad-hoc exports into a unified, large-scale data management system, enabling integration with agent frameworks for automated market monitoring.
- Analyzed over 50 live transactions and integrated SWOT assessments on major tech firms using large language models to translate raw data into actionable strategy notes that informed business roadmaps.
- Engineered recurring market reports from large datasets using SQL and Tableau, leveraging prompt engineering to automate summarization and contributing to a 12% improvement in client portfolio returns.

DJS Compute | *Competitive Programming Mentor*   **Sep 2023 - Dec 2023**
- Mentored over 25 students in competitive programming, enhancing their analytical skills and problem-solving abilities through structured guidance, clear documentation, and a demonstrated passion for organization.

PROJECT EXPERIENCE

Food2Recipe: Image-to-Recipe Data Pipeline   **Dec 2024 - May 2025**
- Built an end-to-end pipeline converting food photos into normalized recipe records with YOLOv8 + CNN features.
- Lifted ingredient detection +23% using FPN, custom heads, curated augmentations, and balanced sampling.
- Orchestrated data versioning, training, and inference with Airflow/DBT for reproducible runs.
- Validated outputs at 82% BLEU; documented failure modes and improved label quality.

Expert Patient Interaction Language Model   **May 2024 - Oct 2024**
- Fine-tuned Falcon-7B via Q-LoRA; improved response coherence and empathy +18% on internal prompts.
- Built a modular PyTorch training/eval harness for prompt tuning, ablations, and rapid iteration.
- Implemented BERTScore + calibration; maintained ~85% quality across cohorts and scenarios.
- Automated ingestion, preprocessing, and bias-mitigation checks; packaged for deployment.

Virtual Sales Assistant: A Chrome Extension   **Mar 2024 - Jun 2024**
- Shipped an extension delivering personalized recommendations to 200+ users.
- Designed a CF pipeline: behavior features 'candidate generation 'ranking (87% offline accuracy).
- Demonstrated 22% faster product search and 15% simulated conversion lift using interaction logs.
- Added telemetry, A/B hooks, and guarded rollout for safe, iterative improvements.

Uber Ride Data Using Big Data Technologies   **Jan 2024 - May 2024**
- Unified 10M+ ride logs into a curated Hadoop/Spark ETL for demand, surge, and routing analytics.
- Achieved 3× query speedups vs SQL-only; produced star-schema tables and materialized aggregates.
- Built Tableau dashboards for peak-hour planning and route allocation.
- Delivered insights on driver utilization and revenue distribution adopted by operations.

SKILLS

- **Languages:** Python, SQL, C/C++, Java, JavaScript, R
- **Data Engineering:** Airflow, DBT, Hadoop, Spark, ETL Pipelines, Snowflake
- **Data Analysis & Visualization:** Tableau, Power BI, Pandas, NumPy, Statistics, Data Modeling, Relational Databases
- **Cloud & Platforms:** AWS, Azure, Linux, Git, Docker
- **Machine Learning & AI:** Model Development, XGBoost, scikit-learn, Predictive Analytics, Data Preprocessing, Time Series Forecasting, GenAI, TensorFlow, Large Language Models, Agent Frameworks, Prompt Engineering
- **Core Concepts:** Automation, Reliability Analytics, Data Pipelines, Problem Solving, Cross-functional Collaboration, Passion For Organization

CERTIFICATIONS

- Amazon Web Services – Cloud Foundations — Cloud Architecting — Data Engineering

PUBLICATIONS

- N. Chheda, P. Saval. Virtual Sales Assistant: A Chrome Extension for E-Commerce Websites. IEEE ICAC2N 2024
- Gala, P., N. Chheda. Expert Patient Interaction Language Model. IEEE ICAC2N 2024