ARJUN ADUKATHIL

College Station, TX | (979) 575-9041 | arjun.adukathil@tamu.edu | LinkedIn | GitHub

EDUCATION

Texas A&M University, College Station, Texas	**Aug 2024 - May 2026**
Master of Science in Management Information Systems	GPA: 3.7/4.0
University of Mumbai, Mumbai, India	**Aug 2017 - May 2021**
Bachelor of Engineering in Computer Science	GPA: 3.76/4.0

PROFESSIONAL EXPERIENCE

LTIMindtree, Mumbai, India	**Jun 2021 - Jul 2024**

Data Engineer - Client: Onsemi

- Architected Snowflake dimensional models using SQL and DBT that reduced analyst report generation time by 25%, enabling faster decision-making for 50+ business stakeholders across finance and operations
- Reduced data ingestion latency by 30% through fault-tolerant Airflow pipelines processing 100GB+ daily datasets, ensuring 99.9% SLA adherence and enabling real-time business reporting
- Cut data validation effort by 80% from 20 to 4 hours weekly through Python data quality and governance framework, preventing production data incidents across 100+ datasets
- Saved 5+ engineering hours weekly through Bash-based monitoring tools that prevented 95% of storage related server failures, maintaining 99.5% data pipeline uptime

Data Engineer - Client: ServiceNow

- Boosted query performance by 40% through clustering and materialized views, reducing executive dashboard load times for 200+ daily business users
- Spearheaded migration of 20TB+ data from SAP HANA, SQL Server and PostgreSQL to Snowflake, reducing average query execution time by 60% for 50+ data analysts
- Deployed Apache Kafka + Spark pipeline processing 1 billion records/hour, reducing data latency from 24 hours to real-time for customer support ticket analysis
- Crafted 10+ KPI dashboards in Tableau processing 2M+ records, enabling executives to monitor sales performance, customer health, and product adoption trends

ACADEMIC PROJECTS AND PAPERS

**SafePlate	AWS Lambda, Amazon Bedrock, API Gateway, Flutter (Link)**	**Jan 2026 - Feb 2026**

- Achieved <5-second allergen classification serving 32M+ Americans with food allergies through serverless system using AWS Lambda, API Gateway, and Amazon Bedrock to analyze restaurant menus
- Engineered multi-modal AI pipeline with Claude Vision to classify menu dishes into 3 allergen safety categories, reducing dining decision time by 95%

**TerraLens	PyTorch, NumPy, Scikit-learn, NetworkX, Jupyter (Link)**	**Oct 2025 - Dec 2025**

- Achieved 56x faster similarity search by building a scalable image-embedding pipeline using CNN features and HNSW indexing, reducing query latency from 37 ms to 0.66 ms
- Improved clustering quality from ARI 0.57 to 0.62 by constructing a 27k node k-NN graph and applying Louvain community detection to uncover high-purity data partitions

**New York City Trips Analytics	GCP, GCS, BigQuery, DBT, Looker Studio (Link)**	**Jul 2025 - Aug 2025**

- Built scalable analytics platform using BigQuery, DBT, Medallion architecture and star-schema modelling that processed 10GB+ taxi data, enabling sub-second queries on trip pattern analysis
- Supported data-driven decision-making by building interactive Looker Studio dashboards visualizing trip volumes, fare distributions, and geographic trends

POS Cloud Database System| AWS, Galera, MongoDB, MariaDB (Link)	**Aug 2024 - Dec 2024**

- Deployed 3-node MariaDB Galera cluster on AWS EC2 achieving 25% faster read-write performance and zero downtime failover

TECHNICAL SKILLS & CERTIFICATIONS

- **Programming languages:** Python, SQL, Dart, JavaScript, Bash, C/C++, Java
- **Data Engineering:** Snowflake, DBT, Apache Airflow, Spark, SAP HANA, MongoDB, SQL Server, Hadoop, Oracle
- **Cloud & DevOps:** AWS, GCP, Azure, Docker, Kubernetes, Jenkins, Git, Bitbucket, Linux, Jira
- **Analytics:** Tableau, Power BI, Looker Studio, Data Science libraries (Numpy, Pandas, Matplotlib), Microsoft Excel
- **Certifications:** Snowflake Snowpro Core, Professional Scrum Master 1, IT Automation with Python by Google
- **Publications:** Housing Society Management System using Flutter and Firebase in IRJET, March 2021