Arya Lokesh Gowda

+1 (214)972-5529 | itsaryalg@gmail.com | LinkedIn | GitHub | Portfolio

EXPERIENCE

Data Analyst | *Nelnet* | Nebraska, USA
Aug 2024 – Present
- Developed financial models and analyzed KPIs, improving forecast accuracy by 5%
- Automated 10+ financial reports using SQL-based ETL workflows on Snowflake, decreasing work hours by 88%
- Streamlined ledger management by integrating financial reports across systems and migrating customer invoicing to a unified platform, reducing processing time by 70%

Data Processor | *Northeastern University* | Boston, USA
Oct 2023 – Present
- Optimized queries to extract and analyze applicant data for efficient processing to present findings in reports, enhancing transparency, communication, and drive admission decisions
- Extracted key metrics to drive insights, refining selection criteria and planning with agile methodologies

Data Engineer | *Infosys* | Bangalore, India
Oct 2021 – Aug 2023
- Architected a scalable data warehouse handling 10M+ daily records, with data governance protocols that reduced data quality issues by 40%, resulting in a **$150K (3%) reduction in operational costs**
- Orchestrated end-to-end ETL pipelines processing 500GB+ daily data across 12 source systems, reducing processing time from 6 hours to 45 minutes
- Led development of ML models for customer segmentation and predictive analytics, generating **$122K additional revenue**

Data Science Intern | *Entuple Solutions* | Bangalore, India
Jun 2020 – Jan 2021
- Deployed machine learning models for product optimization, resulting in **15% performance improvement** and **$50K annual cost savings** across manufacturing operations
- Visualized real time metrics for 10+ stakeholders and initiated automation processes for multiple instances

SKILLS

Programming Languages: Python, SQL, Java, R, MATLAB
Data: PostgreSQL, NoSQL, DBeaver, MySQL, Oracle SQL, MongoDB, Hadoop, MS Excel (VLOOKUP, Pivot Tables, Macros), SAS, Apache Spark, Hadoop, Hive, Jupyter
Business Intelligence: Tableau, Power BI, MS Access, Salesforce, Netezza
Cloud : Amazon Web Services (EC2, Redshift, Lambda, Glue, Athena, Sagemaker, IAM), Google Cloud Platform, Docker
Engineering: Airflow, JIRA, ControlM, GitHub, GitLab, Talend, Jenkins, Kafka, Snowflake

PROJECTS

Graid - *Chatbot to aid your grades*
Jan 2025-Present
- Pioneered an interactive AI-powered chatbot leveraging GenAI, Speech-to-Text LLMs (Groq), Pinecone, FastAPI, and RAG models to enhance student learning from video lectures
- Innovated a novel RAG framework for efficient contextual retrieval and personalized learning experience

Covid-19 Vaccine Sentiment Analysis | *VADER, Time Series Analysis*
Aug 2024-Nov 2024
- Engineered scalable pipeline processing 100K+ tweets daily, implementing custom NLP preprocessing
- Developed dashboards correlating sentiment trends with 50+ major vaccine events, identifying 3 key message types that increased vaccine acceptance by 12% in target demographics

Instagram Reach Analysis | *Passive Aggressive Regressor, Seaborn, Statsmodels*
Jun 2024-Aug 2024
- Built ensemble ML model analyzing 20+ engagement metrics across 10,000 posts, achieving **92.8% accuracy** and identifying optimal posting windows that **increased reach by 35%**
- Used A/B testing for content strategy, increasing engagement rate to 14% and 25% in follower acquisition

Fetal Health Classification System | *PySpark, scikit-learn, MLflow, Autoencoders*
Jan 2024-Apr 2024
- Designed a distributed ML pipeline processing 5000+ medical records daily, achieving **97% sensitivity** in critical case detection, potentially impacting 1000+ patient outcomes annually
- Engineered custom Autoencoder architecture for anomaly detection, reducing false positives by 40% and achieving 99% precision in high-risk pregnancy identification

Customer Data Stratification | *NumPy, Matplotlib, TensorFlow*
Sep 2023-Dec 2023
- Led 5-person team in developing RFM segmentation model analyzing 50,000+ customer profiles, **generating $70,000 additional revenue** through campaigns targeting 4 distinct customer segments
- Refined and deployed BERT-based NLP model processing 100,000+ customer interactions, achieving 96% accuracy in behavior prediction and **driving 13% sales increase**

EDUCATION

M.S in Data Analytics Engineering | *Northeastern University* | Boston, MA | GPA: 3.8/4
May 2025
Focusing: GenAI, MLops, Data Mining, Data Warehousing and Data Management in Analytics

B.E in Electronics and Communication | *Visvesvaraya Technological University* | *India*
Jul 2021