PORTFOLIO OVERVIEW DOCUMENT
AYUSH THONIPARAMBIL SASEENDRAN

Portfolio - Github

SUID :486477692 | Email : athonipa@syr.edu

Introduction

This document serves as a comprehensive reflection of my academic journey, professional experiences, and key projects developed throughout my career and graduate studies. It showcases the practical application of data science, engineering, and analytics skills through real-world scenarios—ranging from building scalable ETL pipelines to deploying cloud-based applications and developing interactive dashboards.

With a strong foundation in both theoretical concepts and hands-on implementation, this portfolio is structured to highlight my ability to solve complex problems, communicate data-driven insights effectively, and continuously grow in an ever-evolving technological landscape.

Table of Contents

1. Education

Degree	Institution	Duration
M.S. in Applied Data Science (GPA: 3.9/4)	Syracuse University, NY	Aug 2023 – May 2025
B.E. in Electronics and Telecommunication (GPA: 3.0/4)	University of Mumbai, India	Jul 2015 – May 2019

Master of Science in Applied Data Science

 ⚲ *Syracuse University, School of Information Studies | Aug 2023 – May 2025*
GPA: 3.9/4.0

Throughout the program, I've gained a comprehensive foundation in data-driven problem-solving, statistical modeling, and cloud-based data engineering. My coursework focused on real-world applications of data science techniques using tools such as Python, R, SQL, and cloud services like AWS and Azure.

Key Courses & Learnings:

- **Data Science**: Learned end-to-end data workflows including data cleaning, feature engineering, statistical analysis, and machine learning model development.

- **Database Management Systems (DBMS)**: Designed relational schemas, normalized databases, and optimized queries using SQL Server and PostgreSQL.

- **Applied Machine Learning**: Applied algorithms such as SVM, Random Forest, and Logistic Regression to classification, regression, and NLP tasks.

- **Data Warehouse**: Built star/snowflake schema models using Snowflake and dbt, integrating business intelligence principles for large-scale analysis.

- **Big Data Management**: Worked with Hadoop, Spark, and Kafka for distributed data processing, and designed ETL pipelines for real-time analytics.

- **Natural Language Processing (NLP)**: Built sentiment classifiers using NLTK and Scikit-learn, and evaluated text-based models using metrics like F1-score and ROC.

- **Cloud Management**: Deployed data pipelines and machine learning models on cloud platforms (AWS Lambda, Azure Functions), with experience in scalable architecture design.

These projects and hands-on assignments not only strengthened my analytical and technical capabilities but also helped me understand how to operationalize models and communicate insights effectively to stakeholders.

Bachelor of Engineering in Electronics and Telecommunication

⚲ *University of Mumbai, Vidyalankar Institute of Technology | Jul 2015 – May 2019*
GPA: 3.0/4.0

As an undergraduate, I built a strong foundation in mathematical modeling, electronics systems, and signal processing. Though the program was focused on hardware and communication technologies, I was introduced to programming, data analysis, and automation early on, which sparked my interest in data science.

Key Learnings:

- Mastered fundamentals of logic design, microcontrollers, and signal processing.

- Developed technical reports, managed lab data analysis, and programmed basic control systems using MATLAB and C.

- Completed a senior design project on drone-based mapping using LiDAR, integrating Raspberry Pi and ROS for environment scanning.

2. Professional Experience

Position	Company	Duration
Data Science Intern	BioMarin Pharmaceutical Inc	May 2024 – Nov 2024
Data Engineering Analyst	Accenture, Mumbai	Jan 2020 – Aug 2023

Data Science Intern – BioMarin Pharmaceutical Inc.

⚲ *Novato, CA | May 2024 – Nov 2024*

At BioMarin, I led the migration of structured risk register data from traditional Excel spreadsheets into a fully normalized SQL Server database using SSMS. This transition streamlined data retrieval and improved data governance, resulting in a 40% boost in operational efficiency.

I was responsible for both logical and physical data modeling—defining entity relationships, building ER diagrams, and implementing constraints to ensure referential integrity. Additionally, I developed comprehensive documentation and audit triggers to monitor changes across critical tables, significantly enhancing data traceability and security.

To improve usability, I built an interactive PowerApps application that allowed non-technical users to log in, browse, update, and manage data through form-driven interfaces and galleries. This resulted in a 35% increase in user satisfaction, as measured by internal surveys. I also collaborated with data governance teams to define user role-level access controls and ensured compliance with internal data policies.

Data Engineering Analyst – Accenture

⚲ *Mumbai, India | Jan 2020 – Aug 2023*

At Accenture, I specialized in building robust and scalable ETL pipelines using Informatica PowerCenter, managing large volumes of structured and semi-structured data for global enterprise clients. My role involved orchestrating batch jobs using shell scripts, which enhanced data processing speeds by 30% and ensured on-time data availability for reporting.

I utilized Informatica's parameter files to design reusable workflows, enabling the configuration of mapping variables, connection strings, and control flags. This allowed seamless integration with FTP servers and relational databases.

To optimize performance in Oracle 12c environments, I designed and implemented strategic indexing, materialized views, and partitioning schemes, which reduced query response times and improved dashboard refresh rates.

As part of several migration and modernization projects, I also coordinated with cross-functional teams, including SMEs and business analysts, to ensure that data models and pipelines aligned with evolving business needs. Furthermore, I led code reviews and implemented automation scripts that reduced manual dependencies and improved data pipeline reliability.

3. Key Projects

Project Name	Description
⏍ Online Recipe-Sharing Platform	Developed a multi-cloud recipe-sharing platform with advanced search, ratings, and user-authored content. Used PowerApps for UI, SQL for data storage, and integrated cloud services (AWS, Azure, GCP).
US FEC Individual Contribution Data Analysis	Analyzed FEC 2007-08 data using Snowflake and dbt. Created contribution and expenditure fact tables, and designed Power BI dashboards to visualize donor and geographic trends.
🌍 Finding Frontiers Travel Management System	Built a SQL-backed travel management app using PowerApps. Created conceptual/logical ERDs and scripts. Developed BI dashboards for booking, reviews, and management.
💡 Household Energy Consumption Forecasting	Forecasted energy demand using regression models in R. Integrated weather and usage data, developed Shiny app dashboards, and recommended policy strategies for efficiency.

4. Technical Skills

- Programming: Python, R, SQL, Scala, Shell Scripting
- Cloud/DB: Snowflake, AWS, GCP, Oracle, PostgreSQL
- ETL: Informatica, Airflow, dbt, SSMS
- Visualization: Power BI, Tableau, Shiny
- Tools: Git, Bitbucket, Jira, PowerApps

5. Certifications

• AWS Certified Solutions Architect – Associate (In Progress)

This certification validates expertise in designing scalable, fault-tolerant, and cost-optimized architectures on Amazon Web Services (AWS). Through this learning path, I'm gaining hands-on experience with core AWS services like EC2, S3, RDS, Lambda, and VPC. It's strengthening my ability to architect cloud-native solutions, manage secure deployments, and make strategic decisions for cloud-based systems—skills that are essential in today's data-driven organizations.

• Microsoft Certified: Azure Fundamentals

This certification demonstrates a foundational understanding of Microsoft Azure's core services, governance, and compliance features. It covers key concepts such as cloud computing models, Azure pricing and support, and cloud security principles. It has helped me build a solid baseline for working in hybrid and multi-cloud environments, and enabled me to navigate Azure Data Factory, Azure SQL, and Azure Functions confidently within academic and personal projects.

6. Conclusion

This portfolio highlights my journey as a data professional, showcasing a blend of academic excellence, hands-on project experience, and practical industry exposure. Through my master's program at Syracuse University and professional roles at BioMarin and Accenture, I have developed a well-rounded skill set in data engineering, statistical analysis, machine learning, and cloud-based application development.

From transforming raw data into actionable insights to building user-centric applications, each project reflects my ability to combine technical rigor with strategic thinking. Whether working on federal contribution analysis using Snowflake and Power BI, or developing intuitive platforms like a recipe-sharing app and travel management system, I've consistently delivered impactful, scalable solutions.

Looking ahead, I am committed to continuous learning, contributing to ethical and innovative applications of data science, and driving data-informed decisions that create real-world value. I'm excited about the future and eager to take on opportunities where I can grow, lead, and make a meaningful difference in the world through data.