

# HYMA ROSHINI GOMPA

Baltimore, MD | hymarog1@umbc.edu | 940-843-6984 | [www.linkedin.com/in/gompa-hyma](https://www.linkedin.com/in/gompa-hyma) | [github.com/Roshinigompa](https://github.com/Roshinigompa)

## EDUCATION

**Masters in Data Science** | University of Maryland Baltimore County | GPA:3.95/4.0

May 2025

**Bachelors in Computer Science** | Lovely Professional University | GPA:3.9/4.0

May 2022

## SKILLS

**Programming:** SQL, Python (NumPy, Pandas, Scikit Learn), R (ggplot2, dplyr), C++ | **Web Technologies:** HTML, CSS, Java Script | **Database Management:** MySQL, TSQL, PostgreSQL, Snowflake | **Statistics & ML:** Hypothesis testing, A/B testing, Causal Inference, Regression | **Tools:** Tableau, Power BI, MS Excel, MATLAB, Salesforce Apex, LWC | **DevOps & CI/CD:** Docker, Jenkins, Kubernetes, Git | **Collaboration Tools:** JIRA, Confluence, GitHub.

## RELEVANT EXPERIENCE

**Business Analyst Intern / Bizinc**

Oct 2024 – Dec 2024

- **Authored 15+ technical documentation assets**, enhancing cross-team onboarding by 30% through clearer project specifications and reducing misunderstanding across teams.
- **Led 10+ requirement-gathering sessions** with product and business teams, ensuring project alignment with client needs, contributing to a **25% boost in user engagement**.
- **Designed a Tableau and SQL-based performance dashboard** that tracked **50+ KPIs**, reducing manual reporting time by 40% and improving decision-making efficiency by **40%**.
- **Built an automated data pipeline in SQL, Python, and JavaScript**, reducing weekly reporting time by **40%** and improving data processing efficiency for the management team by **30%**.

**Data Analyst / Josh Technology Group**

Jun 2022 – Jul 2023

- **Improved product review classification accuracy by 20%** through a logistic regression model with **85% accuracy**, using NLP preprocessing on **10,000+ reviews** for clients **Medstar, Bwell, and Walgreens**.
- **Developed 5 data marts by transforming complex SQL queries**, achieving a **95% data integrity** rate and enhancing data readiness for analysis within the data science team.
- **Optimized 20+ Tableau dashboards**, reducing report generation time by **35%** and increasing report accuracy by **25%** by identifying key data insights and streamlining dashboard workflows.
- **Executed high-volume data analysis for datasets of up to 100,000 rows**, utilizing advanced Excel techniques to meet operational needs with **98% accuracy**, maintaining data handling efficiency on tight deadlines.

**Data Analyst Intern / Josh Technology Group**

- Extracted and cleaned customer transaction data **using SQL and performed exploratory data analysis** to identify key features for **clustering algorithms (K-means, DBSCAN)**, improving segmentation for customer profiling.
- Developed client-facing **Tableau dashboards and financial reports**, leveraging **Excel and PowerPoint** to enhance report accuracy and clarity, resulting in a **20% improvement** in financial reporting efficiency.
- Collaborated with business stakeholders and cross-functional teams, using **NLP-based** insights to respond to data requests and translating over **30% of data requests into actionable reports**, increasing operational **efficiency by 20%**.
- Maintained and managed **RDBMS systems**, ensuring **data integrity**, while supporting the finance team's data needs.

## RESEARCH PROJECTS

**Survival Analysis for Breast Cancer Prediction**  
**Research Under Prof. Zeynep Kacar, UMBC**

Aug 2024 – Dec 2024

- Developing survival models (Kaplan-Meier, Cox PH, Random Survival Forests) for breast cancer prognosis using Python and MATLAB, improving predictive **accuracy by 15%** on the METABRIC dataset. Identifying key prognostic factors to support clinical decisions.

## ACADEMIC PROJECTS

**Anomaly detection using Robust Graphical Lasso**

Oct 2024 – Dec 2024

- Designed a framework using Robust Graphical Lasso to handle outliers in high-dimensional datasets by combining RPCA with sparse covariance estimation. Leveraged **ADMM optimization** to achieve robust performance, validated on synthetic and real-world datasets.

**End-to-End Machine Learning Workflow with MLflow and AWS CICD**

May 2024 – Jun 2024

- Integrated MLflow for experiment tracking and deployed containerized models using Docker and AWS EC2. Automated CI/CD with GitHub Actions, reducing deployment time by **40%** and achieving **99.9% uptime**.

**Real-Time Voting Analysis System**

Apr 2024 – May 2024

- Built a scalable, real-time voting system with **Python, Kafka, and Spark Streaming**, capable of processing **500,000+ records/minute**. Deployed a live dashboard with Streamlit, providing insights in under 1-second processing time.