

Sanjaikumar Balasubramaniyan

Data Science Professional

My Portfolio | (812) 671-6737 | bsan@iu.edu | [LinkedIn](#) | [GitHub](#) | [LeetCode](#)

PROFESSIONAL EXPERIENCE

Research Data Scientist

Indiana University, Bloomington, IN

Jan. 2025 – Present

- Investigated funding patterns across **157,000+** BIPOC nonprofit grants to uncover equity gaps.
- Mapped funding networks and classified recipients by demographic focus using **advanced retrieval techniques**.
- Identified systemic disparities in grant allocations through **network analysis** and **statistical modelling**.
- Presented actionable funding strategy recommendations to university stakeholders to promote equitable distribution.

Data Scientist Intern

Systech Solutions Inc., Chennai, India

Apr. 2024 - July 2024

- Tasked with enhancing real-time analytics and customer segmentation for a mid-size analytics firm.
- Built dynamic **Power BI dashboards** and automated **ETL pipelines** with **Python** and **SQL**.
- Deployed **machine learning models** that improved customer segmentation accuracy by **25%**.
- Reduced data processing time by **35%** and accelerated decision-making speed by **40%** across departments.

Data Analyst

8Queens Software Technologies Pvt Ltd., Chennai, India

May 2022 – Jan. 2023

- Supported the sales department's need for real-time performance tracking and health risk prediction.
- Developed interactive **Tableau dashboards** and built **machine learning models** in **Python** and **R**.
- Enhanced diabetes risk prediction accuracy by **15%**, improving health analytics insights for client outreach.
- Enabled strategic sales initiatives by integrating **predictive analytics** into daily operations.

TECHNICAL SKILLS

Programming: Python, R, SQL, C/C++

Data Visualization: Tableau, Matplotlib, Seaborn, Power BI, Gephi

Big Data & Cloud: AWS (S3, EC2, QuickSight), GCP (BigQuery, DataFlow), Azure

ML & Algorithms: Regression, Classification, Clustering, Model Evaluation, Hyperparameter Tuning

Database Management: MySQL, PostgreSQL, MongoDB, SQLite, Data Warehousing (Snowflake, Redshift), ETL Processes

PROJECTS

Intelligent Resume Screening System | **NLP, BERT, Streamlit, Python, AWS Lambda**

- Developed an **NLP-based AI screening tool** using **BERT embeddings** to classify and rank resumes based on job descriptions, achieving **93% matching accuracy** across diverse candidate profiles.
- Deployed a real-time web app with **Streamlit** and **AWS Lambda**, reducing manual HR screening time by **70%** and supporting dynamic role-based filtering for 10,000+ resumes.

AI Agent for Operating Backend System | **LLM, Flask, SQLite, Google Gemini-pro**

- Engineered an AI agent leveraging **Gemini-pro LLM** to automate backend workflows, improving operational efficiency by **50%** and reducing manual interventions by **80%** at **95% accuracy**.
- Developed and deployed a **Flask API** integrated with **SQLite**, reducing system response times by **40%**, supporting **1,000+ concurrent queries**, and boosting backend performance by **30%**.

Advanced ANPR for Vehicle Management | **YOLO, OCR, Python, SQL**

Published in Adalya International Journal, [DOI: 10.37896/aj13.3/001](#)

- Designed a **YOLO-based object detection model** combined with **OCR** for real-time license plate recognition, achieving **96% detection accuracy** and enhancing blurred image clarity by **85%** through advanced preprocessing.
- Built a scalable end-to-end vehicle management system with automated tracking, parking management, and real-time entry/exit logging, attaining **92% system-wide accuracy**.

EDUCATION

Master of Science in Data Science

Indiana University, Bloomington, IN

Aug. 2024 - May 2026

- Coursework: Statistics, Applied Algorithms, Advanced Database Concepts, Machine Learning, LLMs, Data Mining