

# Sanjaikumar Balasubramaniyan

Data Scientist

[www.sanjai-11.github.io](https://www.sanjai-11.github.io) | [bsan@iu.edu](mailto:bsan@iu.edu) | [linkedin.com/in/sanjaikumar-b-309066202/](https://linkedin.com/in/sanjaikumar-b-309066202/) | [github.com/sanjai-11](https://github.com/sanjai-11)

## EDUCATION

### Master of Science in Data Science

Indiana University, Bloomington, IN | May 2026

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

### B. Tech in Artificial Intelligence & Data Science

SRM Valliammai Engineering College, Chennai, India | April 2024

- Relevant Coursework: AI, Business Intelligence, Natural Language Processing, Numerical Linear Algebra, Data Structures, Computational Linguistics, Probability & Statistics, Image Processing

## PROFESSIONAL EXPERIENCE

### Data Scientist Intern

Systech Solutions Pvt Ltd., Chennai, India | Apr 2024 - July 2024

- Built **real-time Power BI dashboards**, enhancing decision-making speed and visualization quality by **40%**.
- Automated data workflows with optimized **ETL pipelines**, cutting data processing time by **35%**.
- Deployed machine learning models, boosting customer segmentation accuracy by **25% for predictive insights**.
- Provided scalable data solutions, improving cross-team operational efficiency by **30%**.

### Data Analyst Intern

8Queens Software Technologies Pvt Ltd., Chennai, India | May 2022 – Jan 2023

- Streamlined real-time data processing by **transforming** datasets, improving accessibility and accuracy.
- Created interactive Tableau dashboards for sales, **driving a 15% improvement** in performance insights.
- Deployed supervised ML model for **diabetes risk prediction**, boosting predictive accuracy in **health analytics**.
- Delivered actionable insights through data-driven solutions, supporting strategic decision-making processes.

## TECHNICAL SKILLS

**Programming:** Python, R, SQL, C/C++, Scala, JavaScript

**ML & DL Algorithms:** Regression, Classification, Clustering, XGboost, CNN, RNN, LSTM, MLPs, GRU, GANs

**Technical Tools:** PyTorch, TensorFlow, Tableau, PowerBI, Pandas, NumPy, SAS, Stata, Scikit Learn, Hugging Face, Computer Vision

**Cloud & Databases:** PostgreSQL, MSSQL, MongoDB, Neo4j, Hadoop, Hive, AWS (S3, EC2, SageMaker, Glue, QuickSight), GCP (BigQuery, DataFlow, AutoML, VCP)

**Developer Tools:** VS Code, Docker, Postman, GitLab, GitHub, Bitbucket, JIRA

**Certifications:** Google Professional Data Analytics, Python for Data Science (NPTEL, Elite batch), ML by Harvard

## PROJECTS

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

- Built an AI agent using Gemini-pro LLM, automating workflows with 50% improved efficiency and 80% fewer manual tasks at **95% accuracy**.
- Developed a Flask API with SQLite, cutting response times by 40%, supporting **1,000 queries**, and boosting backend efficiency by **30%**.

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

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

- Built a **YOLO**-based model with **OCR** for license plate detection, achieving **96% accuracy** and enhancing blurry images by **85%** through advanced preprocessing.
- Developed a scalable system with **automated vehicle tracking**, entry/exit logging, and **92%** overall accuracy.

### AI-Powered Banknote Identification | Computer Vision, Deep Learning, Python, gTTS

- Developed a deep learning model with **96% accuracy**, reducing recognition errors by 18%, and **integrated gTTS** (Google Text to Speech) for real-time voice feedback, improving accessibility by **40%**.
- Built an inclusive financial tool for visually impaired users, benefiting **100+ testers**, and **presented at RTACT 2024 for its innovation and social impact**.