

Sarah Joby David

Phone: +91 9981115451 | sarahjobydavid@gmail.com | linkedin.com/in/sarahjobydavid | github.com/SarahJobyDavid

EDUCATION

Vellore Institute of Technology (2022 – 2026)
Bachelor of Technology in Computer Science and Engineering (Specialization in Cloud Computing and Automation) |
CGPA: 8.9

SKILLS

- **Programming** - Java, Python, SQL
- **ML & Data** - TensorFlow, Pandas, NumPy, Scikit-learn; EDA, CNNs, Data Preprocessing, Feature Engineering, Model Evaluation
- **Tools** - AWS, Power BI, MySQL, TensorFlow

PROJECTS

Potato Disease Classification (Deep Learning, Mobile & Cloud) (July–Aug 2025)

- Built end-to-end ML pipeline: Preprocessed/augmented image data; trained CNN with TensorFlow for early/late blight detection (agricultural crop loss mitigation).
- Deployed scalable model to mobile app via AWS for real-time farmer diagnostics; optimized for robustness & accuracy.

Student Performance Analysis (EDA & Supervised ML) (Oct 2025)

- Analyzed 1K+ records with Pandas/SQL: Engineered features linking demographics to pass rates (>60% avg.); 57% female vs. 43% male passers.
- Visualized insights (bar/KDE plots) on parental education impact (15% score boost) & test prep (25% pass increase); automated pipelines for iteration.

Data-Driven Hospitality Management System (June 2025)

- Developed Power BI pipelines: Ingested/transformed data with Power Query/DAX; improved occupancy forecasting by 15% via simulated scenarios.
- Built interactive dashboards (RevPAR/ADR KPIs) for cross-team insights & decision-making.

Customer Segmentation Analysis (Unsupervised ML) (Nov 2025)

- Applied KMeans clustering on mall customer dataset (200+ records) using Sklearn/Python: Standardized features (Annual Income, Spending Score); identified optimal 5 clusters via elbow/silhouette scores.
- Visualized segments with scatter plots; derived insights on profiles (e.g., high-income high-spenders at 82 avg. score) for targeted marketing strategies.
- Automated preprocessing, evaluation, and summary pipelines for scalable business segmentation.

Research Work

AI-Driven Resilience & Privacy in Cloud IoTBook Chapter (Scopus-Indexed, Scrivener/Wiley) – Aug 2025

- Co-authored on ML for threat detection/privacy in IoT; co-presented at International Scientific Research Congress (Turkey, Virtual – Dec 2024).
- Explored energy modeling & resilience strategies from research papers.

Activities & Certifications:

- **Extracurriculars:** Cloud Zone Club Designing Team- Core Member
- **Certifications:** Data Science with Python (iamneo) - April 2025
 - Oracle Certified Professional: Java SE 11 Developer