# Swetha Gayatri Kandikonda

Senior Data Analyst
+917993539574 | swethagayathri1964@gmail.com | LinkedIn: linkedin.com/in/swetha-gayatri-kandikonda-26286212a

## SUMMARY

Data Analyst with 3+ years of experience in Azure, Databricks, Py Spark, SQL, and Power BI. Skilled in building scalable data pipelines, developing dashboards, and applying ML techniques including XGBoost, deep learning, and NLP. Proven track record of reducing manual effort by 85%, improving data accuracy by 20%, and achieving 94% ROC AUC in sentiment analysis.

## EXPERIENCE

**Senior Data Analyst** Dec 2020 - Jul 2024

LTIMindtree

- Processed and transformed 100GB+ of raw data daily by orchestrating distributed ETL pipelines with PySpark on Azure Databricks, enabling timely downstream analytics.
- Improved reliability of production pipelines by integrating PySpark jobs with Databricks Workflows for scheduling and monitoring, reducing pipeline failures by 30% and improving SLA compliance.
- Cut debugging time by 50% by implementing logging and exception handling in PySpark-based ETL pipelines, ensuring faster issue resolution.
- Elevated query performance by 15% by optimizing SQL joins, window functions, and indexing strategies within Azure Databricks.
- Accelerated executive decision-making by 30%, developing interactive Power BI dashboards with drill-through KPIs, automating manual reporting.
- Increased stakeholder adoption of analytics tools by 25% by designing self-service Power BI dashboards, integrating row-level security for reliability.
- Boosted actionable insights by 20% by applying statistical functions, pivot tables, and lookup formulas in Excel to analyze business datasets.
- Improved customer response time by 35% by implementing event-triggered workflows in Azure Logic Apps, automatically processing support tickets as soon as they were raised.
- Eliminated 80% of manual intervention by automating cross-application integrations through Azure Logic Apps improving process reliability.
- Amplified customer efficiency by 30% by processing 100+ daily HTTP requests with Azure Function Apps, enabling scalable, serverless operations.
- Reduced cloud infrastructure deployment time by 40% by automating resource provisioning and configuration using Azure PowerShell scripts.
- Expanded Azure community reach to 680K+ users by contributing 770+ technical solutions on Stack Overflow as an Azure Collective member.

## PROJECTS

**Loan Defaulter Prediction using Machine Learning** Jan 2025 - Apr 2025

- Deployed a loan defaulter prediction model on imbalanced dataset, boosting model reliability by removing 40% features through feature engineering.
- Enhanced dataset quality by conducting exploratory data analysis, imputing missing values with KNN, and removing 0.1% outliers using Scikit-learn.
- Built end-to-end ML workflows incorporating categorical encoding, feature scaling, and an XGBoost model achieving a Recall score of 0.83.

**Amazon Product Recommendation Engine** Jan 2025 - Apr 2025

- Built a hybrid recommendation engine for video games, combining SVD-based collaborative filtering with content-based methods to improve results.
- Enhanced recommendation relevance by generating game descriptions via prompting, embedding with Sentence Transformers, and applying cosine similarity, boosting content-based recall by 15%.
- Evaluated model effectiveness using Precision@k, achieving ~20% precision, demonstrating measurable potential to increase user engagement.

**Interactive Tableau Dashboard: Analyzing Urban vs Rural Education Disparities** Aug 2024 - Dec 2024

- Uncovered key factors influencing education in urban vs. rural areas by developing interactive Tableau dashboards highlighting disparities.
- Identified correlations between school facilities (washrooms) and student retention rates, enabling data-driven insights into education accessibility.
- Enhanced stakeholder understanding of education gaps by visualizing multi-dimensional datasets in Tableau, driving actionable discussions on policy and resource allocation.

## SKILLS

**Programming** Python, C#, Microsoft SQL Server, KQL, ADO.NET, Data Structures, Algorithms, Problem solving, Py Spark

**Machine Learning** Machine Learning, Regression, Classification, Data Cleaning, Data Preprocessing, Data Transformation, Feature Engineering, Data Visualization, Cross Validation, Deep Learning, Neural Networks, Natural Language Processing, Sentiment Analysis, Generative AI, LLMs

**Libraries** Pandas, Matplotlib, Seaborn, Scikit-Learn, xgboost, nltk, Pytorch, Optuna, transformers, ollama, Numpy, Scipy, Spacy

**Cloud Computing** Microsoft Azure, Databricks, Logic Apps, Powershell, Azure CLI, Azure Runbooks, Azure Service Bus, Data Lake, Azure Synapse

**Others** Microsoft Power BI, SQL Server Management System, MS Excel, Microsoft Power Automate, ASP.NET MVC, ASP.NET Web API, Tableau

## EDUCATION

**Master of Science - Data Science** Aug 2024 - Apr 2026

Vellore Institute of Technology

GPA : 9.16

**Bachelor of Technology - Computer Science** Jul 2016 - Aug 2020

CVR College of Engineering

GPA : 8.68