SARAN RAJ KRISHNAPILLAI

+91 8870245241 - saranrajkrishna98@gmail.com – LinkedIn-Saranraj-K – github-Saranraj-K – Leetcode-Saranraj-k/ – Hackerrank-Saranraj-K

OBJECTIVE

Passionate Data Science Professional with a strong background in leveraging data to drive business decisions. With experience in collaborating with global leaders, I specialize in designing scalable data engineering solutions, building machine learning models, and delivering actionable insights through analytics and visualization.

EDUCATION

Bachelor of Engineering (ECE) - Jeppiaar Engineering College, Chennai (2016 to 2020) - CGPA: 7.7/10.0

SKILLS

Programming:	Python, SQL, OOPS, Data Structures & Algorithms
Data Science:	Machine Learning (Regression, Classification & Clustering), Deep Learning, Natural Language Processing, Transformers, Statistics
Libraries & Frameworks:	Scikit-learn, TensorFlow, Keras, XGBoost, MLlib, NLTK, PySpark, Flask, Pandas, Seaborn, Scipy, MLflow
Cloud Platform:	AWS - EC2, S3, Lambda, Glue, SageMaker & Azure – Databricks, Azure ML
Big Data:	Apache Spark, Hadoop, Hive, Kafka, Airflow.
Visualization Tool:	Power BI

EXPERIENCE

Associate (Skills - Python, PySpark, Azure Databricks, Big Data Analytics, Machine learning, Power BI) Oct 2024 - Present

PwC *Bangalore, KA*

- As part of PwC's Risk Analytics team, collaborating with a global retail leader to manage and audit big data from multiple countries, leveraging advanced analytics and risk modelling to provide actionable insights for mitigating risks.
- Implementing Medallion Architecture on data using Azure Databricks, Python, PySpark, and SQL for efficient data engineering.
- Developing and automating pipeline workflows in Databricks to streamline the data engineering process.
- Performing in-depth data analytics and creating impactful visualizations using Power BI to derive actionable insights.
- Designed and deployed machine learning models on Azure Databricks while leveraging Azure DevOps for version control and continuous integration to streamline data-driven solutions.
- Participated in IAM 2025 in Mexico alongside client auditors from all regions across the globe, representing the analytics team and presenting AI/ML and analytics solutions.

IT Analyst (skills - Python, Data Analytics, Machine learning, SQL, Power BI) Apr 2021 - Sep 2024

NTT DATA *Bangalore, KA*

- Worked with the team on data migration projects for Fidelity Investments, to transition millions of records from a legacy Oracle platform to a cloud infrastructure. This involves meticulous configuration transfer and application migration. Leveraging client insights to enhance user experience for multinational industry leaders.
- Cleaned and pre-processed large datasets, improving data quality and extracting insights to drive informed decision-making.
- Developed responsive Python web applications using Flask, reducing manual efforts and streamlining workflows.
- Built predictive models using machine learning to forecast configuration app failures, increasing system reliability.
- Created complex SQL queries for data research, analysis, and troubleshooting, enhancing data-driven decision making. Utilized Power BI and Excel for analytics visualization, enabling clear and actionable insights.
- Collaborated with cross-functional teams to deliver actionable insights, supporting business improvements with technical expertise.

PROJECTS

Data Comparison Tool (Nov 2023 to Jan 2024): Developed a Flask-based tool using NLP, Python, SQL to extract and compare data from various document formats, which helps in reducing validation time for the testing team.

Error Log Analyzer (Mar 2023 to Aug 2024): A machine learning application to classify error logs generated during data load in migration, using classification algorithms to identify patterns and anomalies for efficient troubleshooting.

Banking Credit Risk Modelling (Oct 2024 to Mar 2025): YesBank internal project at PwC - Developed a machine learning model to predict potential loan defaulters using classification algorithms on banking data. Performed threshold calibration to balance precision and recall, enhancing decision-making for risk mitigation. Improved model interpretability and business alignment by tuning probability thresholds instead of relying solely on default cutoffs.

COURSEWORK

*Machine Learning with Python *NLP Specialization *Deep learning Specialization *MLOPS with AWS *Big Data Engineering with Azure Databricks