M. Saikumar

Email Id : saikumar.raj007@gmail.com
LinkedIn: https://www.linkedin.com/in/m-sai-kumar-b84080148/
Phone : +91 - 9392647303

Profile Summary:

❖ Results-Oriented and Highly skilled Data Professional with 8 years of Experience in Data Science, Senior Data Analysis, and Big Data Engineering. Possessing Expertise in AWS and GCP platforms, PySpark, and Big Data Technologies, I have successfully delivered data-driven solutions and actionable insights to drive business growth and operational efficiency. Adept at leveraging advanced analytics techniques and machine learning algorithms to uncover patterns, trends, and valuable insights from complex datasets.

❖ Proven history of Designing and Implementing scalable Data Architectures, optimizing Data pipelines, and ensuring Data Accuracy and quality. Strong leadership and communication skills, with a collaborative approach to cross-functional teamwork and stakeholder engagement. Committed to continuous learning and staying updated with emerging technologies and industry best practices. Seeking a challenging role as a Data Scientist to leverage my expertise in Data Science, Analytics, and Big Data Engineering to contribute to the success of forward-thinking Organizations.

WORK EXPERIENCE:

❖ **Manager, ICICI Group of Companies [Apr 2022 - Feb 2023]**

- Data Management and Architecture: Oversaw the end-to-end data management process, including data ingestion, storage, processing, and transformation using PySpark and Hadoop Big Data technologies.

- PySpark and Hadoop Expertise: Demonstrated advanced proficiency in PySpark and Hadoop ecosystem tools such as HDFS, MapReduce, Hive, or Spark SQL, enabling efficient data processing and analysis at scale.

- Data Pipeline Development: Designed and implemented data pipelines using PySpark and Hadoop, ensuring smooth data flow, ETL operations, and data integration across multiple sources.

- Big Data Processing: Utilized distributed computing techniques and parallel processing to handle large volumes of structured and unstructured data, optimizing performance and scalability.

- Data Quality Assurance: Developed and implemented data quality checks and validation processes to ensure data accuracy, consistency, and integrity throughout the data management lifecycle.

⚑ As a Data Manager in ICICI Group of Companies, I successfully led a team in effectively managing client databases, ensuring data integrity, and supporting strategic decision-making. Through the

implementation of robust data management policies and practices, I optimized data quality and efficiency, improving overall data governance. My expertise in data analysis and data systems contributed to enhancing data security, privacy, and performance. Additionally, I played a key role in fostering collaboration with cross-functional teams and external partners to drive data-driven insights and achieve business objectives.

❖ SENIOR DATA ANALYST, SKI CAPITAL Co. Ltd [Apr 2021 - Nov 2022]

- Data Analysis and Insights: Conducted comprehensive data analysis using both AWS and GCP platforms, leveraging tools like Amazon Redshift, Amazon Athena, Google Big Query, or Google Data Studio to extract valuable insights and support data-driven decision-making.

- Data Modeling and Optimization: Developed and implemented efficient data models and optimized data storage and retrieval processes, ensuring scalability, performance, and cost-effectiveness.

- Machine Learning and Predictive Analytics: Utilized machine learning techniques and predictive analytics models to uncover patterns, trends, and future predictions, enabling proactive decision-making and strategic planning.

- Data Integration and ETL: Designed and implemented data integration pipelines, performing ETL (Extract, Transform, Load) operations to gather, transform, and load data from multiple sources, ensuring data accuracy and consistency.

- Cloud Platform Expertise: Demonstrated proficiency in both AWS and GCP platforms, including services such as Amazon S3, EC2, Redshift, Google Cloud Storage, Compute Engine, BigQuery, enabling seamless data processing, storage, and analysis.

As a Senior Data Analyst in Ski Retail Capital Co. Ltd, I successfully leveraged my expertise in exploratory data analysis, classification modeling, and data visualization to drive data-driven decision-making and support strategic initiatives. Through my ability to analyze and interpret complex datasets, I provided valuable insights that contributed to process optimization, improved operational efficiency, and informed business strategies.

❖ SENIOR DATA ANALYST (L2), Shriram Group of Companies [May 2019 - Jun 2020]

- Advanced Data Analysis: Conducted advanced data analysis using AWS platform tools, such as Amazon Redshift, Amazon Athena, or Amazon EMR, to derive meaningful insights and support decision-making processes.

- Data Visualization and Reporting: Created visually compelling and interactive dashboards and reports using tools like Tableau or Power BI, effectively communicating complex data analysis results to stakeholders.
- Data Governance and Quality Assurance: Implemented data governance practices, ensuring data accuracy, integrity, and security throughout the analysis process. Conducted data quality assessments and developed strategies for data cleansing and validation.

- AWS Cloud Services Expertise: Demonstrated proficiency in utilizing various AWS Cloud services, such as Amazon S3, EC2, and Glue, for data storage, processing, and ETL (Extract, Transform, Load) operations.

- Business Impact: Highlight specific instances where your data analysis and insights led to measurable improvements in business performance, such as revenue growth, cost savings, or operational efficiency.

- As a Senior Data Analyst (L2) in Shriram Group of Companies, I successfully utilized my expertise in data analysis, regression modeling, and SQL to contribute to the organization's data-driven decision-making process. My ability to understand complex business requirements, perform advanced data manipulations, and deliver actionable insights played a crucial role in enhancing operational efficiency and driving business success.

❖ DATA ANALYST, HGS PVT Ltd [Jun 2017 - Apr 2019]

- Data Analysis: Conducted in-depth analysis of large datasets using AWS Cloud tools and technologies, such as Amazon Redshift and Amazon Athena.
- Data Modeling: Designed and implemented effective data models to optimize data storage and retrieval, ensuring efficient data analysis processes.
- Data Visualization: Created visually appealing and informative dashboards and reports using tools like Tableau or Power BI, enabling stakeholders to make data-driven decisions.
- Data Cleaning and Preparation: Performed data cleaning, data transformation, and data preprocessing tasks to ensure data accuracy and quality for analysis.
- AWS Cloud Services: Leveraged AWS Cloud services, such as Amazon S3, EC2, and EMR, to store, process, and analyze large-scale datasets, ensuring scalability and cost-effectiveness.

- During my tenure at HGS Ltd, I gained valuable experience in data management, analysis, and visualization, contributing to data-driven decision-making and process improvements. My attention to detail, problem-solving skills, and ability to communicate complex findings in a clear and concise manner were instrumental in driving successful data initiatives within the organization.

❖ DATA ANALYST (Internship), Reliance Group of Companies [Dec 2014 - Apr 2017]

- Managed claims database, treated missing values, and applied data visualization techniques on the organization's dashboard.
- Conducted data cleaning and preprocessing tasks to ensure the accuracy and integrity of the claims database.
- Collaborated with the team to identify patterns and trends in the data, providing valuable insights for decision-making.
- Implemented data visualization techniques, creating interactive dashboards and visualizations to communicate complex information effectively.
- Utilized statistical analysis methods to analyze the data and identify areas for improvement in claims processing and customer satisfaction.

🞧 This internship experience allowed me to apply my data analysis skills in a real-world setting, gain hands-on experience with data management, and collaborate effectively with cross-functional teams.

PROJECTS EXPERIENCE:

❖ Recommending Ways to Increase Revenue Of A Grocery Store

The project involves conducting a thorough analysis of Point of Sale (POS) Data for providing recommendations through which a grocery store can increase its revenue by coming up with attractive combo & discount offers for customers.

Skills & Tools Covered

Market Basket Analysis, Big Data, Hadoop, Apache, Pyspark, Exploratory Data Analysis, KNIME, Python

❖ Understanding Customers' Buying Patterns for an Automobile Parts Manufacturer

This project aims to find the underlying buying patterns of the customers of an automobile part manufacturer based on the past 3 years of the Company's transaction data and hence recommend customized marketing strategies for different segments of customers.

Skills & Tools Covered

RFM Exploratory Data Analysis Python Pyspark Big Data Analysis.

❖ Online Retail Orders Analysis

This project is based on the order management functionality of an online retail store in which you are provided with the "orders" database, and you are asked some queries related to it. Answers to these queries will help the company in making data-driven decisions that will affect the overall growth of the online retail store.

Skills & Tools Covered

Joins, Sub Queries, SQL-clauses-statements-conditions, SQLite using DB Browser and MySQL Workbench

❖ Election Exit Poll Prediction and U.S.A Presidential Speech Analysis using Machine Learning

This project is based on 2 case-studies: Vote Prediction and Text Analysis. The first project is to predict which party a citizen is going to vote for on the basis of their age and according to the answers given by the citizens to the questions asked in a survey conducted. The second project is based on the analysis of the inaugural U.S.A. Presidential speeches. One has to draw inferences based on the analysis done on these speeches.

Skills & Tools Covered

Text Mining Analytics, Support Vector Machine - K Nearest Neighbor - Naive Bayes, Ensemble Techniques, Logistic Regression, Linear Discriminant Analysis.

❖ Visualizing Insurance Claims using Tableau

This project explored the art of problem-solving with the aid of visual analytics. Tableau's data visualization tools were used to create interactive dashboards to provide high-level insights to the CEO of an Insurance company to drive the company's policymaking.

Skills & Tools Covered

Business Intelligence, Tableau, Dashboard Designing

TECHNICAL CONCEPTS & SKILLS:

1	Programming Languages:	Python, R, SQL
2	Data Manipulation and Analysis	NumPy, Pandas, Stats- Models
3	Machine Learning Libraries	scikit-learn, TensorFlow, Kera's, PyTorch
4	Statistical Analysis	Statistical modeling, hypothesis testing, regression analysis
5	Data Visualization	Matplotlib, Seaborn, ggplot2, Tableau
6	Big Data Processing	Apache Hadoop, Apache Spark, Hive
7	Deep Learning	Neural networks, convolutional neural networks (CNN), recurrent neural networks (RNN)
8	Feature Engineering	Dimensionality reduction, feature selection, feature extraction
9	Model Evaluation and Validation	Cross-validation, AUC-ROC, precision-recall, confusion matrix
10	Data Mining and Exploration	Exploratory data analysis (EDA), data preprocessing, data cleaning
11	Time Series Analysis	ARIMA/SARIMA models, forecasting, seasonality analysis
12	Optimization and Simulation	Linear programming, Monte Carlo simulation
13	Cloud Platforms	Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure
14	Version Control Systems	Git, GitHub
15	Database Systems	MySQL, PostgreSQL, MongoDB
16	Data Wrangling	Data extraction, transformation, and loading (ETL)
17	Deployment and Productionization	Docker, Flask, REST APIs
18	Data Governance and Ethics	Privacy, security, compliance, ethical data handling
19	Collaborative Tools	Jupyter Notebook, Google Colab, Slack, Trello

EDUCATION QUALIFICATION:

Institute Name	University/Course	Year of Completion	Grade
McCombs School Of Business	The University of Texas at Austin. (PGDSBA)	2023	B
Nrupatunga Degree College	Osmania University of AP Graduation (B.com)(Computers	2014	B