MANASA BIJJALA

+1 (737) 615 2652 | manasabijjala.ds@gmail.com| [Linkedin](http://www.linkedin.com/in/bijjalamanasa)

**SUMMARY**

**Data Engineer with 3+ years of experience designing scalable data pipelines, optimizing big data processing, and implementing cloud solutions. Hands-on expertise in Python, SQL, Apache Spark, and Kafka, with a strong background in AWS, Azure, and Snowflake. Focused on real-time data streaming, ETL development, and performance tuning to enhance data reliability and efficiency. Driven about building resilient data architectures that drive business insights and decision-making.**

**EDUCATION**

**Master of Science in Computer Science**  
**University of Southern Mississippi August 2023 – May 2025**

**Bachelor of Science in Computer science**  
**Jawaharlal Nehru Institute of Technology | July 2018 – April 2022**

**TECHNICAL SKILLS**

* + **Programming Languages**: Python, SQL, C/C++, R.
  + **Cloud Platforms**: AWS (S3, Redshift, Glue, EMR), Azure.
  + **Big Data Technologies**: Apache PySpark, Hadoop, Kafka.
  + **Databases & Data Warehousing**: PostgreSQL, MySQL, MongoDB, Snowflake, Google Big Query.
  + **Data Visualization Tools**: Tableau, Power BI (DAX Commands), Excel, Looker.
  + **Libraries**: NumPy, Pandas, Matplotlib, Scikit-Learn, PyTorch, TensorFlow.
  + **Version Control Tools**: Git, GitHub.
  + **Data Modeling & Warehousing:** Star/Snowflake schema, dimensional modeling, OLAP/OLTP.
  + **Machine Learning**: Supervised & Unsupervised Learning, Classification, Regression, Clustering, Deep Learning (CNN), Natural Language Processing (NLP).
  + **Data Analysis & Processing**: Data Cleaning, Data Transformation, Data Wrangling
  + **Data Science & Analytics:** Data Science, Predictive Modeling, Statistical Analysis**,** Prescriptive Modeling, Quantitative Analysis
  + **Business & Analytical Skills**: Business Intelligence, Statistical Tools, Business Reporting, Data Analysis

**EXPERIENCE**

**WESTPAQ SOFTWARE SOLUTIONS | DATA ENGINEER**

**Hyderabad, India | Remote July 2022 – July2023**

* Built and maximized scalable data pipelines handling 20GB daily using PySpark, HBase, SQL, MySQL, ensuring seamless data integration and processing.
* Augmented big data performance by optimizing Spark RDDs, Data Frames, and SQL joins, leveraging caching, serialization, and shuffle optimization, achieving 20% faster processing and 15% upgraded query runtime.
* Led cloud migration from Cloudera to AWS (EMR, S3, Redshift), improving throughput by 20%, enabling scalable, high-performance data infrastructure.
* Integrated Apache Kafka and Spark for real-time data ingestion, streamlining pipeline orchestration using Oozie, boosted system uptime to 99.99% across development and production environments.
* Facilitated data-driven decisions through Tableau visualizations, presenting findings to fix the three biggest causes of data pipeline failures, increasing data accessibility and reliability scores by 25%.

**Virtual Internship by AWS Academy January 2021 – December 2022**

* Mastered AWS cloud services through a 10-week virtual internship backed by AWS Academy, culminating in proficient cloud computing and infrastructure management for application deployment.
* Leveraged AWS services, including EC2, S3, and VPC, to build a cost-effective data pipeline prototype; reduced projected data processing costs by 30% compared to traditional on-premise solutions.
* Learned about AWS storage solutions (S3, RDS, DynamoDB) and data processing tools, crucial for handling large-scale data pipelines.

**VERZEO |Machine Learning Intern**

**Hyderabad, India |Remote March 2020 – January2021**

* Programmed and optimized machine learning algorithms using NumPy and Pandas, reducing model training time by 40% and improving overall efficiency of machine learning projects.
* Refined and implemented 3 data preprocessing steps, achieving a data cleaning accuracy rate of 95%, enabling more reliable and consistent results in subsequent modeling tasks.
* Fabricated interactive Tableau dashboards by extracting data from 4+ personal projects, spotlighting key statistical insights and promoting data-driven narrative; enhanced client adoption by 15%.

**PROJECTS**

**RAMIE CROP DISEASE MONITORING| NSF FUNDED PROJECT**

**Hattiesburg, Mississippi August 2023 – December 2024**

* Designed data-driven precision agriculture pipelines using reinforcement learning, IoT sensors, and cloud computing, optimizing resource allocation and reducing consumption by 40%.
* Built real-time data pipelines for drone and IoT data processing using Python, SQL, AWS (S3, Redshift), improving pest control and cutting crop losses by 50%.
* Established a real-time alerting system via IoT sensors, reducing critical drone navigation failures by 25% and improving overall safety compliance with FAA regulations, impacting public safety.

**DATA ANALYSIS BY WEB SCRAPING USING PYTHON**

**Hyderabad, India August 2021 – May 2022**

* Created a Python script using Pandas to automate formatting and standardizing data for new data sources, helping new team members onboard to data processes 50% faster.
* Engineered robust ETL pipelines using Apache Kafka and Spark, converting unstructured web data into actionable insights; improved data processing speed by 60% and reduced query latency by 45%.
* Constructed interactive dashboards visualizing web-scraped data, delivering actionable intelligence to stakeholders, utilizing Python libraries (Pandas, Beautiful Soup, Selenium, Matplotlib); facilitated a 20% increase in efficiency.

**CERTIFICATIONS**

* **Certified in Data Structures in Python** – Expertise in algorithms & data manipulation.
* **Achieved PCAP in Python** – Validated Python programming skills.
* **Certified in C++ (CPA)** Strong OOP foundation.
* **Achieved Blue Prism Certification** – Proficient in RPA.012++
* **Certified in Machine Learning** – Internshala Trainings, 2021