

Sudheerkumar Immaneni

United States | +1 9012063385 | immanenisudheerkumar1998@gmail.com | [LinkedIn](#)

PROFESSIONAL SUMMARY

Software Engineer with 3 years of experience building machine learning models and data pipelines using AWS SageMaker, Azure Data Factory, PySpark, and SQL. Proven ability to design scalable ETL workflows, automate validation processes, and deliver predictive analytics solutions. Experienced in working with global clients like Credit Agricole and RBC on high-impact projects across finance and operations. Skilled in feature engineering, model evaluation, and cloud-based deployment using Docker and Kubernetes. Recently completed a Master's in Data Science (May 2025) with a focus on machine learning, forecasting, and real-time data processing.

WORK EXPERIENCE

Capgemini

May 2021 - Dec 2023

Data Analyst

Bangalore, India

- Designed, trained, and deployed machine learning models in a cloud environment using AWS SageMaker, enabling predictive analytics that led to a 25% improvement in forecast accuracy and a 40% reduction in model processing time.
- Performed extensive feature engineering, data cleaning, and transformation using Python (Pandas, NumPy) and SQL to prepare structured and unstructured datasets for modeling, reducing data preprocessing time by 30%.
- Built and optimized recommendation systems to improve internal user workflows, resulting in a 20% boost in task completion rates and a 15% decrease in user navigation time.
- Implemented automated model validation and testing pipelines using Python, ensuring consistency across environments and reducing production deployment errors by 20%.
- Used data visualization tools like Power BI and Matplotlib to create dashboards and reports that tracked key model metrics (precision, recall, AUC), enabling non-technical stakeholders to interpret outcomes with ease.
- Collaborated with cross-functional teams (analysts, business managers, cloud engineers) to define KPIs, collect data requirements, and translate business goals into analytical problems.
- Optimized cloud deployment infrastructure using Docker and Kubernetes, improving deployment scalability and reducing release cycles by 45%.

Capgemini

May 2021 - Dec 2023

Data Engineer | Credit Agricole Group Infrastructure Platform (CA-GIP)

Bangalore, India

- Built and optimized ETL pipelines using Azure Data Factory, ADLS, and PySpark, reducing data latency by 40% and improving processing speed by 45%.
- Automated data validation and reconciliation, increasing data accuracy by 30% and cutting manual QA time by 25%.
- Developed Spark applications to process large datasets, improving performance and saving 20+ hours of manual work monthly.
- Created SQL-based transformations with window functions and joins, enhancing job efficiency by 30%.
- Designed data workflows and ingestion pipelines for structured and semi-structured data, ensuring 99.9% data availability.

University of Memphis

Jan 2025 - May 2025

Technical Assistant – Main Campus

Memphis, TN

- Enhanced system efficiency by providing technical support and maintaining data tools (Python, R, SQL, Jupyter, MySQL) across lab systems.
- Supported research students with dataset cleaning, basic analysis, and environment setup for data projects, enhancing their research efficiency and accuracy.
- Analyzed helpdesk logs to identify patterns and optimize support workflows using data insights.

EDUCATION

University Of Memphis

Jan 2024 - May 2025

Masters, Data Science

- GPA: 3.93

PROJECTS - ACADEMICS

Analyzing Pediatric Hospitalization Data (2010-2019) and Recommendations:

2010 - 2019

- Analyzed 3+ million pediatric patient records using EDA, dimensionality reduction techniques, statistical testing, and K-Means clustering to identify patterns and sub-populations influencing health outcomes, visualized findings for clear communication.

Movie Recommendation system

- Built a Movie recommender system using collaborative filtering algorithm in Octave on Movie Lens, developed 3 Dashboards and recommended movies with 94% accuracy.

LLM Powered E-Commerce Support Assistant

- Built a chatbot using GPT-4 and T5 to handle over 10,000 daily e-commerce queries with 92% accuracy, significantly improving

customer support efficiency for order, product, and return inquiries.

- Performed Retrieval-Augmented Generation (RAG) to deliver accurate, personalized information from a catalog of 500,000+ products, increasing response precision by 30% and enhancing customer satisfaction.

CERTIFICATION

- Microsoft Certified: Fabric Data Engineer Associate | [Link](#)

SKILLS

- **Programming Languages:** C, Python, R, Power BI, SQL(MySQL, Cassandra), CSS, HTML, PySpark
- **Data Engineering:** Azure Data Factory (ADF), Azure Data Lake Storage (ADLS), Databricks, Apache Spark, ETL Pipeline Design, Data Ingestion, Data Validation, Data Reconciliation, Workflow Automation
- **Data Science & Machine Learning:** Supervised Learning (Regression, Classification), Time Series Forecasting, Recommendation Systems, Feature Engineering, Model Tuning, Model Deployment (AWS SageMaker), Model Evaluation (Accuracy, AUC, Precision, Recall)
- **Natural Language Processing (NLP):** Text Classification, Hugging Face Transformers, Basic Retrieval-Augmented Generation (RAG), Tokenization, Preprocessing
- **Visualization:** Power BI, Matplotlib, Seaborn, Dashboards, Web Scraping
- **Libraries/Frameworks:** Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Pytorch, TensorFlow, Git, MLflow, Airflow (Basic), Post-migration Support, Technical Documentation
- **Cloud & DevOps:** AWS (SageMaker, S3), Azure (ADF, ADLS, Synapse), CI/CD Pipelines
- **Big Data Technologies:** Awareness of Hive, Hadoop, Spark, Snowflake, Air flow
- **Soft Skills:** Team Collaboration, Leadership, Problem-Solving, Communication, Agile Methodologies, Cross-functional Coordination