SHREYANTH S
DATA ENGINEER

CONTACT INFO

✉ shreyanth0810@gmail.com

🌐 www.shreyanth07.info

in www.linkedin.com/in/Shreyanth07

EDUCATION

2021 - 2023
MTech in Data Science & Engineering
Birla Institute of Technology and Science - Pilani

2015 - 2019
BE in Electronics & Communication Engineering
Anna University - Chennai

EXPERIENCES

September 2020 - Present
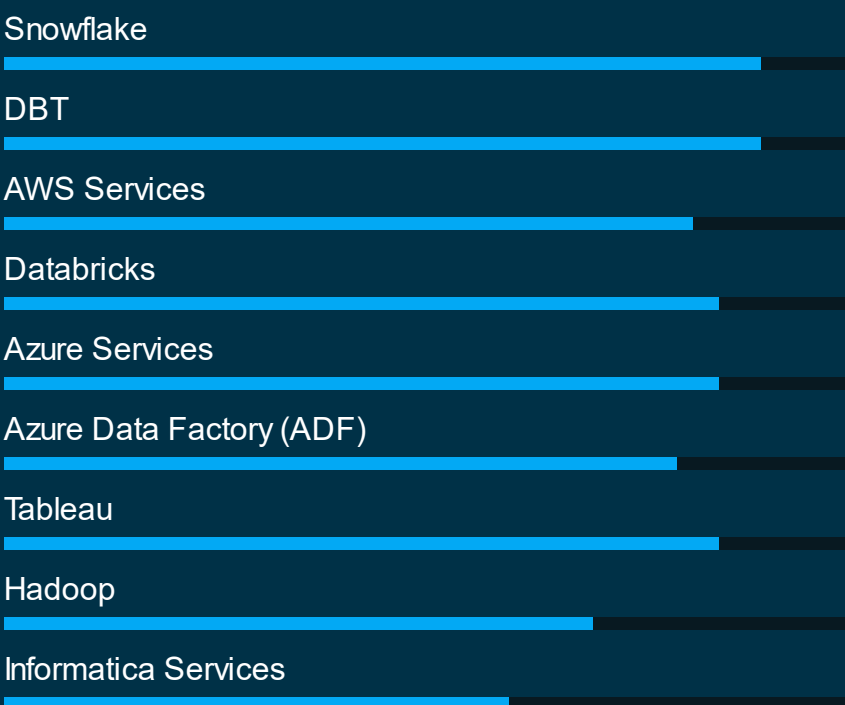Data Engineer
Indium Software (India) Pvt Ltd

May 2019 - May 2020
Product Architect
Sciencotonic Edu India Pvt Ltd

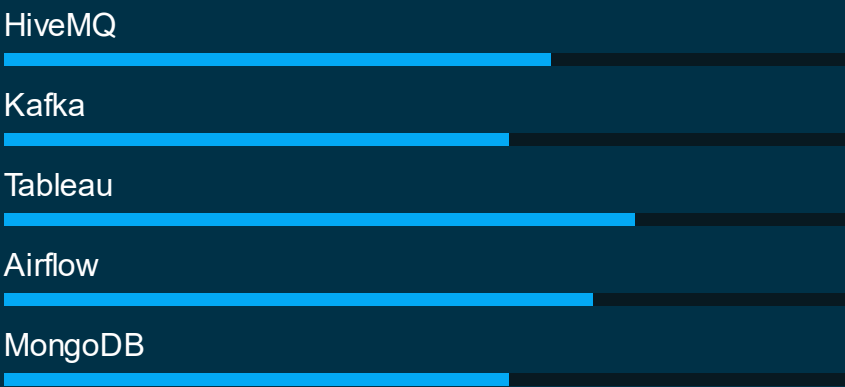**November 2018 - Present**
Software Developer
Upwork Inc

PROGRAMMING SKILLS

Python

SQL

PySpark

Java

JavaScript

PRIMARY TOOLS/TECHNOLOGY

Snowflake

DBT

AWS Services

Databricks

Azure Services

Azure Data Factory (ADF)

Tableau

Hadoop

Informatica Services

SECONDARY TOOLS/TECHNOLOGY

HiveMQ

Kafka

Tableau

Airflow

MongoDB

LANGUAGES

English

Tamil

Hindi

AWARDS

🪪 Gem Award - Q1 2023 (Indium)

🪪 Spot Award - May 2023 (Indium)

🪪 Data Ninja Q1 2022 (Indium)

🪪 Data Ninja H2 2021 (Indium)

PROFILE

I am a Data Engineer with 4 years of expertise in ETL (Extract, Transform, Load), data warehousing, and developed several successful projects in data engineering. I have extensive experience in data architecture, modeling, and analysis and am passionate about building innovative solutions that help businesses leverage their data. I have experience developing and deploying data-driven solutions to manage and analyze data, as well as building data pipelines to access and transform data.

I have experience in working with databases/data cloud such as Snowflake, Oracle and SQL Server, and am proficient in programming languages such as Python, SQL, and Java. Additionally, I have hands-on experience in deploying solutions using Spark, Databricks, AWS and Azure. I am passionate about solving complex data problems and leveraging analytics to create value for organizations.

EXPERIENCES

SEP 2020 - PRESENT
INDIUM SOFTWARE (INDIA) PVT LTD

DATA ENGINEER

- Working for major and leading clients in Supply chain management, Affiliate Marketing and Healthcare and Power Resource Management domains
- Create a centralized platform, integrating data sources from heterogenous database systems to create a unified source.
- Primarily focused on ETL/ELT data from Oracle to Snowflake using DBT for Transformation and dimensional modelling for transformed data in Snowflake.
- Designed, implemented, and optimized data pipelines using Apache Spark and Databricks on Azure, enabling efficient ETL processes for large-scale datasets.
- Develop Snowpipe for real-time ingestion from AWS S3 to snowflake for Development from Semi-Structured and Unstructured data formats.
- Implemented scalable analytics platform from the ground up, including Snowflake, Rivery, dbt transformations, and Tableau visualizations.
- Create new business data models based on star schema methodology, allowing for scalable reporting in Snowflake warehouse.
- Create and maintain warehouse permissions for multiple users and software as part of data governance.
- Implement multiple strategies and development for automation of procedures in python and snowflake stored procedures for undergoing inhouse master data management.
- Formulating a procedure for automated standardization of data and maintain, improve and develop new or existing Python scripts for automated pipeline, reporting and Data model creation.
- Automate and develop data lineage, data quality and data masking on the master data as per the security policies and governance.
- Create Snowpipe functionality inside of Snowflake to automatically read file changes to Amazon S3 buckets and custom APIs for data sources using Rivery and Precog ETL interfaces.
- Spearheaded the migration of on-premises data infrastructure to Azure cloud, resulting in a 40% reduction in operational costs and a 30% increase in data accessibility.
- Implemented data security measures, including encryption at rest and in transit, ensuring compliance with industry regulations and best practices.
- Optimized Databricks clusters for cost-effectiveness and performance, achieving a 20% reduction in infrastructure expenses while maintaining high processing speeds.
- Design data architecture for a custom platform to allow for a scalable marketplace.
- Build, define, and manage team KPIs and OKRs for quarterly achievements, and build dimensional models of transformed data in snowflake.
- Automate and consolidate business reporting processes to ensure ownership of data and accurate metric reporting along with automation of procedures in Snowflake using python and SQL for inhouse Master Data Management.
- Implementation of custom metadata fields across multiple sources to ensure accurate financial reporting.
- Explore POCs on tools and technologies as per the requirement and worked on different SCM modules like order management, quality control, order tracking and more.
- Led end-to-end design, implementation, and execution of projects utilizing Amazon SageMaker, effectively showcasing the platform's capabilities and demonstrating data pipeline orchestration, model deployment, and real-time data processing.
- Developed scalable and efficient data pipelines for POCs, leveraging AWS services like Amazon S3, Glue, and Lambda to extract, transform, and load (ETL) data, enabling seamless integration with SageMaker for advanced analytics and machine learning experiments.
- Collaborated cross-functionally with data scientists and developers to define requirements, select appropriate machine learning algorithms, and build predictive models within SageMaker, resulting in actionable insights and data-driven decision-making for the organization.
- Worked in Agile Methodology of Scrum along with Bitbucket, Github and Git for Versioning.
- Nominated and worked in multiple projects and Sub projects of clients, formulating the business requirements and transforming business problems to the most suitable solutions.
- Mentored junior team members and leading the Data Engineering team in my recent projects.

SKILLS

Snowflake, Amazon SageMaker, AWS S3, EC2, Glue, Amazon Kinesis, DBT, AWS, Databricks, Azure, ADF, Rivery, Apache Spark, Tableau, Data Migration & manipulation, Analysis & Modeling, Data Warehousing - Pipeline & Architecture, ELT/ETL, Data Preparation, HiveMQ, Kafka, Airflow

MAY 2019 - MAY 2020
SCIENCOTONIC EDU (INDIA) PVT LTD

PRODUCT ARCHITECT

- Perform design, development and support of existing and new products.
- Understand product vision and business needs to define product requirements and product architectural solutions.
- Develop architectural and design principles to improve performance, capacity, and scalability of products.
- Develop and maintain the system roadmap, ensuring scalability, performance, and user experience.
- Design and implement algorithms and models to personalize content recommendations based on user behavior and preferences.
- Work closely with data engineers and data scientists to integrate data pipelines and optimize the data infrastructure for the recommendation system.
- Conduct regular system performance evaluations and implement improvements to enhance the accuracy and effectiveness of the recommendation algorithms.
- Provide technical guidance and mentorship to the development team, ensuring adherence to architectural standards and best practices.
- Planning and execution of new product releases and consulting the business management team to clarify objectives and functional requirements for new or modified products.
- Maintain product roadmap and architectural standards that assure product development projects optimally align with business objectives.
- Define product validation policies so as to deliver products that meet system and project expectations and facilitate the creation, review, and sign off of project deliverables.
- Analyse market segments and customer base to develop market solutions and define product requirements that address market opportunities.
- Work with the technical team in development and issues needing technical expertise or complex programming and systems knowledge.

SKILLS

Digital Transformation, Team Management, Requirement Analysis, Data Migration, AWS, SDLC, Data Analysis and Visualization, Competitive Analysis, Product Design, Software as a Service (SaaS), Project Management, Project Planning, Technical Architecture, Product Development

NOV 2018 - MAY 2021
UPWORK INC.

SOFTWARE DEVELOPER

- Successfully completed 19 diverse projects, achieving an impressive 5/5 rating on 11 projects, highlighting exceptional client satisfaction. Maintained an exceptional average rating of 4.75, reflecting consistently high-quality work and client service.
- Proficiently utilized D3.js and Python to create data analysis charts, enabling clients to gain valuable insights from customer data showcasing a strong grasp of data analytics and visualization, essential for modern software development.
- Designed user-friendly interfaces for contact management, integrating Google Maps for location visualization with advanced filtering capabilities. Developed e-signature interfaces using Java and JavaScript, enhancing user experience and efficiency.
- Leveraged data engineering techniques to monitor production processes, create data processing jobs, and resolve bugs and issues in customer products. Demonstrated an aptitude for maintaining data integrity and optimizing data-driven systems for enhanced decision-making processes.
- Skillfully developed interfaces for static graphs using Matplotlib and Python, showcasing competence in data visualization libraries beyond D3.js for handling various graphical data representation tasks.

SKILLS

Project Management, AWS, Data Analysis, Data Visualization, D3.js, Python, User Interface Design, Google Maps Integration, Java, JavaScript

CERTIFICATIONS

CS50X, CS50P, CS50AI Harvard University Distinction - 91%	**Statistics, Algorithms, CS 101** Stanford University Distinction - 85%	**ML, DBMS, Discrete Mathematics** NPTEL - IIT Distinction - 80%
Applied Data Science I & II WorldQuant University Distinction - 95%	**Advanced Algorithms & Big Data** Coursera - UCSD Distinction - 100%	**Python for Everybody** Coursera - UMich Distinction - 98%
Computer Vision Basics Coursera - UB SUNY Distinction - 100%	**Mathematics for ML** Coursera - ICL Distinction - 100%	**ML with Python** FreeCodeCamp Distinction - 100%