UJJAWAL SINHA

Senior Software Engineer

Phone: +91 7873878103 | **Email:** ujjawalsinhacool16021998@gmail.com

LinkedIn: https://linkedin.com/in/sinha-ujjawal | **GitHub:** https://github.com/Sinha-ujjawal | **Portfolio:** https://ujjawal-sinha.vercel.app

---

EXPERIENCE

Sr. Data Engineer | Circana

Bengaluru, Karnataka | January 2024 – Present

- Created an NLP model for coding new products with market structure attributes using tools like Python, Pyspark, Scikit-Learn and used Optuna for Hyperparameter tuning.
- Developing and maintaining data pipelines to process UPCs and impressions data from clients, and combine with our internal data to generate final reports showing key KPIs like total sales, buyers, and trips. Technologies used: Python, Apache Spark, Hadoop, Hive, PostgreSQL, AWS S3.
- Automate code deployment using Jenkins adhering to CI/CD principle.
- Migrate existing hive workloads to Pyspark and scheduling them on Airflow.
- Mentoring four junior developers and helping them completing their tasks, clarifying their doubts on how the code is working and helping them clear their blockers
- Technologies: Python, Apache Spark, Hadoop, Hive, PostgreSQL, Azure, Scikit-Learn, Optuna, Jenkins

Data Engineer | Swiss Re

Bengaluru, Karnataka | February 2021 – January 2024

- Automating tagging of similar entities (insured and broker company names) into clusters using Jaccard distance algorithm. Tools used – Python, PySpark, Palantir Foundry (Code repo, Object explorer, Data lineage etc.).
- Optimized existing data pipeline written in R by analysing SQL queries, removing redundancy, parallelizing etc. adhering to clean code principles.
- Automated reports and email alerts for DQ health scores in the internal CRM tool using Python, PySpark, SQL, and Palantir Foundry.
- Automated RFM analysis to calculate overall health scores of Managed Customers & Brokers based on different KPIs (Premium earned, profitability, Cross-Selling, Claims etc.), using methods like mean and percentile rank. Tools used – Python, Pyspark, Palantir Foundry (Code repo, Object explorer, Data lineage etc.).
- Made a ML model to classify NPS responses to a predefined list of classes/labels using SparkML, pyspark.
- Tools: Python, PySpark, Palantir Foundry (Code repo, Object explorer, Data lineage)

Data Engineer - 1 | Nissan Digital

Trivandrum, Kerala | July 2019 - February 2021

- Developed ETL script for a tool which recommends time-phased safety stock using statistical methods (modified King's formula) to help maintain optimized inventory. Tools used were Python, Prefect, SQLAlchemy, AWS Services like Lambda, RDS, and Docker.
- Made reusable modules for interacting with AWS services (S3 file upload/ download, triggering AWS Batch jobs, interacting with RDS, etc.).
- Developed ETL script for a tool which estimate volume of different vehicle models, in different regions. Tools used were Python, Prefect, SQLAlchemy, AWS service like EC2.
- Tools: Python, Prefect, SQLAlchemy, AWS (Lambda, RDS, EC2), Docker

---

TECHNICAL SKILLS

- **Programming:** Python, Golang, Bash scripting
- **Databases:** SQL, PostgreSQL, Oracle, SQL Server
- **Big Data:** Apache Spark, Hadoop, Hive, PySpark, Pandas, Polars
- **Cloud/DevOps:** AWS (S3, Lambda, EC2, Batch), Azure, Jenkins, Docker
- **Tools:** Apache Airflow, Py-Spark, Scikit-Learn, Palantir Foundry

---

ACHIEVEMENTS

- Special recognition for optimizing R scripts (Swiss Re Triannual Awards)
- Spotlight award for project contribution (Nissan Digital)
- Finalist in 7th CSI In-App National Project Awards

EDUCATION

B.Tech, Computer Science KIIT University, Bhubaneswar, Odisha CGPA: 8.9 | 2015-2019

CERTIFICATIONS

Certified Palantir Foundry Data Engineer Professional Palantir | 2023