

# Sudeep Tripathy

## Software Engineer

+91-982-366-7554  [sud.tripathy@gmail.com](mailto:sud.tripathy@gmail.com)  [www.linkedin.com/in/tsudeep](https://www.linkedin.com/in/tsudeep)

### SKILLS

#### Area of Expertise:

- Automation framework design and development
- Big Data Test Engineer
- Data Pipelines and Engineering
- Storage Systems
- Tools Development
- Python, R, SQL
- Hadoop, Hive, Spark, Kafka
- SAN, NAS, iSCSI, FC, NFS
- API Automation
- Quality Assurance
- Agile Methodologies
- Virtualization
- Dockers and Kubernetes

#### Work-Like-Experience:

- Business Intelligence
- Machine Learning Models
- Predictive Analysis Lifecycle
- Exploratory Data Analysis
- Time series forecasting
- NLP - text analysis
- Tableau

### VISA

USA-B1 (Business Visa) valid till 2026

### ACADEMICS

#### Post Graduation Program in Business Analytics & Business Intelligence

McCombs School of Business,  
University of Texas

#### Master in Big Data Analytics

Great Lakes Institute of Management

### PROFILE

- **Software Engineer** with more than **8.6 years** experience in **designing automation strategy** for end to end software systems, **backend, api, middleware, hardware, sub component performance** and **architecture** testing using **Python**.
- A data enthusiast with more than **5 years** of experience in **data analytics, data engineering, data science** supporting business domains impacting **product quality** and **AI driven customer care** complementing sales.
- Automation experience in **big data analytics real time systems, data pipelines** and technologies like **Kafka, Zookeeper, Redis, Hive, Hadoop, Spark**.
- Testing and automation experience in **Storage Systems, Storage Protocols** and technologies such as **SAN, FC, FCoE, iSCSI, LVM, NPIV, RDM, NFS**.
- Experience in Agile, CI/CD, Version Control practices and hands on tools such as **GIT, JIRA, Confluence, Jenkins**.
- Experience in working with **globally distributed teams** and travelling onsite for project coordination, knowledge transfer. Contact point for any engineering issues, bug triage, bug tracking into releases.
- Exposure to **ML models** and **web sentiment analysis**.

#### NetApp / Member Technical Staff

DECEMBER 2018 - PRESENT, BANGALORE

Project Title - Customer Retention Model using Auto Support

Objective - To identify high risk in product functionality leading to customer attrition and predict year-over-year adoption of product features through an integrated interactive dashboard generating reports.

Outcome - Automatic query generator to replicate customer configurations. Integration with product defect databases to predict likelihood of defects yielding increased quality metrics helping retention.

Techniques - Data Mining, Data Engineering, Logistic Regression, WSA

Tools Used - R, Python, HIVE, SQL, Tableau

Project Title - Netapp In Place Analytics

Objective - Enable NAS storage to run in-place analytics

Outcome - Driver for Hadoop clusters to access NFS data.

Techniques - File Systems, Data Analytics, System Engineering, Hadoop

Tools Used - Python

## Post Graduation Diploma in Advance Computing (DAC)

C-DAC:Centre for Development of Advanced Computing, Pune, India

## Bachelor of Engineering (BE) in Computer Science

G.H.Raisoni College of Engineering (GHRCE), Nagpur University, Nagpur, Maharashtra

### DATA SCIENCE/ML PROJECT

**Project Title** - Recommendations Proposal for Supply Chain Company

**Outcome** - Data driven business insights for supply chain impacting timely delivery of goods. The model implementation helped the business retain 650 customers worth of \$139M AUM.

**Techniques** - Machine Learning, Logistic Regression, Xgboost, Tableau

**Tools Used** - R/Python

**Project Title** - India Credit Risk

**Outcome** - The project involved developing a credit risk default model to predict loan defaults.

**Techniques** - Logistic Regression, Random Forest

**Tools Used** - R/Python

**Project Title** - Build a forecasting model to predict monthly gas production.

**Outcome** - The project involved developing an ARIMA model to forecast gas production level for final 12 months.

**Techniques** - Time series forecasting, ADF Test, ARIMA

**Tools Used** - R/Python

**Project Title** - Choosing preferable mode of transport for employees.

**Outcome** - The project involved deciding on the mode of transport employees prefer while commuting to office.

**Techniques** - KNN, Naive Bayes

**Tools Used** - R/Python

## Deutsche Bank / Associate

MARCH 2018 - DECEMBER 2018, PUNE

**Project Title** - Global Transaction Banking

**Objective** - Imbibe intelligence in test framework reporting and troubleshooting.

**Outcome** - Designed, developed and implemented end to end automation framework involving api and services automation, data validation in a BDD framework while matching data from two different databases.

**Techniques** - Data Engineering, Hadoop

**Tools Used** - Python, SQL, HIVE, Robot, Scoop

## Msys Technologies / Senior Engineer

FEBRUARY 2015 - MARCH 2018, PUNE

**Project Title** - Infosight (Predictive Analysis of Storage Arrays) [Nimble]

**Objective** - Develop tools for data pipeline validation and monitoring.

**Outcome** - Data extraction, count of records in database, pipeline subsystem verification, visual automation using prometheus and Grafana real time charts. Making infosight pod architecture compliant.

**Techniques** - Data Engineering, Storage

**Tools Used** - Python, Kafka, Redis, Volt-db, Vertica-db, SQL

**Project Title** - Infinite Memory Engine [Data Direct Networks]

**Objective** - Design SSD failure hardware automation suite and integrate with benchmarking tools.

**Outcome** - Fault injection techniques applied with benchmarking tools to validate the software system against industry standards like mdtest, Bonie++, iozone.

**Techniques** - Storage, System Engineering, Automation

**Tools Used** - Python, Lustre, PFS, SSD

## Cumulus Systems [Hitachi] / Member Technical Staff

MAY 2012 - FEBRUARY 2015, PUNE

**Project Title** - Hitachi Common Code Module

**Objective** - Develop cli automation suite.

**Outcome** - Every failure condition of the cli was tested.

**Techniques** - Storage, System Engineering, Automation

**Tools Used** - Python

**Project Title** - SAP-Landscape Virtualization Management

**Objective** - Develop automation framework

**Outcome** - Minimal acceptance test suite with heterogeneous server OS

**Techniques** - Storage, System Engineering, Automation, Virtualization

**Tools Used** - Python

---