Deepak Kumar

Data Engineer

+14089088074 ◇ deepak.tech@icloud.com ◇ Fremont, CA, United States ◇ LinkedIn

SUMMARY

Results-driven software engineer specializing in data platforms with over 8 years of experience. Skilled in implementing and optimizing robust data infrastructure and software solutions across multiple domains including social media, transportation, banking, airlines, and e-commerce. I have experience in big data technology stacks, ETL processes, Database management, Cloud computing implementation and web-based application development & testing. Research experience covering quantitative methodology, finding research gap, survey design, statistical analysis, hypothesis testing etc.

SKILLS

Programming Languages: Python, Java
Database Management: MySQL, PostgreSQL, Oracle, Mango DB, Hive
Web Development: HTML, CSS, NodeJS
Testing and Automation: Selenium, Cucumber, Test NG, Pytest
Cloud and Big Data Technologies: AWS, Docker, Hadoop, Redshift, Spark, Presto, Kafka, Informatica, Airflow, TigerGraph Db, GCP, Data Proc

EXPERIENCE

Data Engineer   Oct '21 — Present
Meta (Contract)   Menlo Park, United States
- Implemented data processing pipelines to clean, transform, and structure raw data, ensuring it meets quality & integrity standards and efficiently migrated a petabyte-scale dataset from one geographic region to another.
- Spearheaded compliance initiatives, focusing on identifying and addressing instances of violence and unethical content within user videos and reels based on Internal Monitoring System report.
- Led efforts in deprecating outdated source data, proactively updating data ingestion pipelines to align with the evolving landscape of absolute source data and optimizing resource utilization for enhanced performance & efficiency.
- Created insightful dashboards utilizing Tableau and Unidash, providing stakeholders with a visually intuitive and comprehensive understanding of key performance indicators and data insights.
- Developed DAGs for both production-ready and experimental datasets utilizing Python, Java, SQL, APIs, Presto, Spark, and machine learning modules.
- Developed framework to detect fraud on Marketplace platform using Spark, TigerGraph Db, Kafka
- Performed A/B testing, experimentation, curated & provided training data for LLM and other AI models

Data Engineer   Mar '20 — Oct '21
i2Data Systems worked for client Lyft   San Francisco, United States
- Conducted analysis of business requirements for booking, rides, and daily spend reporting for Accounting Team.
- Developed a comprehensive data ingestion framework using Apache Airflow, Glue, EMR, Redshift, Athena, and Lambda on AWS, facilitating the ingestion of data from Oracle FAH, MySQL, and CRM into HDFS and S
- Developed business reports & dashboard encompassing various Key Performance Indicators (KPIs) using MODE and Tableau which includes near real-time data tracking related to booking, and ride.

Data Engineer   Nov '18 — Dec '19
Facebook (Contract)   Menlo Park, United States
- Worked for video and feed data engineering team to develop & maintain efficient ETL process using big data stacks.
- Developed dashboard, executed data migration, compliances implementation and ad revenue calculation.
- Involved in backfill, migration, data infra resources real-time monitoring and on-call resources for Global Ops team.

Hadoop Developer   Mar '16 — Dec '17
Pro-Teck Consulting on assignment at Wells Fargo   Fremont, United States
- Worked on design and implemented data models, schemas, and storage solutions optimized for Hadoop environments.
- Created efficient and reliable code to process and analyze data, ensuring data quality, integrity, and security standards.
- Collaborated with business stakeholders to translate their requirements into technical solutions while adhering to Agile.

Associate IT Consultant   Aug '11 — Apr '15
ITC Infotech Ltd.   Bengaluru, India
- Used Java, JSP, and Servlets framework for creating web-based application.
- Developed UI using HTML, JavaScript, and JSP, and developed Business Logic and Interfacing components using Business Objects, XML, and JDBC.
- Used MAVEN to build the project into WAR and EAR files and Clear Case for version controlling.
- Responsible for coding SQL Statements and Stored procedures for back-end communication using JDBC.
- Worked extensively on bug fixes on the server side and made cosmetic changes on the UI side.

- Developed testing framework using Java and selenium for regression & cross platform testing and performance monitoring

EDUCATION

Doctor of Philosophy in Information Technology, University of the Cumberlands (GPA: 3.7)

Jan '18 — Dec '22
Williamsburg, United States

- Completed 83 credits related to Information Technology *Link*

Master of Science in Computer Science, San Francisco Bay University (GPA: 3.8)

May '15 — Dec '16
Fremont, United States

- Completed 36 credits related to Computer Science

PROJECTS

Factor related to adoption of IoT in Smart Home, University of the Cumberlands Link

May '21 — Dec '22
United States

- Completed dissertation as the part of PhD program

AWARDS

Global Recognition Awards
Global Recognition Awards

Sep '24

CERTIFICATIONS

Judge, Globee® Awards for Business — Oct '24
Advance SQL for Data Scientists — Nov '18
Introduction to Cloud Computing — Nov '18

PUBLICATIONS

Research Paper
Link: ORCID

Optimized Support Vector Machine Based Fused IoT Network Security Management
IEEE

ChOs_LSTM: chebyshev osprey optimization-based model for detecting attacks
IEEE

An Advanced Wasserstein-Enabled Generative Adversarial Network Enabled Attack Detection for Blockchain-Assisted Intelligent Transportation System

Jul '24

An Efficient DDoS Attack Detection using Attention based Hybrid Model in Blockchain based SDN-IoT
IEEE

Jul '24

Enhancing Cyber Security Using Quantum Computing and Artificial Intelligence: A Review
International Journal of Advanced Research in Science, Communication and Technology

Jun '24

VOLUNTEERING

Mentor , ADPList

United States
Jan '24 — Present

I have been serving as **mentor**in Computer Science and Information Technology to students, and software engineers.

Reviewer

United States
Jan '24 — Present

I have reviewed numerous research papers worldwide in the fields of Computer Science and Information Technology. I have served on the review boards of various IEEE conferences in Malaysia, China, the UK, India, and Egypt. Additionally, I have completed more than 50 reviews for MDPI, Springer, and IEEE conferences.

MEMBERSHIP

Senior Member, IEEE, USA
Member, SigmaXi, USA
Member, International Association of Engineers, Hong Kong
Program Committee Member: PyTorch Conference 2025 from The Linux Foundation
Guest Reviewer- MDPI, Springer Nature, IEEE, IOP