

# Deepana Dhakshinamurthy

## Senior Data Engineer

Boston | +1 (857)-379-6706 | [deepanadhakshinamurthy@gmail.com](mailto:deepanadhakshinamurthy@gmail.com) | [Deepana GitHub](#) | [LinkedIn](#) | [Portfolio](#)

### SUMMARY

Data Engineer with 9 years of expertise in SQL, Data Engineering, Data Modelling, Business Intelligence, ETL Pipelines, Cloud platforms (AWS, Azure and GCP), Performance Tuning and Data Visualizations. Proven track record of delivering solutions for US clients across the **healthcare, insurance, and financial** industries

### TECHNICAL SKILLS

- **Programming language:** Python, SQL, HTML, C, C++, UNIX, Excel (VLOOKUP, Pivot, Charts)
- **Data Engineering Tools:** Power BI, Tableau, AWS Glue, ADF, Databricks, DBT, Airflow, Docker, SSIS, SSRS, AWS S3
- **Python Libraries:** Pandas, Numpy, Matplotlib, Scikit-learn, TensorFlow, SciPy, Seaborn, Plotly, PyTorch
- **Database Management System:** SQL Server, Oracle, PostgreSQL, NoSQL, MySQL, Azure SQL, AWS RDS, Redshift
- **Project Management & Tools:** Agile (Scrum), SDLC, Jira, Azure DevOps, Confluence, TFS, GitHub, Tortoise, SVN
- **Office Suite:** Microsoft Word, PowerPoint and Excel, Google sheets, Microsoft 365, Outlook, OneNote
- **Machine Learning:** Classification, Regression, Random Forest, Linear Models, Decision Trees, Neural Networks, LLM

### PROFESSIONAL EXPERIENCE

#### Colorado Department of Transportation

June 2024 to Dec 2024

Denver, Colorado

##### Data Analytics Intern (*SQL, Power BI, Python*)

- Constructed interactive **Paginated Power BI reports** and Dashboards (**DAX Queries**) to analyse Colorado's state bridge inspections, enhancing financial decision-making and reducing **inspection delays by 25%**
- Developed ETL pipelines to integrate a new data model into existing system, improving data accessibility by 35%
- Participated in **Scrum** ceremonies of **Agile** Methodology, backlog tracking using **Jira** as per **Sprint** goals

#### Cognizant Technology Solutions

July 2022 to June 2023

Chennai, India

##### Senior Associate/Technical Lead (*SQL (SQL Server, MySQL), SSIS, Azure, SSMS, Python*)

- Led Data Engineering projects in Insurance, Banking and Healthcare by collaborating with stakeholders, cross-functional teams to understand the business requirements, and mentored 5+ junior resources on their assignments
- Engineered ETL pipelines using complex SQL procedures (window functions, CTEs, complex joins), SSIS, Azure Data Factory, **Databricks**, Data Lake, CDC to extract and transform data (healthcare, insurance) **reducing processing time by 40%**
- Enhanced ETL workflows by reducing processing time (**Performance Tuning**) by 50% through Indexing SQL Server procedures, Execution Plan, **Logging, QA checks**, and Data Analysis of **complex Datasets** using Python, SQL and Excel
- Administered data workflows involving **multi-terabyte financial and insurance datasets** using **AWS Glue**, Data Lake, Lambda, S3, Redshift, Snowflake; utilized **dbt models** and **Snowpipe** for ingestion, maintained pipelines in **Snowpark**
- Automated Python data pipelines for data loading and transformation and orchestrated workflows using **Airflow in Docker** containers with **CI/CD** deployment via **GitHub Action/Azure DevOps**, resulting in **60% faster deployment cycles**

#### Associate (*SQL, Power BI, Tableau, Snowflake, Excel*)

Nov 2017 to June 2022

- Launched **conceptual, logical, and physical Data Models** for 20+ entities with the help of Redgate, SQL Server; Proficiently resolved **70% of critical production** environment issues by implementing **data governance** and data **quality** standards
- Migrated ETL packages, SQL solutions from SSIS (On prem) to ADF and AWS leading to **35% reduction** in compute costs
- Designed **20 + SSRS** reports for monthly and bi-weekly insurance billing and launched **OLAP** data models utilizing SSAS
- Crafted Dashboards using **Power BI DAX queries, Looker, Tableau** for financial, healthcare and insurance Products, monitored **KPI** metrics which boosted operational efficiency by **30%**

#### Merit Software Services Private Ltd

Oct 2016 to Oct 2017

Chennai, India

##### SQL Developer (*SQL, SSIS, C# (Script Task), SSRS*)

- Created SQL queries, SSRS reports, and SSIS packages aligned with backend requirements of a web portal
- Automated a manual data loading process using SSIS packages, improving data loading efficiency by 40%

#### Capgemini

July 2014 to Sep 2016

Navi Mumbai, India

##### Senior Software Engineer (*MSSQL, SSIS, SSRS, PLSQL*)

- Analyzing data using SQL, Excel and reports generation through SSRS
- Enhanced File Transfer Process (FTP) by 40% (utilized SSIS and replaced the usage of command prompt)

### ACADEMIC PROJECT (*Northeastern University*)

#### Diabetes Readmission Prediction System (*ML, Python, Big Query, GCP, Looker, Airflow, Docker*)

Jan 2025 to Apr 2025

- Conducted **statistical analysis** on patient records via hypothesis testing, correlation, feature selection to identify key factors
- Optimized and evaluated Machine Learning models (**XGBoost, Random Forest, Decision Tree**) using **MLFlow** and evaluated performance using key metrics like **MSE, RMSE, and R-squared**; utilized **DVC** for data version control
- Implemented unit and integration testing for data transformation, model build pipelines using **Pytest**
- Persisted model prediction outputs to **GCP BigQuery** and built interactive Looker dashboards to monitor response time
- Orchestrated data workflows with Apache Airflow in **Docker** containers on **GCP VM**, deployed ML models on **GCP buckets** and managed CI/CD pipelines via GitHub Actions for streamlined deployment

**Chatbot Implementation (*Snowflake, LLM, Streamlit*)***Jan 2025 to Apr 2025*

- Built an interactive multi-turn recommendation chatbot with **Streamlit** UI, enabling personalized user experiences
- Integrated data from **AWS S3** into **Snowflake** via **Snowpipe**, designed normalized models using **DBT** data models
- Implemented **LangChain + Ollama LLM** for natural language understanding; generated MiniLM embeddings stored in Snowflake and used **FAISS** for efficient semantic search and personalized recommendations

**S&P 500 Stock Analysis (*ML, Python, Tableau*)***Jan 2024 to April 2024*

- Conducted **Predictive Analysis** of S&P 500 stock values by collecting historical data from S&P associated firms
- Architected Tableau dashboards and conducted **exploratory data analysis (EDA)** to visualize stock performance
- Predicted stock values using ML models such as Regression, **KNN, Random Forest** and **Decision Tree** and evaluated the performance using statistical metrics F1 Score, MSE, R-Squared achieving an excellent fit with an R-squared of 0.9

**FITHUB Management System (*NoSQL, MySQL, Python*)***Sep 2023 to Dec 2023*

- Designed hybrid database using MySQL for structured data (users, transactions), NoSQL for unstructured data (activity logs)
- Developed stored procedures, queries, functions, EDA using Python to manage and analyze user details and daily activities

**EDUCATION****Northeastern University, Boston***Aug 2025***Master of Science in Data Analytics Engineering***(Expected)***Coursework:** Data Mining, Gen AI/LLM in Data Engineering, Machine Learning Operations (MLOps)**GPA: 4****SRM Valliammai Engineering College, Chennai***April 2014***Bachelor of Engineering in Electrical and Electronics****ACHIEVEMENTS**

Received **STAR** award, interim hikes, client appreciations for delivery excellence at Cognizant