Deepana Dhakshinamurthy
Senior Data Engineer

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

SUMMARY

Data Engineer with 9 years of expertise in SQL, Data Engineering, Database Management, Data Modelling, **Business Intelligence**, ETL Pipelines, Data Warehousing, Cloud platforms (AWS, Azure and GCP) and Data Visualizations. Specialized in leading Performance Tuning, Machine Learning, Cloud Migration projects under **Financial**, Insurance and Healthcare domain

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, Alteryx
- **Cloud Platform & Storage:** AWS, Azure, GCP, AWS S3, Azure Data Lake
- **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
- **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*
Data Analytics Intern *(SQL, Power BI, Python)*   *Denver, Colorado*

- 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 SQL and Python ETL pipelines** to integrate a new data model into existing 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*
Senior Associate/Technical Lead *(SQL (Azure SQL, MSSQL), Azure, SSIS, Python, AWS)*   *Chennai, India*

- Engaged in Financial, Insurance, Healthcare related Data Engineering projects by collaborating with Stake holders and cross-functional teams to understand business requirements and providing technical guidance
- Demonstrated expertise in building complex **ETL pipelines** using SQL procedures, **SSIS**, **ADF, AWS Glue, DBT Models** to extract Data from **XML, Json, API** and perform Data Integration, Transformations and Loading
- Enhanced ETL workflows by reducing processing time **(Performance Tuning) by 50%** through **Indexing**, SQL Profiler, Execution Plan analysis, Unit Testing, and **Data Validation** utilizing SQL, Python, and Excel
- Built Python ETL pipelines for data transformations and orchestrated Airflow workflows in Docker containers (**CI/CD Github**)
- Engineered ML models (e.g., **Regression, Random Forest**) using **Python** to predict Insurance claims and evaluated Metrics

Associate *(SQL, Power BI, Tableau, Snowflake, Excel)*   *Nov 2017 to June 2022*

- Migrated ETL pipelines to cloud platforms such as Azure, AWS, and Snowflake resulting in a **35% reduction** in computational
- Monitored workflows management using **Azure DevOps (ADO)** for seamless CI/CD integration and proficiently resolved 70% of critical production issues by implementing **Data Governance** standards
- Crafted Dashboards using **Power BI** DAX Queries and **Tableau** in Healthcare and **Insurance** Products, monitored **KPI** metrics resulting in a notable **30%** improvement in operational efficiency

Merit Software Services Private Ltd   *Oct 2016 to Oct 2017*
SQL Developer *(SQL, SSIS, C# (Script Task), SSRS)*   *Chennai, India*

- 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%
- Implemented Error Handling in the automation process by triggering emails automatically and notifying relevant teams

Capgemini   *July 2014 to Sep 2016*
Senior Software Engineer *(MSSQL, SSIS, SSRS, PLSQL)*   *Navi Mumbai, India*

- 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**
- Monitored model performance, response time and resource utilization, through **Looker dashboards**, **PyTest** test cases to validate functions, **DVC** for version controlling; Evaluated performance using key metrics like **MSE, RMSE, and R-squared**
- 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 ChatBot system with **Streamlit** UI, integrating AWS S3 data into a **Snowpipe**, designing Data Models via DBT
- Applied **LangChain, MiniLM** embeddings for LLM-based recommendations in **Snowpark** with **FAISS** for efficient text retrieval

S&P 500 Stock Analysis *(ML, Python)* *Jan 2024 to April 2024*
- Conducted Predictive Analysis of S&P 500 stock values by collecting historical stock values from S&P associated firms
- Deployed 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 *Jul 2025*
Master of Science in Data Analytics Engineering
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