

Dinakar Kolli

Data Engineer

Mechanicsburg, PA, USA | +1 (717) 608 -1194 | dinakarkolli0824@gmail.com | [LinkedIn](#) | [GitHub](#)

Summary

- Data Engineer with 3.5 years of experience in data modeling, ETL processes and data warehouse design to drive data driven decision making.
- Proficient in Microsoft tools for building scalable and secure data solutions.
- Skilled in Python programming for data processing, automation and statistical analysis ensuring seamless data transformation and Integration.
- Experienced in implementing data governance models and compliance frameworks to manage sensitive data securely and effectively.
- Expertise in creating interactive dashboards and data visualizations to uncover actionable insights and support strategic goals.
- Proven ability to optimize data workflows, reducing latency and improving system performance by up to 30%.
- Collaborative team player with strong stakeholder management skills, capable of aligning technical solutions with business objectives.
- Passionate about leveraging analytics to uncover hidden patterns and enable smarter decision making for clients in public and private sectors.

Technical Skills

- Programming & Query Languages: Python, PySpark, C, K-SQL, DAX, HTML, CSS
- Databases & Cloud: MySQL, Azure Data Factory, Azure Synapse, Azure Data Lake
- Data Processing & ETL: ADF Pipelines, Data Modeling, Data Migration
- Visualization & Reporting: Power BI, Excel (Pivot, Power Query, Power Pivot)
- Data Management: Delta Lake, Databricks, Azure Storage Explorer
- Version Control & Collaboration: Git, Azure DevOps

Professional Experience

Graduate Assistant, Shippensburg University, USA

Jan 2025 – Dec 2025

- Assisted faculty in delivering undergraduate level courses by conducting lab sessions, tutorials and grading assignments for classes of 30+ students.
- Provided academic support through one to one guidance, helping clarify concepts and improve performance in subjects like Python, Machine Learning.
- Developed supplementary teaching materials including practice problems and slide decks to enhance student understanding and engagement.
- Facilitated group discussions and problem solving sessions, improving class participation and helping students achieve a 15-20% increase over all course performance.

Data Engineer, Tata Consultancy Services, India

May 2021 - Aug 2024

- Designed and implemented robust data models and warehouse structures to enable efficient data storage and retrieval in a 20% improvement in query performance.
- Automated end to end ETL pipelines using Azure Data Factory, integrating multiple data sources and ensuring consistent data quality across systems.
- Build dynamic and interactive dashboards using Power BI to visualize KPIs, enhancing decision-making real-time analytics capabilities.
- Implemented data governance frameworks to ensure compliance with organizational and industry specific regulations, minimizing risk exposure.
- Conducted root cause analysis of performance bottlenecks in data workflows, achieving a 30% improvement in processing efficiency through query optimization and parallel execution techniques.
- Collaborated with business teams to identify core requirements and designed scalable analytics solutions to support growth initiatives.
- Spearheaded knowledge sharing sessions to train team members on advanced ETL processes and Power BI features.

ML Engineer Intern, Cognibot, India

Apr 2020 – Jul 2020

- Worked on a machine learning project involving 10,000+ data points, using Pandas and NumPy for data cleaning, transformation, and feature engineering.
- Implemented and evaluated multiple ML algorithms (e.g., Linear Regression, Decision Tree, KNN), achieving up to 90% accuracy on test datasets.
- Improved model performance by 20-30% through hyperparameter tuning and cross-validation techniques using Scikit-learn.
- Used Matplotlib and Seaborn to create 15+ visualizations for EDA and model insights, supporting data driven decision making.

Projects

Azure Data Engineering Pipeline (Azure Data Factory, Databricks, SQL, PowerBI)

- Build an end to end Azure data pipeline using Azure Data Factory, Databricks (PySpark) and ADLS Gen2 following the Medallion architecture (bronze – silver – gold).
- Developed scalable PySpark transformations and implemented data validation, incremental loads and Delta Lake optimizations for reliable, analytics ready data.
- Delivered curated datasets for reporting and integrated the pipeline with Power BI to support automated, accurate business insights.

Custom SVM Model for Weather Classification [Python, Pandas, Numpy, Matplotlib]

- Implemented a Support Vector Machine (SVM) classifier completely from scratch (without scikit-learn) to predict weather conditions (Rain vs Sunny) using temperature and humidity features.
- Developed and coded multiple kernel functions - Linear, Polynomial, RBF and Sigmoid evaluated their effect on decision boundaries, support vectors, and model accuracy.
- Performed detailed mathematical calculations for hyperplanes, kernel transformations, coefficients, and decision functions to demonstrate a deep understanding of SVM theory and optimization.

Certifications

- Microsoft Certified: Azure Data Fundamentals (DP-900)
- Microsoft Certified: Power BI Data Analyst Associate (PL-300)
- Microsoft Certified: Azure AI Fundamentals (AI-900)

Education

Master of Science in Computer Science

GPA:3.71 - Dec 25

Shippensburg University, Shippensburg, Pennsylvania, USA

Bachelor of Technology in Computer Science and Engineering

GPA:3.73 - Jul 21

Acharya Nagarjuna University, Andhra Pradesh, India

Achievements

- Served as Head of the Graduate Appeals Committee, leading case reviews, ensuring policy compliance and facilitating fair resolutions for student appeals.
- Received the GEMS Award for out standing performance and commitment to organizational excellence.