

DEEPAK DULAL

Tallahassee, FL | (850) 300-9050 | ddmelamchi@gmail.com | github.com/deepak-dulal | linkedin.com/in/deepak-dulal

PROFESSIONAL SUMMARY

Results-driven Data Engineer skilled in building scalable ETL pipelines and automating workflows with Airflow and Pentaho. Proficient in SQL for data extraction, transformation, and optimization. Experienced in developing data architectures and collaborating with teams to deliver reliable solutions that improve efficiency and support data-driven decisions.

EDUCATION

Master of Science in Data Science, GPA: 3.93/4

08/23 - 05/25

Florida State University

Tallahassee, FL

Courses: Advanced Data Mining, Applied Machine Learning, Applied Regression Modeling, Mathematics for Data Science, Deep and Reinforcement Learning, Advanced Data Science, Data Ethics

SKILLS

Programming Languages	Python, SQL, R
Big Data & ETL Tools	PySpark, Delta Lake, Databricks, Hadoop, Airflow, DBT
Cloud Platforms	AWS (S3, Lambda, Glue), Azure (Data Factory, Data Lake)
Databases & Warehouses	PostgreSQL, Snowflake
Other Tools & Frameworks	Git, CI/CD, Docker, Kafka, Terraform, Power BI, Tableau

EXPERIENCE

Junior Data Analyst

05/22 - 06/23

Greative Soft Private Limited

Kathmandu, Nepal

- Conducted **business analysis** and implemented ETL pipelines to process financial transactions from multiple sources, analyzing salary patterns and financial behaviors to assess creditworthiness for **200K+ customers**.
- Built BI dashboards and data visualizations for financial datasets, ensuring standardized data models to deliver actionable insights to stakeholders, facilitating data-driven decision-making and boosting marketing **ROI by 20%**.
- Collaborated with cross-functional teams** to define reporting needs and translated them into technical solutions by using **SQL** to extract, clean, and transform data, ensuring high quality and delivering actionable insights
- Optimized **SQL queries** and **automated workflows** using **Airflow and Pentaho**, reducing processing times by **25%** and improving operational efficiency.

PROJECTS

Sales Performance and Profitability Insights Using Power BI

06/24 - 11/24

- Developed **2 interactive dashboards** analyzing **500K+ sales records**, identifying revenue trends and profitability drivers.
- Implemented **DAX measures** for profit margin, YoY growth, and KPI tracking, enhancing reporting clarity.
- Optimized **ETL pipeline** using Power Query, reducing data processing time by **30%** and improving data accuracy.

Enhanced Answer Retrieval System Powered by OpenAI LLM and RAG [GitHub](#)

06/24 - 08/24

- Integrated **OpenAI LLMs** with **RAG techniques** to enhance answer relevance and contextual understanding, increasing retrieval accuracy by **40%**.
- Built an **automated data pipeline** combining **web scraping and NLP**, improving large-scale data extraction, preprocessing, and scalability.

Data-Driven Diamond Price Estimation Using R [GitHub](#)

01/24 - 05/24

- Built a regression model in R to estimate **diamond prices**, achieving an R^2 of **95.39%**, with feature selection (AIC, BIC) improving accuracy.
- Performed **data cleaning** and addressed multicollinearity using **Weighted Least Squares (WLS)** and log transformations, reducing LOOCV-RMSE to **0.134**.

CERTIFICATIONS

Machine Learning Specialization: Stanford University

05/24 - 09/24

Ask Questions to Make Data-Driven Decisions: Google

03/23 - 04/23

Excel Skills for Business: Essentials: Macquarie University

11/22 - 12/22

INVOLVEMENT/AWARDS

Scholarship: Received **Full Ride Scholarship** from the Indian Embassy for undergraduate studies at VIT University.

Initiatives: **Organizer** of Melamchi Food Bank, a temporary community food donation hub launched during COVID-19

Leadership: Served as **Treasurer** of the Mithial Social Club, overseeing finances, budgeting, and fundraising.