

Utsabi Dangol

utsabidangol123@gmail.com | Vermillion, SD (Open to Relocation) | 605-202-1708 | [Linkedin](#) | [Portfolio](#)

TECHNICAL SKILLS

- **Programming & Scripting:** Python (Pandas, PySpark, NumPy), SQL, JavaScript
- **Big Data & Stream Processing:** Apache Spark, PySpark, Databricks, Apache Kafka, Hadoop, Apache Airflow, dbt
- **Cloud & Data Platforms:** GCP, AWS, Azure, Snowflake, Google BigQuery, Redshift
- **Data Engineering & Warehousing:** Data Modeling, Schema Design, Partitioning, Incremental Processing
- **Analytics & BI:** Looker, Tableau, Power BI
- **Machine Learning & AI:** PyTorch, LLMs, RAG (Retrieval-Augmented Generation), Feature Engineering
- **DevOps & CI/CD:** Git, GitHub Actions/GitLab CI, Docker, Kubernetes, Automated Testing (pytest)

WORK EXPERIENCE

Logic US

Software Engineer

Sep 2021 – Aug 2024
Minnesota, US (Remote)

- Owned end-to-end design and scaling of distributed ELT pipelines ingesting data from 10+ sources.
- Led a team of 9 engineers, driving architecture decisions, code quality standards, and best practices.
- Reduced pipeline runtime by 40% and compute costs by 30% by implementing partitioning and workload strategies.
- Designed scalable data warehouse schemas and data marts supporting 50+ dashboards used for decision-making.
- Improved data reliability by 35% by building automated validation, monitoring, and alerting frameworks.
- Implemented RBAC and PII masking, ensuring secure and compliant access to sensitive datasets across teams.

Impacters

Junior Software Engineer

Apr 2021 – Sep 2021
Kathmandu, Nepal

- Built an ML system for real-time crop health monitoring, improving decision-making using CNNs and transfer learning.
- Designed data processing pipelines for sensor data, enabling actionable insights for farmers.
- Deployed ML models on resource-constrained, solar-powered devices, ensuring reliable offline inference.

Asterdio Inc.

Associate Software Engineer

Jul 2020 – Mar 2021
Kathmandu, Nepal

- Improved customer targeting and product insights by developing clustering algorithms on behavioral patterns.
- Reduced query latency by 40% by optimizing backend data workflows and database queries in Node.js and MongoDB.
- Built internal data ingestion and transformation pipelines enabling scalable product analytics.

EDUCATION

MS, Computer Science, University of South Dakota

GPA: 4.0 Relevant coursework: Reinforcement Learning, Agentic AI, Data Analysis, Information Storage & Retrieval

Aug 2024 – May 2026

BS, Computer Science, Tribhuvan University

Recipient of Fr. Eugene L. Wartin S.J. Memorial Outstanding Student Award

Frs. Locke and Stiller Research Award for thesis on crime analysis using KNN and risk matrix

Nov 2016 – Jun 2020

ACADEMIC PROJECTS

LLM-Based PDF Question Answering using RAG

Python, LangChain, IBM Watsonx, ChromaDB, Gradio

Aug 2025 – Dec 2025

- Engineered a RAG pipeline using Watsonx embeddings and ChromaDB to enable semantic search over PDFs.
- Deployed a Gradio interface for live document ingestion and top k similarity querying, reducing hallucinations.

E-Commerce Analytics: Product Success and Price Prediction

(Python, FAISS, HuggingFace, FastAPI, SHAP, SentenceTransformers)

Aug 2025 – Dec 2025

- Built predictive models for product ratings and launch risk, achieving 82.98% classification accuracy.
- Implemented RAG-based price prediction pipeline with dense, sparse, and hybrid retrieval using FAISS and BM25.

PUBLICATIONS

Uncertainty-Aware LLM-Guided Policy Shaping for Sparse-Reward Reinforcement Learning

Accepted — IEEE CAI 2026

Assessment of Criminal Activities in Kathmandu Valley using Risk Matrix and KNN Classifiers

Mini Thesis | ISSN: 2705-456X

LEADERSHIP & AWARDS

- **NSF I-Corps Entrepreneurial Lead (2025):** Conducted 40+ customer discovery interviews to validate AI product ideas, including LLM-based legal assistant and music transcription system.
- **Professional Director, Rotaract Club of NewRoad City:** Led programs empowering members with professional skills and career development.