Dipen Kumar Maheshwari

+1-780-708-5514 | dipenkum@ualberta.ca | linkedin.com/in/dipen-kumar | github.com/dipenkumarr | Portfolio Website

EDUCATION

University of Alberta Edmonton, AB
Bachelor of Science Honors in Computer Science Expected: December 2026

Relevant Coursework: Software Engineering, Data Structures & Algorithms, Relational Database Management, Machine Learning, Artificial Intelligence

EXPERIENCE

Machine Learning Engineer Intern Jan 2025 – Present
SAP Vancouver, BC
- Engineered an end-to-end data processing and unsupervised **NLP** pipeline on **Azure Databricks** (PySpark) to process ~**38k** raw tickets reducing message volume by ~**85%** via **FAISS**-optimized semantic de-duplication.
- Implemented comprehensive data quality workflows using **NLTK/Spacy**, reducing noisy data by **15%** and significantly improving data consistency for model training.
- Leveraging **LLMs (GPT-4o/Gemini)** with structured Pydantic outputs to evaluate and categorize tickets, projecting for nearly **60%** in reduction in manual triage time.
- Prototyping an advanced **RAG pipeline** using **DSPy and ColBERT** to enhance an enterprise system, exploring late-interaction retrieval to generate highly accurate summaries from large-scale partner documentation.

Software and Research Intern May 2024 – June 2024
University of Alberta - ANCL Edmonton, AB
- Led the **containerization** of autonomous navigation systems using **Docker**, reducing development environment setup time by **50%** and enhancing system reliability.
- Refactored autonomous navigation algorithms in **ROS, Linux, C++**, and **Python**, boosting readability and reducing runtime errors.
- Implemented **Git** and **GitHub** version control practices, including branch protection and pull request reviews, reducing merge conflicts.

PROJECTS

SpotLotto - Event Lottery System | *Java, Android Studio, Firebase, JUnit, Git, GitHub, XML* | **GitHub**
- Built an **Android** app using **Java** and **Firebase** to automate event registrations via lottery system, integrating Google ML Kit for **QR code scanning** and geolocation features.
- Led a **6-member Agile team**, to design a **multi-role user** system and implement Firebase Cloud Messaging for **real-time push notifications**, achieving 80% test coverage with **JUnit** through test-driven development.

UniWays - Google Hacks Winner | *Next.js, React, TypeScript, Prisma, PostgreSQL, Google Maps APIs* | **GitHub**
- Developed a **full-stack** campus navigation web application using **Next.js**, and **TypeScript**, achieving a **20% reduction** in estimated student travel time through **Google Maps API** integration and real-time route optimization.
- Designed **REST APIs** with **Prisma** and **PostgreSQL**, incorporating **OAuth** for robust user authentication and enhanced data protection.

Bloom.ai | *Next.js, React, TypeScript, NextAuth, TailwindCSS, MongoDB, Zod, Gemini-AI API* | **GitHub** | **Live**
- Developed an **AI-powered** anonymous full-stack feedback platform, reducing message composition time by **40%** through **Gemini-AI LLM** integration.
- Implemented a secure backend and authentication system using **MongoDB**, **NextAuth**, and **Zod** schema validation for enhanced data integrity and security.

Classification on NotMNIST-RGB and NotMNIST DL using CNN | *Python, PyTorch* | **GitHub**
- Engineered a convolutional **neural network deep learning model** with modified **LeNet architecture**, achieving **97.5%** accuracy in image classification; optimized **object detection** on NotMNIST-DL dataset to **82.8%**.

TECHNICAL SKILLS

Languages: Python, JavaScript, TypeScript, Java, Go, C/C++, SQL, NoSQL, HTML, CSS
Frameworks/Libraries: Next.js, React.js, Node.js, Express.js, Tailwind CSS, LangChain, Numpy, Pandas, PyTorch
Developer Tools: Git, GitHub, VS Code, Android, Linux/Unix, Docker, Azure Databricks, Jira, Confluence
Databases: MongoDB, MySQL, SQLite, Firebase, PostgreSQL