

DEVANSH MISTRY

35 King St W, Kitchener, ON N2G 1A1
+1 647-462-3187 • devanshmistry890@gmail.com
[Portfolio](#) • [linkedin](#) • [Github](#)

PROFESSIONAL SUMMARY

AI Application Engineer focused on building high-value, production-ready AI systems that drive customer retention and business efficiency. Specialized in integrating Voice Agents and RAG pipelines into modern web environments. Strong ability to bridge the gap between technical logic and product design, leveraging high-end visualization skills to create engaging, polished demos. Experienced in identifying high-impact AI use cases and delivering privacy-focused, browser-based solutions.

TECHNICAL SKILLS

AI & ML: PyTorch, Edge AI (WebLLM, Transformers.js), RAG Pipelines, Azure OpenAI, Document Intelligence, Computer Vision, Prompt/Context Engineering.

Web Integration & APIs: React.js, Next.js, REST APIs, GraphQL, MERN Stack, Authentication/Security (JWT).

Languages & Tools: Python, SQL, MongoDB, Vector Databases (Pinecone/Chroma), Docker, Git/GitHub Actions, Adobe Creative Suite.

FEATURED PROJECTS

Real-Time AI Voice Agent | Tech: Python, Gemini AI, ElevenLabs, Twilio/WebSocket

- Built a <150ms latency voice agent capable of handling real-time customer support conversations.
- Engineered the orchestration layer to handle Voice-Activity-Detection (VAD), interruption handling, and TTS/STT synchronization.
- Deployed a custom WebSocket server to stream audio chunks bidirectionally, ensuring human-like responsiveness.

Smart Recipe Retrieval (RAG System) | Tech: Python, VectorDB, Gemini API

- Built a RAG pipeline using MongoDB Atlas, Google Gemini LLM and embeddings APIs to generate tailored recipe recommendations.
- Implemented vector search, semantic retrieval, and prompt optimization to improve contextual accuracy in recipe recommendations.

Edge AI NLP Analyzer | Tech: React, Transformers.js, WebGPU

- Developed a privacy-centric text analysis tool that runs BERT-sized models entirely in the user's browser (Client-side).
- Implemented WebGPU acceleration to achieve near-native inference speeds for sentiment analysis and entity extraction without server costs.
- Eliminated data privacy risks by processing 100% of user data locally, removing the need for cloud API calls.

PROFESSIONAL EXPERIENCE

Data Scientist (Intern) (Remote)

iNeuron.ai.

May 2022 – Feb 2023

- Deployed predictive machine learning models on Azure ML, reducing manual analysis time for internal teams by automating data insights.
- Optimized data preprocessing pipelines, reducing feature engineering latency by ~20% and improving model inference accuracy.
- Containerized models using Flask + Docker, automating deployment and scaling across environments.
- Designed interactive dashboards to visualize performance metrics and streamline insights for decision-makers.

Machine Learning Intern (Full-time)

F(x) Data Labs, Ahmedabad, India.

Jan 2023 – Feb 2023

- Researched and reported on AI applications in aviation, synthesizing academic + industry data into actionable insights.
- Prototyped NLP and predictive models, validating feasibility for real-world aviation operations.
- Translated complex ML outputs into actionable business strategies, directly supporting senior leadership in evaluating AI adoption in aviation.

Content & Analytics Assistant (Part-time)

Conestoga College Esports, Waterloo, ON.

Aug 2024 – Apr 2025

- Produced analytics-driven engagement reports using performance data to guide strategy decisions.

PROFESSIONAL DEVELOPMENT AND TRAINING

Post-Graduate Certificate, Web Development | Conestoga College, Waterloo, ON

Bachelor of Engineering, Information Technology | L.D. College of Engineering, India

Relevant Coursework:

- Data Science & Artificial Intelligence
- Pattern Recognition & Probability and Statistics
- Database Management Systems, Cryptography and Network Security, DSA

Deep Learning & Neural Networks | Independent Study (Andrej Karpathy Curriculum)

Microsoft Certified: Azure AI Engineer Associate (AI-102) | Status: Active

Microsoft Certified: Azure Data Scientist Associate (DP-100) | Status: Active

Microsoft Certified: Azure AI Fundamentals (AI-900) | Status: Active

IBM Deep Learning and Reinforcement Learning | Status: Active

IBM Machine Learning Specialization | Status: Active