M S Gulshan

Ingénieur logiciel principal en IA

Bengaluru, Inde | +91 74833 15810 | gulshanms.sde@gmail.com | gulshanms.dev | linkedin.com/in/gulshanms | github.com/gulshanmsgit

Profil

Ingénieur en IA qui dirige l'équipe IA de Vrtta Green Solutions : traitement de documents par LLM et recherche pour les rapports d'empreinte carbone, en lien direct avec les clients. En dehors du travail, j'ai créé et publié AnswerGate, un outil libre de test pour assistants IA, sur PyPI, et Exsy, une application de gestion de l'argent.

Expérience

Vrtta Green Solutions — Montréal, Canada (à distance)
Ingénieur logiciel principal en IA (promu après le poste de développeur logiciel en IA) — Sept. 2025 – aujourd'hui

- Je dirige la planification et le développement de l'équipe IA, et gère les projets clients jusqu'à la livraison.
- J'ai construit le pipeline de l'outil d'empreinte carbone des produits de l'entreprise : il lit des nomenclatures de plusieurs pages et produit les rapports d'émissions qui étaient auparavant préparés à la main.
- J'ai affiné des LLM libres (LoRA, QLoRA) pour extraire et classer les matériaux, avec 98,7 % de précision de lecture.
- J'ai construit une recherche hybride dans Qdrant (vecteurs et filtres de métadonnées) qui associe chaque matériau à plus d'un million d'entrées d'ACV.
- J'ai réduit la latence de recherche de 40 % grâce aux budgets de jetons et à la diffusion NDJSON/SSE depuis le serveur Flask.

Mu Sigma — Bengaluru, Inde
Scientifique décisionnel stagiaire — Juin 2025 – sept. 2025

- J'ai construit des pipelines de données et des modèles d'analyse pour de grandes entreprises, et présenté les résultats.

EY Global (par Edunet) — À distance
Stagiaire développeur MERN — Févr. 2025 – mars 2025

- J'ai développé des applications web MERN servant plus de 100 utilisateurs simultanés.

Projets

AnswerGate | Python, Ollama, Qdrant, GitHub Actions | github.com/gulshanmsgit/answergate — 2026

- Outil libre en ligne de commande et GitHub Action, publié sur PyPI, qui pose des questions de test à un assistant IA et fait échouer la demande de fusion quand les réponses se dégradent.
- Note les réponses avec un LLM local via Ollama et vérifie les fuites de renseignements personnels, l'injection de requêtes et les réponses hors sujet.
- A détecté 36 des 42 défauts introduits sur 1 070 questions de test, sans fausse alerte, et signalé un changement de taille des segments qui a réduit l'exactitude d'un assistant FastAPI + Qdrant.

Exsy – carnet d'argent familial | JavaScript, Firebase, application web installable | github.com/gulshanmsgit/Exsy- — 2026

- Application web installable pour suivre l'argent d'une famille : un carnet par personne, comptes d'argent comptant, de banque et de carte, prêts et mensualités avec taxe.
- Gère une tontine (chit fund) : collectes mensuelles, liste des retards, tirage au sort parmi les membres qui n'ont pas encore gagné, et relevés PDF partagés sur WhatsApp.
- Remplit une écriture à partir d'un texto bancaire ou d'un message UPI collé; fonctionne hors ligne, avec connexion Google et synchronisation privée par Firestore.

Feux de circulation adaptatifs, avec la police de la circulation de Bangalore | Python, LSTM, SQL — 2024 – 2025

- J'ai entraîné un LSTM sur les données de capteurs de circulation en direct pour ajuster la durée des feux; la congestion a baissé de 30 % en simulation.

Compétences

IA et LLM : affinage de LLM (LoRA, QLoRA), Ollama, RAG, Qdrant, ingénierie de requêtes, évaluation de LLM
Serveur : Python, Flask, FastAPI, SQL, API REST, pipelines asynchrones, server-sent events
Interface et outils : React, JavaScript, Firebase, MongoDB, Git, GitHub Actions, pytest

Formation

Global Academy of Technology — Bengaluru, Inde
Baccalauréat en génie (B.E.) en sciences et génie de l'information, moyenne 9,73/10 — 2021 – 2025

Distinctions

- Médaille d'or, promotion 2025, Global Academy of Technology
- 3e rang national, International Science Talent Examination, 2019