

AI & GenAI Solution Architect

Location: Delhi/NCR

About the Role:

We are seeking a highly motivated and experienced **AI Solution Architect** to join our growing team. In this critical role, you will play a pivotal role in designing and delivering innovative AI solutions for our customers. You will collaborate with key stakeholders including CXOs to understand their challenges and craft comprehensive solutions that leverage the power of AI/ML and Generative AI.

Responsibilities:

- Partner with customers to deeply understand their business needs and pain points related to AI/ML and Generative AI.
- Design robust and secure solution architectures, encompassing conceptual, technical, and security considerations. This includes high-level and low-level designs, implementation roadmaps, and integration plans with existing workflows.
- Develop compelling solution proposals, including high-level approach and architecture diagrams with clear narratives that explain key design choices, such as:
 - Selection of appropriate AI/ML models (e.g., supervised learning, unsupervised learning, deep learning) based on data characteristics and problem requirements.
 - Strategies for data pre-processing, feature engineering, and model training.
 - Considerations for model explainability, bias mitigation, and fairness.
- Collaborate with engineering teams to create proof-of-concept (POC) demonstrations that validate the proposed solutions, considering factors like:
- Identify and leverage existing AI/ML and Generative AI accelerators and solutions to enhance presales activities, client demos, and POC development.
- Continuously stay updated on the latest advancements in Generative AI through ongoing learning and training initiatives.

Qualifications:

- Bachelor's or master's degree in computer science, Engineering, or a related field.
- 7-10 years of experience in designing and developing AI/ML and Generative AI solutions.
- Proven track record of successfully translating customer requirements into effective AI-powered solutions.

- Strong understanding of machine learning principles, frameworks (e.g., pandas, TensorFlow, PyTorch), and best practices.
- Expertise in cloud computing platforms (Azure, AWS, GCP) for AI/ML deployments.
- In-depth knowledge of Generative AI models and services like Azure OpenAI, Google AI, AWS, Anthropic, Nvidia NeMo, and Microsoft Copilot etc., plus open source LLMs
- Experience with natural language processing (NLP) techniques for tasks like text classification, sentiment analysis, and machine translation.
- Understanding of computer vision (CV) concepts and frameworks for tasks like object detection, image recognition, and image segmentation; and reinforcement learning (RL).
- Familiarity with designing applications with Nvidia CUDA, and familiarity with NVIDIA NeMo framework for GenAI application development.
- Knowledge of model deployment strategies, including containerization and orchestration tools (e.g., Docker, Kubernetes).

Application Process: Please write to contact@cellstrat.com with your resume, the position you're applying for and your name in the subject line of the email.