**Job Title: Senior AI Engineer (Generative AI & MLOps)**

**Company:** Innovatech Solutions

**Location:** Bengaluru, India (Hybrid)

**About the Role:** Innovatech Solutions is seeking an experienced and highly motivated Senior AI Engineer to join our dynamic research and development team. In this role, you will be responsible for designing, building, and deploying cutting-edge generative AI models and robust MLOps pipelines. You will work on challenging problems that push the boundaries of natural language processing and computer vision, and your work will directly impact our next generation of intelligent products.

**Key Responsibilities:**

* Design, develop, and train advanced deep learning models, with a focus on Transformer architectures and Large Language Models (LLMs).
* Fine-tune pre-trained models such as BERT and GPT for specific downstream tasks like text summarization, question-answering, and sentiment analysis.
* Build and maintain scalable MLOps pipelines for continuous integration, continuous delivery (CI/CD), and automated model retraining.
* Utilize cloud platforms like **AWS** or **Azure** to deploy and monitor machine learning models in a production environment.
* Collaborate closely with data scientists, software engineers, and product managers to integrate AI capabilities into our core applications.
* Optimize model performance for latency and throughput, and ensure solutions are cost-effective.
* Stay current with the latest advancements in AI/ML research and apply them to solve real-world business problems.

**Required Qualifications & Skills:**

* Bachelor's or Master's degree in Computer Science, AI, or a related field.
* 5+ years of hands-on experience in machine learning and software development.
* Expert-level proficiency in **Python** and its data science ecosystem (**Pandas**, **NumPy**).
* Deep hands-on experience with at least one major deep learning framework: **PyTorch** or **TensorFlow**.
* Proven experience with the **Hugging Face** library for NLP tasks.
* Strong knowledge of MLOps principles and tools such as **Docker**, **Kubernetes**, and **Git**.
* Experience deploying models on cloud services (e.g., **AWS SageMaker**, **Azure ML**, **GCP Vertex AI**).
* Excellent **problem-solving**, **analytical thinking**, and **communication** skills.

**Preferred Qualifications:**

* Experience with distributed training and data processing using tools like **Apache Spark**.
* Familiarity with vector databases such as Pinecone or Milvus.
* Published research in top-tier AI conferences (e.g., NeurIPS, ICML, CVPR).
* Experience with **FastAPI** for creating model APIs.