**Position:** ML Engineer/Generative AI Engineer

**Overview:**
The core focus of the ML/Generative AI engineer role is to develop cutting-edge generative models and applications. Candidate needs to have a strong background in machine learning and deep learning along with experience in evaluating, building, and deploying generative AI models.

**Job Location:**
Pune/ New Delhi

**Duties and Responsibilities:**
- Design, develop, and implement state-of-the-art generative models for various applications, such as image synthesis, text generation, and content creation.
- Experience working on open source LLMs like Llama, Falcon, BERT etc.
- Pre-process and analyze large datasets to extract meaningful features and patterns suitable for training generative models.
- Optimize and fine-tune generative models for efficiency, scalability, and improved performance.
- Implement strategies for reducing model latency and resource consumption.
- Develop and implement robust evaluation metrics to assess the performance and quality of generative models.
- Document code, models, and methodologies thoroughly to facilitate knowledge sharing and future development.
- Stay updated with the latest advancements in generative models, including GANs (Generative Adversarial Networks), VAEs (Variational Autoencoders), and other relevant techniques.
- Stay at the forefront of the latest research papers and publications in the field of generative AI.
- Contribute to the company's intellectual property by researching and developing novel approaches to address specific business challenges.
- Work closely with cross-functional teams, including data scientists, software engineers, and product managers, to integrate generative models into real-world applications.
- Provide technical expertise and guidance to team members on generative AI-related projects.

**Requirements:**

- B Tech/M Tech in Computer/IT or related field from reputed engineering colleges
- 3 - 6 years of hands-on experience in ML domain.
- Proven experience in developing **generative models**, with expertise in deep learning frameworks such as TensorFlow or PyTorch.
- Strong programming skills in languages like Python and proficiency in machine learning libraries.
- Solid understanding of neural network architectures, optimization techniques, and training methodologies
- Work experience on cloud platforms like AWS/Azure/GCP is desired.

*InvoLead is an equal opportunities employer and welcomes applications from all sections of society and does not discriminate on grounds of race, religion, or belief, ethnic or national origin, disability, age, citizenship, marital, domestic or civil partnership status, sexual orientation, gender identity, or any other basis as protected by applicable law.*