Job Description:

As a System Architect at Beroe, you will be responsible for designing, implementing, and optimizing AI systems that leverage Generative AI and fine-tuning techniques. You will work closely with cross functional teams to create scalable and efficient AI solutions that meet our business objectives. Your expertise in various AI models and fine-tuning methodologies will be critical to our success.

Responsibilities:

1. Develop and document the architecture of AI systems with Generative AI and fine-tuning capabilities.

2. Design, implement, and fine-tune AI models (e.g., GPT, LLaMA, BERT).

3. Lead adoption of best practices, design patterns, and new technologies in AI.

4. Ensure seamless integration of AI models with existing systems, maintaining performance, scalability, and security.

5. Collaborate with project managers on scope, timelines, and deliverables.

6. Provide technical leadership throughout the project lifecycle.

7. Conduct code reviews and ensure quality and integrity of the AI codebase.

8. Mentor and guide junior AI engineers.

9. Foster a culture of continuous learning and improvement.

10. Work with data scientists, engineers, product managers, and other stakeholders to ensure technical feasibility and optimal performance of AI solutions.

Requirements:

1. Bachelor's or Master's degree in Computer Science, Engineering, or related field.

2. Minimum of 10 years in AI architecture, model development, and fine-tuning.

3. Proficiency in Python.

4. Experience with Generative AI models (e.g., GPT, LLaMA, BERT).

5. Expertise in fine-tuning large language models.

6. Experience with machine learning frameworks (e.g., TensorFlow, PyTorch).

7. Knowledge of cloud platforms (e.g., AWS, Azure, GCP).

8. Strong analytical and problem-solving skills.

9. Excellent verbal and written communication skills.

10. Ability to work effectively in a collaborative, fast-paced environment.

Preferred qualifications:

1. Experience with MLOps practices and tools.
2. Familiarity with data engineering and preprocessing techniques.
3. Knowledge of other AI and ML techniques.
4. Previous leadership or mentorship experience.