**Director of Generative AI**

Capital One

**Full job description**

**Chief AI Officer**

**Job Summary**:

Capital One is searching for an enterprising Director of Generative AI capable of directing the development of the company's generative AI platform. If selected, the chosen candidate will lead a team of AI experts focused on optimizing training pipelines, deploying large ML models, and building scalable infrastructure using advanced cloud technologies. This role is critical to Capital One's mission of transforming the banking industry by utilizing the vast potential of cutting-edge generative AI.

**Job Duties and Responsibilities:**

* Take charge of a team comprising of AI engineers and professionals to establish the basis of the Generative AI platform.
* To design an effective distributed training cluster in the public cloud, one must consider the infrastructure requirements, networking considerations, and workload distribution strategies.
* A resilient infrastructure is necessary to ensure that training tasks that take a long time to complete are fault-tolerant.
* Developing a runtime system that executes complex machine learning models efficiently and cost-effectively.
* Our Gen AI platform offers great potential for building anomaly and fraud detection systems and information retrieval mechanisms. We aim to facilitate users in harnessing this potential by creating SDKs and APIs for the community.

**Qualifications and Experience**:

* A technical degree, such as a Bachelor's in Computer Science, Computer Engineering, or a related field, is required.
* For the past seven years, I have been actively engaged in creating and implementing large-scale ML systems, thus acquiring significant hands-on experience in this area.
* With six-plus years of expertise developing AI/ML algorithms in Python or C/C++, I have proven my ability to create cutting-edge solutions.
* A skilled manager possessing 5+ years of experience leading and directing engineering and applied science teams.
* With an in-depth understanding of the entire ML development life cycle, I have successfully leveraged my expertise in utilizing open-source AI/ML frameworks and public cloud solutions.
* Achieving a higher degree in Computer Science or a similar technical field can open up more opportunities for career advancement and enable you to become a subject matter expert.
* Skilled in using machine learning and with proven experience in the large-scale training and deployment of deep neural networks or transformer architectures.
* Browse high-quality research publications in reputable peer-reviewed conferences or recognized open-source contributions from industry.
* Skilled in working with popular machine learning frameworks such as TensorFlow, PyTorch, Lightning, and Mosaic ML, and in executing various machine learning tasks on high-performance GPU clusters.
* An aptitude for agile development in a fast-moving atmosphere with uncertainty.
* Competent and experienced in handling public cloud systems like AWS, Azure, or GCP.
* Knowledge in successfully setting up and using extensive neural network models for demanding production applications.

**Benefits of the Position:**

* Performance-based compensation, which encompasses various incentives like cash bonuses and long-term inducements.
* Employees can access a variety of benefits designed to ensure their overall well-being, such as health and financial provisions.
* Employees at Capital One have access to a variety of resources and programs designed to support career growth and development.
* A work setting that is supportive and inclusive, where employees can grow in both their personal and professional lives.

**About Company**:

Join the team at Capital One and become part of a world-class organization that is dedicated to driving groundbreaking products and scalable, high-performance generative AI infrastructure. We are transforming the banking industry and revolutionizing how we serve our customers and businesses.