Traccia 1

Il candidato discuta in maniera sintetica ed efficace (con non più di 2000 caratteri complessivi per domanda) DUE dei seguenti argomenti, indicando esplicitamente le proprie scelte all'inizio del proprio elaborato:

A) il candidato descriva una piattaforma di High Performance Computing specificatamente disegnata per fornire servizi basati su modelli open di LLM già addestrati. Si discuta in particolare la scelta delle componenti di calcolo e di storage e di network e le metriche di valutazione del sistema. Indichi inoltre le componenti software fondamentali per la messa in produzione del sistema.

B) Il candidato descriva la messa a punto di una pipeline per la gestione in parallelo di più fonti di dati provenienti da sensori. Ne discuta lo stack software che si intende utilizzare, la potenziale scalabilità ed eventuali punti critici per quel che riguarda la latenza e la banda passante. Infine si discutano metodi per garantire la riproducibilità.

C) Il candidato descriva uno o più casi d'uso specifici di algoritmi e strumenti di intelligenza artificiale e non solo, per la gestione di dataset eterogenei (es. integrazione di dati strutturati, immagini e testi), per la creazione di un dataset rappresentativo per il training di modelli di deep learning. Inoltre, ne analizzi lo stack software necessario.

D) L'azienda XY vuole realizzare un POC per migliorare il controllo qualità delle saldature tramite analisi automatica di immagini. Il candidato descriva come definire obiettivi e KPI del POC e del dataset rappresentativo.

Traccia 1

The candidate should discuss concisely and effectively (with no more than 2000 characters total per question) TWO of the following topics, clearly stating the chosen topics at the beginning of the response:

A) The candidate describes a High-Performance Computing platform specifically built to deliver services based on pre-trained open-model LLMs. Discuss,in particular, the selection of compute, storage, and networking components, as well as the metrics used to evaluate the system. Also identify the core software components required to put the system into production.

B) The candidate describes the development of a pipeline for parallel data streams d originating from sensors. Discuss the software stack to be employed, the potential scalability, and any critical points concerning latency and bandwidth. Finally, discuss methods to ensure reproducibility.

C) The candidate outlines one or more specific use cases of artificial-intelligence algorithms and any other tools for managing heterogeneous datasets (e.g., integration of structured data, images, and text) in order to create a representative dataset for training deep-learning models. Additionally, analyze the required software stack.

D) Company XY wants to develop a PoC to improve weld-quality control through automatic image analysis. The candidate explains how to define the PoC's objectives and KPIs, as well as how to build a representative dataset.