

AI/ML Planning Workshop

Product Code: CNS-DAY0-WRK-PLN-GPT

At-a-Glance

Phase: Plan

The Nutanix Artificial Intelligence / Machine Learning (AI/ML) Planning Workshop offers IT teams in-depth and practical guidance to plan and implement infrastructure for AI/ML workloads based on the Generative Pre-trained Transformer (GPT) model for inference workloads running on on-premises Nutanix Cloud Infrastructure (NCI) clusters. This workshop is beneficial during the Planning stage of the Nutanix GPT-in-a-Box Solution journey.

Service Scope

Delivered by highly skilled consultants with strong AI/ML domain expertise and rich experience, the planning workshop begins with a discovery session to understand the operating environment and required outcomes. Critical areas identified during Discovery are then addressed in a customized workshop.

The service includes the following activities:

- Gather and discuss customer AI/ML use case, requirements, and expectations
- Assess and summarize the current state of AI/ML
- Review capacity planning, and scalability considerations for the number of end users who will leverage the GPT-based application
- Identify data management considerations, including the data source, data preparation for AI/ML use, data protection, and security
- Develop a risk management plan
- Learn the importance of choosing the right large language model (LLM) from a list of validated LLMs on Nutanix Cloud Infrastructure (NCI)
- Identify GPU selection and configuration options for inference
- Explore options for training the model, such as using virtual machines (VMs) or container services

Limitations

- Limited to infrastructure planning for AI/ML inference workloads.
- Design and deployment of AI/ML infrastructure workloads are excluded from the workshop
- The planning workshop is limited to one AI/ML use case

Supported Hypervisors

- Nutanix AHV

Prerequisites

- None

Required Product Licenses

- None

Deliverables

- Project Kickoff
- Project Schedule
- Project Status Report(s)
- Workshop
- Optimized Customer Strategy
- High-level Summary Presentation
- Project Closeout

Duration

Typically up to 1 day

Related Products

- Nutanix Cloud Infrastructure (NCI)
- Nutanix Unified Storage (NUS)
- Nutanix Cloud Manager (NCM)

Terms and Conditions

This document contains the entire scope of the service offer. Anything not explicitly included above is out of scope. This service offer is subject to the Nutanix Services General Terms and Conditions, which can be viewed at <https://www.nutanix.com/support-services/consulting-services/terms-and-conditions>

©2024 Nutanix, Inc. All rights reserved. Nutanix, the Nutanix logo, and all Nutanix product and service names mentioned herein are registered trademarks or trademarks of Nutanix, Inc. in the United States and other countries. Nutanix, Inc. is not affiliated with VMware by Broadcom or Broadcom. VMware and the various VMware product names recited herein are registered or unregistered trademarks of Broadcom in the United States and/or other countries. All other brand names mentioned herein are for identification purposes only and may be the trademarks of their respective holder(s).