Pranjli Pandya

Homework 1

MSIS 549

Question 1

Training vs. Inference in ML Models:

> Training: This is the phase where the machine learning model learns from the data. You feed it examples, and it adjusts its internal parameters to minimize errors based on the examples it sees.
>
> Inference: After the model is trained, inference is the stage where the model uses what it has learned to make predictions on new, unseen data.

Transfer Learning:

> Transfer learning is a technique where a model developed for one task is reused as the starting point for a model on a second task. It's important because it can save time and resources, and improve performance, especially when you don't have a lot of data for the new task.

Examples of Computer Vision Tasks:

> Image Classification: Identifying the category of an object in an image.
>
> Object Detection: Locating objects within an image and identifying their boundaries.
>
> Facial Recognition: Identifying or verifying a person's identity from their face.

Question 2

The Deloitte report's observation that "Most organizations are still primarily relying on off-the-shelf generative AI solutions" is intriguing because it reflects a cautious initial approach by companies integrating AI into their operations. This strategy allows businesses to gain immediate efficiencies and build AI familiarity without significant investment in custom solutions. However, this might limit innovation if companies do not evolve toward more tailored AI applications that align closely with their strategic goals. The reliance on ready-made solutions underscores the need for robust AI governance and risk management as businesses scale up their AI capabilities.

Question 3

Tool Used: ChatGPT's "image generator" GPT that runs of GPT-4.

Prompts used: “create a cartoon style portrait of me” and attached a portrait photo of me, and “can you create a webart of the name P Pandya?”

Question 4

Why is it called U-Net?

> U-Net is called so due to its U-shaped architecture. This architecture consists of a contracting path to capture context and a symmetric expanding path that enables precise localization, which together form a U-like shape.

Significance of U-Net beyond biomedical image segmentation:

Beyond biomedical imaging, U-Net has significant implications for other areas of image segmentation where similar precise localization is required. Its ability to work with very few training images and to perform segmentation rapidly makes it useful in fields like satellite image analysis, agricultural imaging, and any area requiring detailed and efficient segmentation.

Important model architectures that use U-Net:

V-Net: Modified version of U-Net for volumetric (3D) image segmentation.

TernausNet: Uses U-Net architecture with VGG11 or VGG16 as the encoder.

R2U-Net: Incorporates recurrent residual convolutional layers into U-Net, enhancing its ability to capture fine details over sequences.

Attention U-Net: Integrates attention mechanisms to focus on specific areas of the input image, enhancing the model's precision.

Question 5

Dataset Used for DDPM Model:

The dataset is "Oxford Flowers102," commonly split into training, validation, and test sets, though exact numbers in each are not specified in the provided text.

Differences in U-Net Architecture in DDPM:

Two Inputs: The network takes both images and time steps as inputs.

Self-Attention Mechanism: Incorporated at certain resolutions for better feature integration.

Group Normalization: Used instead of batch normalization to stabilize training across groups of features.

Residual Blocks: Employed to improve training depth and efficiency.

Noise Prediction: The U-Net predicts noise, which is a deviation from typical U-Nets that predict direct segmentation maps or image features.