POST 7 - GEN AI

UNDERSTANDING GENERATIVE AI IMAGE GENERATION: A STEP-BY-STEP GUIDE MADE EASY

Fundamentals - Part 3

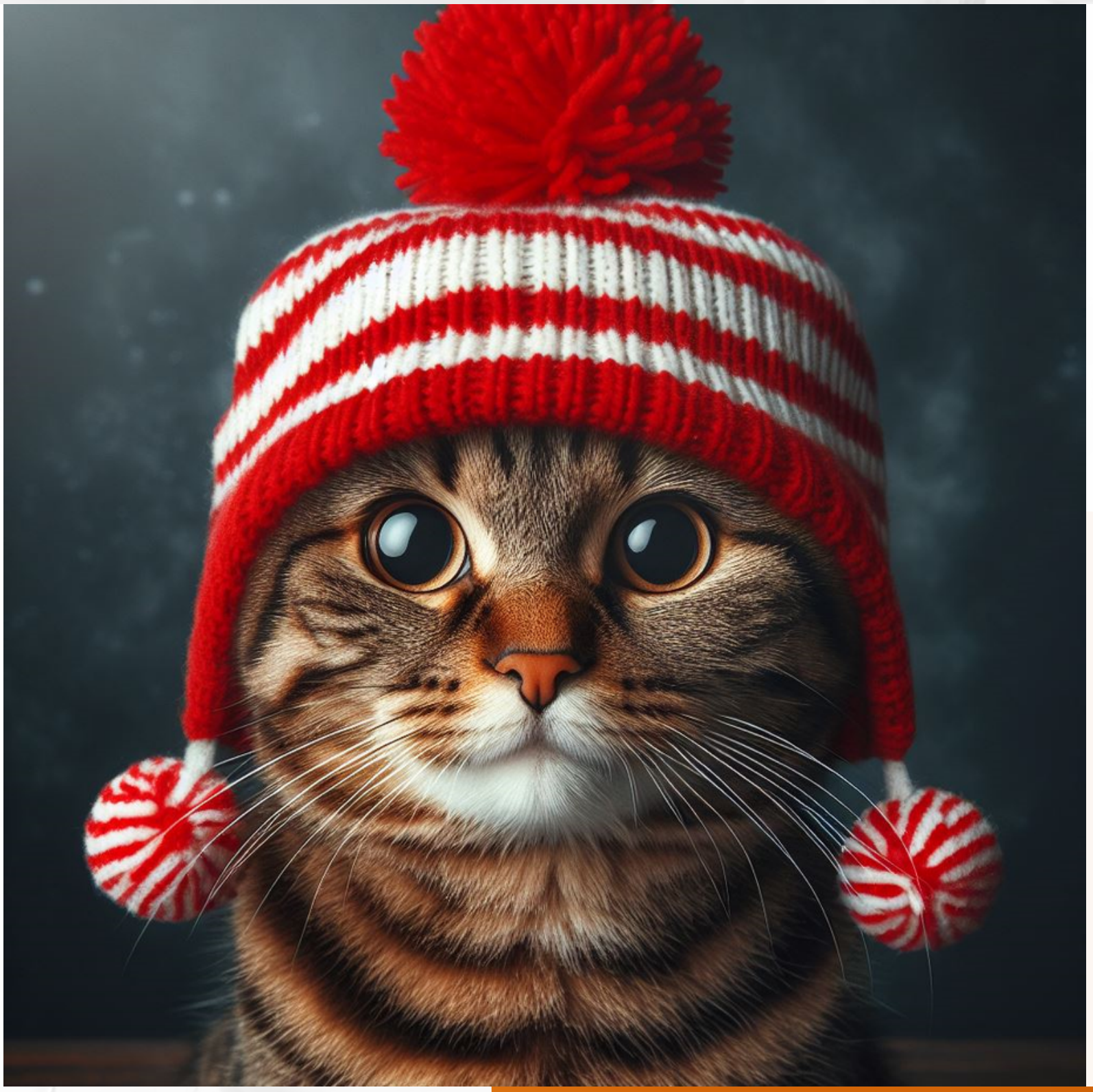
GENERATIVE AI FOR ALL

DINESH LAL
(DIRECTOR, DATA SCIENCE)

DINESHLAL

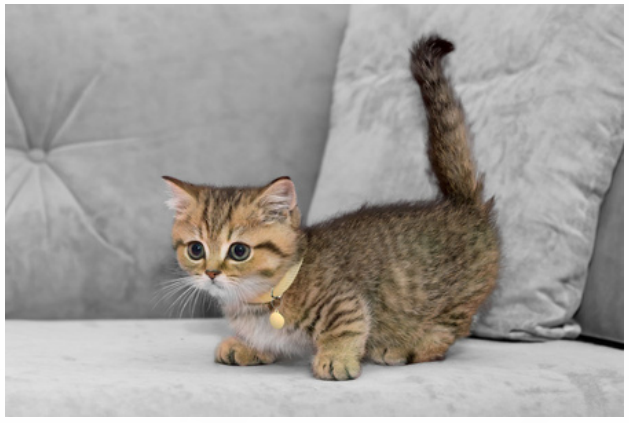
- In this document we will go through the step by step process of image generation by Generative AI
- We will take an example of "create an image of cat with a hat". If we ask this question to any Generative AI bot, how it will work is explained in step by step guide
- The guide is in two sections, one looking it as non technical lens, the other section focuses with technical perspective

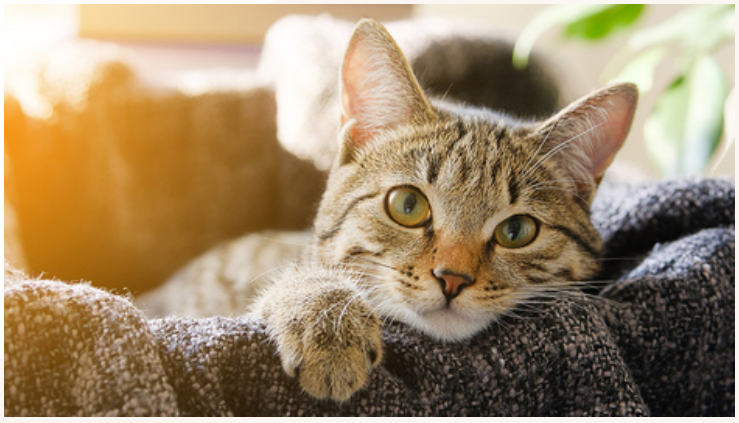
THE OUTPUT



Getting Examples:

- The AI needs to see lots of pictures of cats and hats so it can learn what they look like.





Learning from Pictures:

- It looks at these pictures and tries to figure out what makes a cat and what makes a hat.
- The AI learns from these examples and starts recognizing patterns in the images. It figures out things like shapes, colors, and textures.

Making a New Picture:

- When you ask it to make a picture of a cat with a hat, it uses what it learned to create a new picture that it thinks looks like a cat with a hat.
- It knows how to put together a cat shape, color it purple, etc.

Adjusting if Needed:

- If the picture it makes doesn't look quite right, it tries to fix it based on feedback or things it learned before.

Making it Look Good:

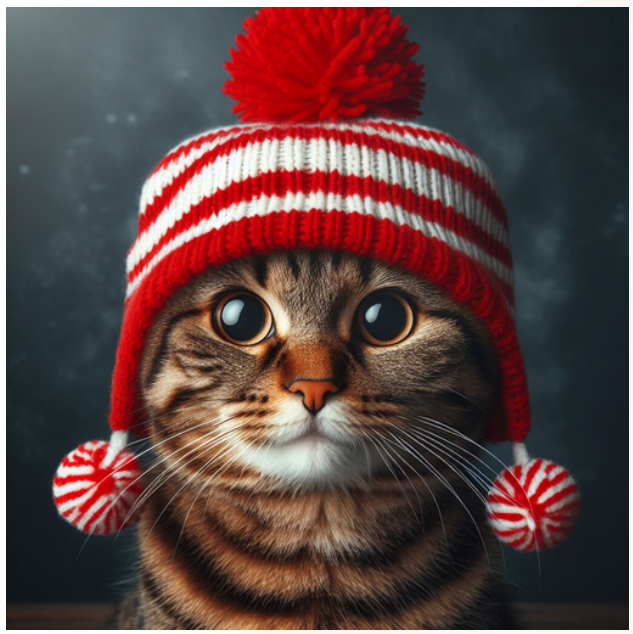
- It tries to make the picture as nice and realistic as possible, so it looks like a real cat with a hat

Checking the Picture:

- After making the picture, it looks at it to make sure it looks like what you asked for—a cat wearing a hat.

Giving You the Picture:

- Finally, it gives you the picture it made of the cat with a hat, and you can use it however you want.



TECHNICAL EXPLANATION

Data Acquisition and Preprocessing:

- The Generative AI model requires a sufficiently large and diverse dataset of labeled images containing cats and hats.
- These images are preprocessed to standardize dimensions, color spaces, and other attributes for consistency during training.

Model Architecture Selection:

- Depending on the complexity of the task, a suitable generative model architecture is chosen, such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), or other deep learning architectures.

Training Procedure:

- The model's parameters are optimized through an iterative training process using techniques such as backpropagation and stochastic gradient descent.
- The model learns to capture the underlying distribution of cats and hats in the dataset while minimizing a predefined loss function.

Latent Space Representation:

- **Images are represented as points in a latent space, typically a lower-dimensional manifold, where each point corresponds to a unique feature vector.**
- **This latent space representation facilitates the generation of new images by sampling from the learned distribution.**

Inference and Sampling:

- During inference, the model generates new images by sampling latent vectors from the learned distribution and mapping them to the image space through the generator network.
- The generator network transforms latent vectors into synthetic images of cats with hats.

Adversarial Training (If applicable) :

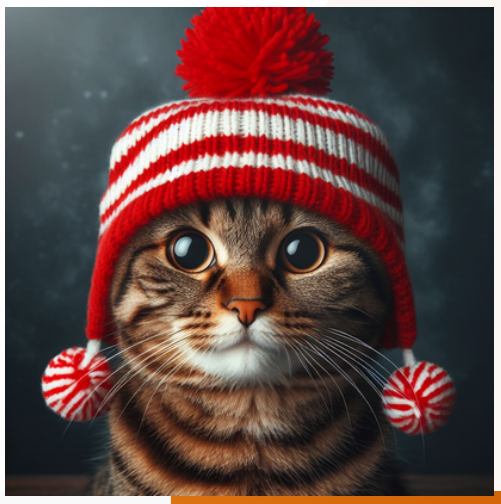
- In the case of GANs, the generator network competes against a discriminator network, which learns to distinguish between real and generated images.
- This adversarial training process helps improve the quality and realism of the generated images.

Evaluation and Validation Metrics:

- The quality and fidelity of the generated images are evaluated using quantitative metrics such as Inception Score, Frechet Inception Distance (FID), or perceptual similarity metrics.
- Additionally, qualitative evaluation by human annotators may be employed to assess the visual realism and coherence of the generated images.

Deployment and Application:

- Once the model achieves satisfactory performance, it can be deployed in production environments for various applications, including image synthesis, content creation, data augmentation, and artistic expression.



Thank You

SPECIAL CREDITS TO OPENAI, CHATGPT,
DALL-E FOR THE CONTENT SUPPORT

DINESHLAL