# Full Fine-Tuning of Stable Diffusion Model

1. **Objective:**
   To fully fine-tune a Stable Diffusion model (e.g., v1.5) on a custom dataset using Google Colab. This approach trains the entire model, allowing it to adapt to new styles, characters, or visual domains, beyond just a single token embedding (as in Textual Inversion).

2. **Requirements:**

   - Google Account with Colab (Pro recommended)
   - Hugging Face account and token (with write access)
   - Dataset of 50–500+ images (at least 512x512 resolution)
   - Knowledge of Python, PyTorch, and Stable Diffusion basics
   - Enough GPU VRAM (Colab T4 = 16GB, A100 = 40GB recommended)

3. **Dataset Preparation:**

   1. Prepare a folder of training images (PNG or JPG format).
   2. Optionally, prepare a `captions.txt` or individual `.txt` files per image.
   3. Zip the dataset and upload to Colab or Google Drive.

4. **Libraries to Install:**

   ```python
   !pip install diffusers transformers accelerate safetensors bitsandbytes datasets ftfy
   ```

5. **Load Base Model:**

   ```python
   from diffusers import StableDiffusionPipeline

   pretrained_model = "runwayml/stable-diffusion-v1-5"
   pipeline = StableDiffusionPipeline.from_pretrained(pretrained_model,
   torch_dtype=torch.float(16)
   ```

6. **Training Script (with Trainer API or Custom Loop):**

   You can either use the `Trainer` API or a custom loop. The recommended method is to use Hugging Face `diffusers/examples/dreambooth/train_dreambooth.py` or `train_text_to_image.py`.

   Steps:
   1. Clone the diffusers GitHub repo:
   ```bash
   !git clone https://github.com/huggingface/diffusers
   cd diffusers
   pip install -e.
   ```

   2. Prepare your data in `/content/data/your_dataset`

   3. Launch training:
   ```bash
   !accelerate launch examples/text_to_image/train_text_to_image.py \
    --pretrained_model_name_or_path="runwayml/stable-diffusion-v1-5" \
    --train_data_dir="./data/your_dataset" \
    --resolution=512 \
   --output_dir="./output" \
   --train_batch_size=2 \
   --gradient_accumulation_steps=2 \
   --learning_rate=5e-6 \
   --max_train_steps=2000 \
   --checkpointing_steps=500 \
   --use_ema \
   --mixed_precision="fp16" \
   ```

7. **Inference After Training:**

   ```python
   from diffusers import StableDiffusionPipeline

   pipe = StableDiffusionPipeline.from_pretrained("./output",
   torch_dtype=torch.float16)
   pipe.to("cuda")

   prompt = "a futuristic building in cyberpunk city"
   image = pipe(prompt).images[0]
   image.save("output.png")
   ```

8. **Upload to Hugging Face:**

   ```python
   from huggingface_hub import notebook_login
   notebook_login()

   from diffusers import StableDiffusionPipeline
   pipe.push_to_hub("your-username/your-model-name")
   ```

9. **Tips:**

   - Increase `max_train_steps` for higher quality
   - Use regularization images for identity preservation (DreamBooth style)
   - Monitor GPU usage and avoid OOM errors by adjusting batch size
   - Save intermediate checkpoints for recovery