Python Coding

```
import torch
from transformers import BlipProcessor, BlipForConditionalGeneration
from PIL import Image
import gradio as gr

# Load BLIP processor and model once (for efficiency)
processor = BlipProcessor.from_pretrained("Salesforce/blip-image-captioning-base")
model = BlipForConditionalGeneration.from_pretrained("Salesforce/blip-image-captioning-base")

# Move model to GPU if available
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
model.to(device)

def generate_caption(image):
    """Generate a caption from an image using BLIP."""
    image = image.convert("RGB")
    inputs = processor(images=image, return_tensors="pt").to(device)

    with torch.no_grad():
        output = model.generate(**inputs)

    caption = processor.decode(output[0], skip_special_tokens=True)
    return caption

# Gradio interface
interface = gr.Interface(
```

```
    fn=generate_caption,
    inputs=gr.Image(type="pil"),
    outputs="text",
    title="🖼 Image Caption Generator",
    description="Upload an image and get a caption generated using the BLIP model (by Salesforce).",
    theme="soft",
    allow_flagging="never"
)

# Launch the web app
interface.launch()
```