horizontal line

**Raspberry AI - Take Home Exercise (AI role)**

# CHALLENGE DESCRIPTION

Our software generates fashion designs in image format based on reference images and text prompts. Customers have specific use cases that may not be honored by out-of-the-box Stable Diffusion image generation models.

The intention is that you don’t spend more than 6 hours on this.

# DELIVERABLE

**Tech spec:** Utilizing the Stable Diffusion open source model as a base, please outline additional methodologies and techniques you would use to achieve the desired customer results. These can be model or non-model (pre/post-processing/prompt engineering) solutions. Please include 2-5 proposed methodologies and any resources/tools required to execute on the methodologies.

**Execution:** Quickly spin up one of your proposed methodologies in whatever platform is easiest for you (e.g Google Colab, Huggingface, etc). If you choose to do this, Raspberry AI will send a $50 gift card to cover any training/GPU costs.

If your approach *didn’t* work, we’d still love to hear about it. Share any notebooks/repos, why you think it didn’t work as you hypothesized, and how you’d iterate on it in the future.

## 

## Consistent outputs

Customer would like to use text-to-image model to create product images with 1) full frontal view 2) realistic photograph 3) on a plain white/grey background. Stable diffusion doesn’t do this very well today, often including mannequins, humans, angled shots, cutting off parts of the garment, or distracting backgrounds even when adding negative prompts and ‘product photo of’ in the prompt.

Desired user flow: text prompt of a desired garment ie ‘black short sleeve crewneck crop top’ (left) or ‘black long sleeve crewneck crop top’ (right)

Example images provided below:



Some datasets you could use:

<https://www.kaggle.com/competitions/h-and-m-personalized-fashion-recommendations/data>

<https://blog.visualdata.io/popular-computer-vision-datasets-fashion-68784fdab723>