

# Generating Product Image from Customer Reviews

Group 2:  
Gilbert Wu  
Sarvesh Siras  
Ashish Palli

# Table of Contents

## [1. Introduction](#)

[Objective:](#)

[Workflow Overview:](#)

## [2. Product Selection and Customer Review Data Collection](#)

[Product Selection](#)

[Reasons for Selection:](#)

[Data Collection](#)

[Sources:](#)

[Data Types Collected:](#)

[Challenges Faced:](#)

## [3. Analysis of Customer Reviews with LLM](#)

[Methodology and Analysis](#)

[Results of Key Insights from the reviews](#)

[Effective Output for Image Generation](#)

## [4. Image Generation with Diffusion Model](#)

[Models Used](#)

[DALLE-3](#)

[Stable Diffusion 2.1](#)

[Prompt Design](#)

[Key Prompting Guidelines](#)

[Example: Initial vs. Final Prompt](#)

[Extracted Features:](#)

[Final Refined Prompt:](#)

[Image Generation Process](#)

[Product Image Variations](#)

[Iterations & Refinements](#)

[Outcome and Observations](#)

[Performance Across Product Categories](#)

[Key Observations](#)

[Limitations](#)

[Recommendations](#)

## [4. AI Agentic Workflow](#)

[Challenges and Lessons Learned](#)

## [5. Conclusion](#)

[Summary of Findings:](#)

[Key Takeaways:](#)

# 1. Introduction

## Objective:

The primary goal of this project is to generate product images based on customer reviews, using advanced AI techniques. This project seeks to answer key questions such as:

- Can AI effectively extract meaningful visual cues from unstructured textual reviews?
- How accurately can diffusion models translate textual insights into detailed product images?

## Workflow Overview:

The project follows a structured methodology to achieve the objective:

1. **Text Analysis using Large Language Models (LLMs):** Extracting relevant features and visual descriptions from product reviews.
2. **Retrieval-Augmented Generation (RAG) and Prompt Engineering:** Enhancing AI-driven text understanding and improving the quality of inputs for image generation.
3. **Image Generation via Diffusion Models:** Creating realistic product images based on structured textual inputs.

The project consists of the following structured phases:

- **Product Selection & Data Collection:** Choosing diverse products with high review volume.
- **LLM-Based Analysis of Customer Reviews:** Extracting key insights from unstructured text.
- **RAG and Prompt Engineering:** Refining queries and optimizing AI-generated prompts.
- **Image Generation Using Diffusion Models:** Producing high-quality images based on textual inputs.
- **Comparative Analysis:** Evaluating the performance of different AI models.

## 2. Product Selection and Customer Review Data Collection

Here, we outline the reasoning behind selecting specific products for AI-generated image testing. It also details the data acquisition strategy, including sources, data types, and challenges encountered. We collected product descriptions, customer reviews, and images from e-commerce platforms, ensuring diverse and balanced datasets to train and evaluate AI models effectively.

### Product Selection

- **Office Chair:** Common ergonomic furniture with distinct design attributes.
- **Dell Monitor:** A technical product with precise visual specifications.
- **Bomber Jacket:** A fashion item with significant texture and design details.

### Reasons for Selection:

- **Diversity:** The products span different categories, including furniture, electronics, and apparel, allowing a broad evaluation of AI capabilities.
- **Popularity:** High review volume ensures a rich dataset for AI analysis.
- **Visual Complexity:** Different challenges in capturing textures, precision, and design help test the effectiveness of the models.

### Data Collection

#### Sources:

- **Amazon:** A popular e-commerce platform with extensive customer feedback.
- **eBay:** Another major online marketplace with varied user reviews.

#### Data Types Collected:

- **Product Metadata:** Name, description, features.
- **Customer Reviews:** Text-based feedback on product experience.
- **Product Images:** Ground truth reference images.

#### Challenges Faced:

- **Diversity of Reviews:** Ensuring a balanced dataset with both positive and negative reviews.
- **Scraping Limitations:** Restrictions on automated data extraction required manual intervention.

### 3. Analysis of Customer Reviews with LLM

In this section, we explore the analysis of customer reviews using Large Language Models (LLMs). We extract key product attributes and visual features through prompt engineering, retrieval-augmented generation (RAG), and text chunking strategies. This analysis ensures that AI-generated images align with customer expectations and product descriptions.

#### Methodology and Analysis

##### Tools Used:

- **OpenAI GPT-4 API:** Analyzing textual descriptions and customer sentiments.
- **LangChain:** Facilitating efficient LLM processing.
- **Vector Database (Pinecone):** Organizing and retrieving product information efficiently.

##### Chunking Strategy:

- **TokenTextSplitter:** Used for token-based chunking, yielding optimal results in processing lengthy reviews.

##### Key Analyses Performed:

- **Product Attribute Summarization:** Extracting detailed insights from product descriptions and reviews.
- **Visual Feature Identification:** Detecting key design elements such as color, shape, and material for accurate image generation.

##### Examples of Prompt Engineering:

- **Query:** “Summarize the product description and customer reviews in a concise bullet-point format highlighting the key features and overall user sentiment for a mesh office chair.”
- **Query:** “Based on customer reviews, identify the key visual features of this office chair, including its material, color, shape, armrest design, backrest style, and any additional ergonomic details.”

# Results of Key Insights from the reviews

## 1. Office Chair

Query: "Summarize the product description and customer reviews in a concise bullet-point format highlighting the key features and overall user sentiment for mesh office chairs."

Output:

```
Here is a summary of the product description and customer reviews for a mesh office chair:

**Product Features:**
- Ergonomic design with adjustable height and tilt functions.
- Breathable mesh back for comfort and temperature regulation.
- Durable construction with a sturdy base and smooth-rolling casters.
- Lumbar support for improved posture during long hours of sitting.
- Contoured seat cushion for enhanced comfort.

**Customer Reviews and Sentiment:**
- Generally positive feedback on comfort and support, particularly for extended use.
- Many users appreciate the breathability of the mesh material.
- Complaints mainly involve assembly difficulties and durability of certain components.
- Overall satisfaction is high, with several reviewers noting the chair's good value for the price.

This summarizes key features and overall user sentiment for the mesh office chair.
```

## 2. Dell Monitor

Query: "Summarize the product description and customer reviews in a concise bullet-point format highlighting the key features and overall user sentiment Dell monitor."

Output:

```
Here is a summary of the product description and customer reviews for a mesh office chair:

**Product Features:**
- Ergonomic design with adjustable height and tilt functions.
- Breathable mesh back for comfort and temperature regulation.
- Durable construction with a sturdy base and smooth-rolling casters.
- Lumbar support for improved posture during long hours of sitting.
- Contoured seat cushion for enhanced comfort.

**Customer Reviews and Sentiment:**
- Generally positive feedback on comfort and support, particularly for extended use.
- Many users appreciate the breathability of the mesh material.
- Complaints mainly involve assembly difficulties and durability of certain components.
- Overall satisfaction is high, with several reviewers noting the chair's good value for the price.

This summarizes key features and overall user sentiment for the mesh office chair.
```

## 3. Bomber Jacket

Query: "Summarize the product description and customer reviews in a concise bullet-point format highlighting the key features and overall user sentiment Bomber jacket."

Output:

```

- Product Description:
  - High visibility insulated bomber jacket designed for road work and construction.
  - Features a weatherproof polyester shell with a quilted lining.
  - Includes a concealed hood, zipper front closure, and reflective material.
  - Offers 5 pockets including a cell phone/radio chest pocket with a clear panel.
  - Available in a range of sizes from Small to 6XLarge.

- Customer Reviews:
  - General satisfaction with the jacket's warmth, quality, and visibility.
  - Mixed feedback on the hood; some find it bulky or not very functional.
  - Positive comments on fast shipping and good material quality.
  - Some issues with sizing discrepancies and customer service responsiveness.
  - Overall, most customers find the jacket to offer good value and effectiveness for its purpose.

```

## Effective Output for Image Generation

### 1. Office Chair

Query: "Based on customer reviews, identify the key visual features of this office chair, including its material, color, shape, armrest design, backrest style, and any additional ergonomic details.",

Output:

```

Based on customer reviews, the key visual features of the office chair include:

- Material: Made from high-quality fabric or synthetic leather, providing durability and comfort.
- Color: Comes in a variety of colors, often including black, grey, and other neutral tones to match office decor.
- Shape: Ergonomically designed with a contoured seat and backrest to support the body.
- Armrest Design: Features adjustable armrests that can be raised or lowered for optimal comfort.
- Backrest Style: Typically includes a mesh backrest for breathability or a padded backrest for added comfort.
- Additional Ergonomic Details: Includes features like lumbar support, adjustable tilt and height mechanisms, and a swivel base with caster wheels for easy movement.

These features are intended to provide comfort during long hours of sitting and to fit various body types and preferences.

```

### 2. Dell Monitor

Query: "Extract the visual characteristics of this Dell monitor from customer reviews, including screen size, bezel thickness, color, stand design, and any notable display features."

Output:

```

The visual characteristics of the Dell monitor based on customer reviews include:

- Screen Size: The monitor is described as being large and suitable for productivity tasks but no specific size is mentioned.
- Bezel Thickness: The monitor has a minimal bezel with about 3/8-inch at the top and sides, and about 3/4-inch at the bottom, offering a nearly full-screen display experience.
- Color: The color quality is described positively, with colors being true and accurate, although some customers experienced issues with color calibration.
- Stand Design: Initially wobbly for some, but stable once the connection between the base and the post is tightened as per the set-up instructions.
- Notable Display Features: The monitor offers a crisp, sharp picture with effective anti-glare coating. It does not support USB-C/Thunderbolt connectivity, limiting it to HDMI and

+ Code + Markdown

```

### 3. Bomber Jacket

Query: "Analyze customer reviews to determine the key visual features of this bomber jacket, such as material (e.g., leather, nylon), color options, fit (slim or oversized), collar style, pocket placement, and any distinctive design elements."

Output:

The reviews highlight several key visual features of the bomber jacket:

1. **\*\*Material\*\***: The jacket has a weatherproof polyester shell with a quilted lining.
2. **\*\*Color Options\*\***: It is available in a high visibility lime green with a black bottom.
3. **\*\*Fit\*\***: It does not specify if it's slim or oversized, but it's available in sizes ranging from Small to 6XL.
4. **\*\*Collar Style\*\***: It includes a concealed hood, but no specific collar style is mentioned.
5. **\*\*Pocket Placement\*\***: There are a total of 5 pockets, with 4 on the outside and 1 on the inside. There is a cell phone/radio chest pocket with a clear panel.
6. **\*\*Distinctive Design Elements\*\***: The jacket features 2" silver reflective material, mic tabs, and pencil pockets, indicating that it is designed for high visibility and work utility.

Overall, the jacket is designed for road work or construction with safety and functionality in mind.



## 4. Image Generation with Diffusion Model

In this section, we do a detailed analysis of image generation using diffusion models from the extracted features, specifically focusing on DALLÉ-3 and Stable Diffusion 2.1. We generated high-quality product images based on extracted customer review features. The methodology involves structured prompt design, iterative refinements, and comparisons between the two models.

### Models Used

#### DALLÉ-3

- A state-of-the-art diffusion model by OpenAI, optimized for high-quality and diverse image generation.
- Strengths: High resolution, superior fidelity, and accurate color representation.
- Weaknesses: Less realistic product representations in some cases.

#### Stable Diffusion 2.1

- Open-source model known for its realistic rendering capabilities.
- Strengths: Greater flexibility and customization.
- Weaknesses: Slower generation time and occasional loss of finer details.

### Prompt Design

The prompt design process was crucial in ensuring accurate product representation. The extracted product features were converted into structured prompts, refined using ChatGPT-4o, and optimized to fit within the 77-token limit for Stable Diffusion.

#### Key Prompting Guidelines

- Structured prompts included specific directives: *"Generate high-resolution, detailed, realistic image."*
- Excluded elements: Faces, people, text, labels, or reviews in the images.
- Ensured professional and clean aesthetics for product representation.

#### Example: Initial vs. Final Prompt

### Extracted Features:

- **Material:** Weatherproof polyester shell with quilted lining.
- **Color Options:** Lime green top with a black bottom.
- **Fit:** Sizes range from Small to 6XL.
- **Collar Style:** Concealed hood.
- **Additional Features:** Five pockets (four outside, one inside), clear-panel chest pocket for phone or radio, two silver reflective strips, mic tabs, and pencil pockets.

### Final Refined Prompt:

*"Generate high-resolution, detailed, realistic image of a high-visibility bomber jacket. Weatherproof polyester shell, quilted lining, lime green top, black bottom. Features concealed hood, five pockets including a clear-panel chest pocket for phone or radio, two silver reflective stripes, mic tabs, and pencil pockets. No text, labels, or people included."*

## Image Generation Process

### Product Image Variations

- Generated 4-6 variations per product.
- Used both DALLÉ-3 and Stable Diffusion 2.1.
- Two prompting styles: *Raw* vs. *Concise*.

### Iterations & Refinements

- Implemented specific directives to enhance image accuracy.
- Adjusted parameters including *num\_iterations* and *guidelines* in Stable Diffusion.
- Improved resolution and image size to enhance visual clarity.

## Outcome and Observations

### Performance Across Product Categories

Product Category	DALLÉ-3 Performance	Stable Diffusion 2.1 Performance
Office Chair	High accuracy & quality	Realistic but slower processing

Desktop Monitor	Excellent fidelity	Good accuracy but missed details
Bomber Jacket	High resolution but unrealistic details	Realistic but lacked accuracy

## Key Observations

- **Popular products (e.g., office chair, desktop monitor)** were well-represented by both models.
- **DALLE-3 outperformed Stable Diffusion 2.1** in terms of accuracy, fidelity, and resolution.
- **Structured product descriptions improved the feature accuracy of generated images.**
- **Fine-tuned prompts produced better results** than simple descriptive summaries.
- **Face and people exclusion improved model performance** due to clear product focus.

## Limitations

- **Bomber Jacket Generation:**
  - Models captured general color and material details but struggled with hoodie shape and design accuracy.
  - DALLE-3 prioritized high resolution over realism.
  - Stable Diffusion 2.1 generated realistic images but missed some crucial product details.
- **Speed Comparison:**
  - **Stable Diffusion 2.1 was slower** than DALLE-3 but did not necessarily generate better images.

## Recommendations

1. **Further fine-tuning of prompts** is required, particularly for products with intricate designs like bomber jackets.
2. **Experimenting with model parameters** (e.g., guidance scale, negative prompts) can help improve fidelity in Stable Diffusion.
3. **Combining both models** could leverage DALLE-3's resolution quality and Stable Diffusion's realism.
4. **Training a fine-tuned diffusion model** on specific product datasets could enhance accuracy and consistency.
5. **Exploring alternative models like Midjourney** may provide additional insights into image generation performance.

## 4. AI Agentic Workflow

We used code based agentic flow which takes image name and generates image from Dalle.

The screenshot below shows the agent function

```
agent_generate_image_product_dalle("mesh chair").
```

The output image is shown in the screenshot and the image file path is also shown.

The agent takes single parameter which is the object name. It queries vector db, creates prompt and generates image from dalle.

```
+ Code + Text
```

```
#Call agent with product name to generate image
image_dalle = agent_generate_image_product_dalle("mesh chair")

# Display the image
image = Image.open(image_dalle)
width, height = image.size
new_size = (width // 4, height // 4)
small_image = image.resize(new_size, Image.LANCZOS)
display(small_image)
```

100%|██████████| 1/1 [00:00<00:00, 1.03it/s]

Here is a concise summary of the product description and customer reviews for the Mesh Office Chair:

**\*\*Key Features:\*\***


- High-density, breathable mesh back with lumbar support and thick cushion for comfort.
- Adjustable, flip-up padded armrests for space-saving and easy storage.
- Features include pneumatic seat height adjustment, tilt lock function, 360-degree swivel, and smooth rolling casters.
- Designed for a maximum capacity of 300 lbs and suitable for various settings such as offices, living rooms, and meeting rooms.
- Comes with installation instructions and video; assembly typically takes 10-20 minutes.

**\*\*Customer Reviews and Overall Sentiment:\*\***

- Users find the chair easy to assemble and express satisfaction with its comfort and quality materials.
- Many customers comment on the chair's aesthetic appeal and value for money.
- The shipping is often praised for being fast, with products arriving in great condition.
- Overall, feedback is very positive, with quick delivery and a high level of customer satisfaction with both the product and the seller.

Generate high-resolution, detailed, realistic image of a mesh chair: Here is a concise summary of the product description and customer DALL-E image URL: <https://oaidalleapiprodscus.blob.core.windows.net/private/org-KaNx240Lxy4bK0Sxxo5SKFe8/user-d8Af2XMg1dsXx1IiRGRRdJap/>

Image successfully saved to /content/gdrive/MyDrive/Colab\_Notebooks/GenAI\_Lab/mesh chair\_dalle\_1.png



The agent function to create image from Dalle is explained below with explanation in comments:  
Multiple helper functions are used for each substep

Agent function to create image based on product name, returns file path of generated Dalle image

```
def agent_generate_image_product_dalle(product_name: str) -> str:

    #Get prompt to generate image from product name using cosine similarity
    in vector db and chat-gpt 4o

    prompt_text = get_image_prompt_for_product(product_name)

    #Clean up the prompt for Dalle input

    final_prompt = create_prompt_image_generation(prompt_text)

    #Final prompt for Dalle with specific instructions

    prompt_final_image = "Generate high-resolution, detailed, realistic image
of a "+ product_name +": "+ final_prompt

    print(prompt_final_image)

    #Call function to generate image from Dalle with product name and final
    prompt, the last parameter is for suffix in image name, default suffix is
    1 eg. Mesh_Chair_Dalle_1.png

    image_path_dalle = generate_image_dalle(product_name,
prompt_final_image,1)

    #Return saved image path

    return image_path_dalle
```

The full code is in Agentic\_Image\_Generation\_from\_ProductReviews.ipynb attached in files.

Note: We need to run the code for storing reviews of products in pinecone db once in section

#Chunk text and store in pinecone db - Run Once

In the jupyter notebook. And then just call the agent function

agent\_generate\_image\_product\_dalle to generate image.

## Challenges and Lessons Learned

- Web Scraping is hard from Amazon, ebay and most marketplaces due to strict policies, we used manual scraping to get reviews
- Agentic AI take about 20-30s to generate image from reviews which is pretty good
- The user has to just input product name and rest of the process is automated by Agentic function.
- This can be further improved by increasing parameters and steps in agentic function.

## 5. Conclusion

### Summary of Findings:

- AI-driven image generation is highly effective for well-defined product categories with structured attributes, such as office chairs and monitors.
- DALL-E-3 consistently outperformed Stable Diffusion 2.1 in terms of resolution and accuracy.
- Apparel products, like bomber jackets, proved more challenging for both models due to intricate design variations and textures.
- Structured prompt engineering and iterative refinement significantly improved image quality and model output consistency.
- Text-to-image generation aligns well with customer reviews, but some finer product details were missed, requiring further optimization.

### Key Takeaways:

- AI-generated images based on customer reviews can be a powerful tool for visualizing products, but optimization is needed for complex designs.
- Prompt refinement plays a crucial role in improving image accuracy.
- DALL-E-3 is more reliable for high-resolution image generation, while Stable Diffusion offers flexibility with more realistic outputs but at the cost of fine detail accuracy.

