Automating Fashion Tagging with GPT-4o:
A Scalable AI Solution
(Project in collaboration with The Webster)

Shanmukh Sri Surya Gopi, Shiva Kumar Midde, Abhiram Reddy Gunutula, Shriteja Salunkepatil
Instructor: Alkiviadis Vazacopoulos

STEVENS
INSTITUTE *of* TECHNOLOGY
THE INNOVATION UNIVERSITY®
1870
Business Intelligence & Analytics

Introduction

THE WEBSTER

The Webster is a U.S. luxury fashion retailer offering curated designer collections through boutiques and an online store.
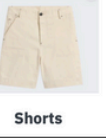
Eg:
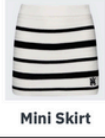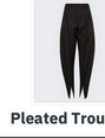
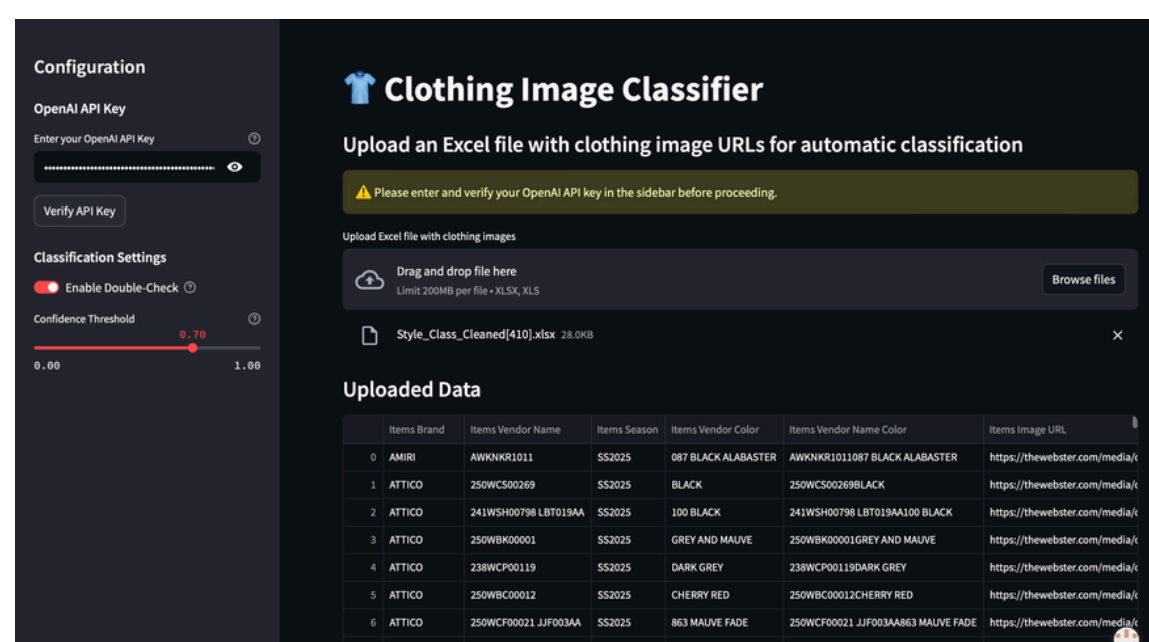| Example 1	Example 2	Example 3	?
Evening Gown	Jumpsuit	Mini Day Dress	

Model learns from labeled products
to *tag new, unseen items.*

- Manual tagging was slow and inconsistent, causing search errors and poor user experience.
- We built a GPT-4o-powered system to automate tagging with speed, accuracy, and scale.

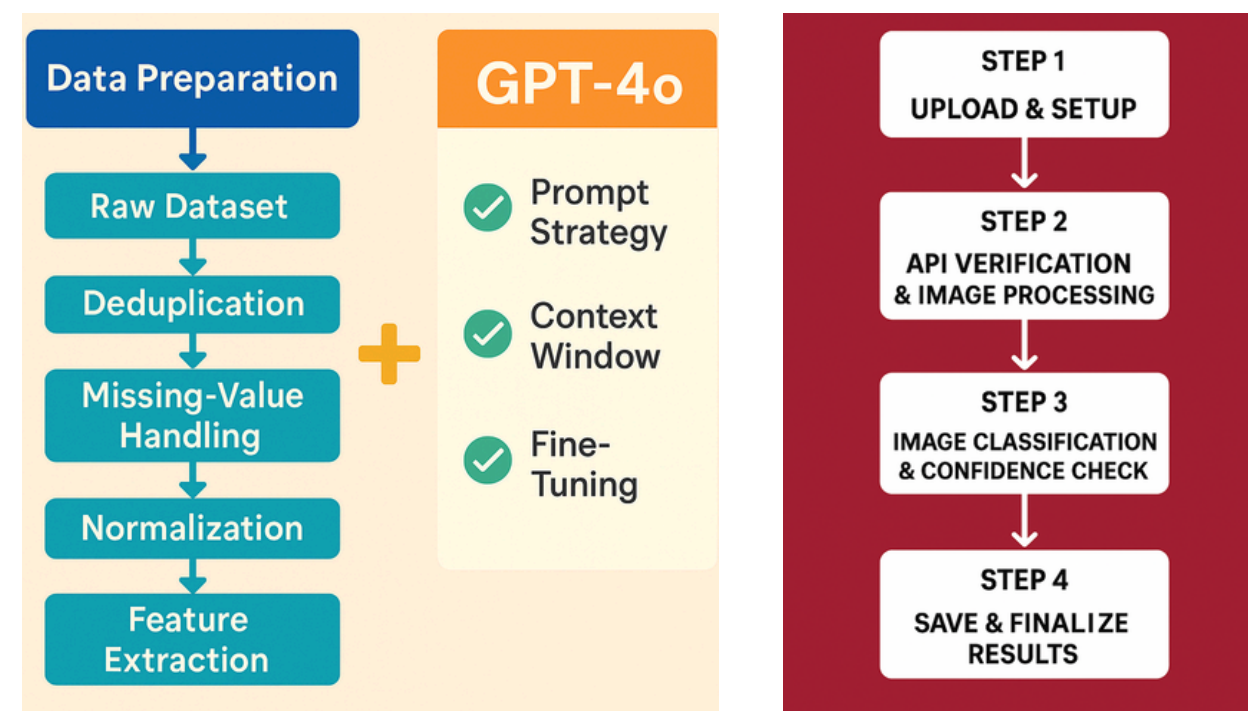## Data Preparation & Category Design

- Merged 94 noisy categories into 10 clean, business-friendly classes
- Cleaned data: removed duplicates & filled missing values
- Extracted features:
1. Text → TF-IDF
2. Image → ResNet-50
3. Metadata → one-hot (brand, price, color)

94 categories were very specific	Example 1	Example 2	Example 3
Outerwear & Jackets	Fashion Coat	Tailored Jacket	Leather Jacket
Bottoms	Shorts	Mini Skirt	Pleated Trousers

GPT-4o Integration & Tagging Pipeline

- Used GPT-4o with 5-shot in-context prompting
- Pipeline: Clean data → Vectors → GPT-4o → Predicted tags
- No fine-tuning needed — zero-shot generalization worked effectively

Data Preparation → Raw Dataset → Deduplication → Missing-Value Handling → Normalization → Feature Extraction

GPT-4o
- ✓ Prompt Strategy
- ✓ Context Window
- ✓ Fine-Tuning

STEP 1 UPLOAD & SETUP
STEP 2 API VERIFICATION & IMAGE PROCESSING
STEP 3 IMAGE CLASSIFICATION & CONFIDENCE CHECK
STEP 4 SAVE & FINALIZE RESULTS

Results

Accuracy ↑ and manual effort ↓ across tagging

Method	Accuracy	High	Medium	Low
Manual	**76%**	76%	24%	**0%**
Early ML Models	**23%**	34%	16%	**50%**
GPT-4o (Final)	**84.5%**	84.6%	11.6%	**3.8%**

🟩 Desired Result (low manual effort)
🟥 Needs Improvement (high manual effort)

Conclusion & What's Next

🚀 Key Outcomes
- 84.5% accuracy with GPT-4o
- 50% reduction in tagging time
- $0.45 per 75 items
- 150% ROI in 3 years

🔧 What's Next
- Add human reviewers for low-confidence outputs
- Enable batch tagging for seasonal surges
- Use few-shot learning for new categories