Appendix A. User study 1: Evaluation of top-5 outfit candidates from text-to-outfit retrieval.

The following figures are some actual examples presented in the user study conducted with Amazon Mechanical Turk to evaluate the satisfaction of our model’s recommendations. As described in Section 4.5, the top-5 results retrieved by our model based on the text queries generated by ChatGPT are displayed for each question.

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. After viewing the recommended results, please indicate if they are satisfactory with a 'Yes' or 'No', and rate them on a scale from 1 to 10.

Provided Text: “casual coat outfit that can be warmly worn in winter.”
Visual Result:
1: 
2: 
3: 
4: 
5: 

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. After viewing the recommended results, please indicate if they are satisfactory with a 'Yes' or 'No', and rate them on a scale from 1 to 10.

Provided Text: “semi-formal style suitable for a wedding party.”
Visual Result:
1: 
2: 
3: 
4: 
5: 

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. After viewing the recommended results, please indicate if they are satisfactory with a 'Yes' or 'No', and rate them on a scale from 1 to 10.

Provided Text: “glittery look that stands out at the club.”
Visual Result:
1: 
2: 
3: 
4: 
5: 

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. After viewing the recommended results, please indicate if they are satisfactory with a 'Yes' or 'No', and rate them on a scale from 1 to 10.

Provided Text: “vintage look suitable for the beach in the summer.”
Visual Result:
1: 
2: 
3: 
4: 
5: 

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. After viewing the recommended results, please indicate if they are satisfactory with a 'Yes' or 'No', and rate them on a scale from 1 to 10.

Provided Text: “soft hippie look perfect for a music festival.”
Visual Result:
1: 
2: 
3: 
4: 
5: 

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. After viewing the recommended results, please indicate if they are satisfactory with a 'Yes' or 'No', and rate them on a scale from 1 to 10.

Provided Text: “campus look with jeans and sneakers.”
Visual Result:
1: 
2: 
3: 
4: 
5: 

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. After viewing the recommended results, please indicate if they are satisfactory with a 'Yes' or 'No', and rate them on a scale from 1 to 10.
Appendix B. User study 2: A/B test between the result of our model and the ground truth

The following figures are some actual examples of A/B tests in the user study conducted with Amazon Mechanical Turk, described in Section 4.5. The outfit sets highlighted in red boxes are the text-outfit retrieval results from our model, while the other outfit sets are the ground truth from the test set.

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. Please choose which recommended result (outfit A or outfit B) you think is better.

Text Query : sunshine monday blue.

Outfit A :

Outfit B :

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. Please choose which recommended result (outfit A or outfit B) you think is better.

Text Query : disco night. Saturday Night

Outfit A :

Outfit B :

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. Please choose which recommended result (outfit A or outfit B) you think is better.

Text Query : winter urban chic. Winter Sunglasses

Outfit A :

Outfit B :

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. Please choose which recommended result (outfit A or outfit B) you think is better.

Text Query : denim trend black jeans. Biker Jeans

Outfit A :

Outfit B :
We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. Please choose which recommended result (outfit A or outfit B) you think is better.

**Text Query : summery blue. Summery Fedoras**

Outfit A:

Outfit B:

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. Please choose which recommended result (outfit A or outfit B) you think is better.

**Text Query : Flea market Sunday outfit**

Outfit A:

Outfit B:

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. Please choose which recommended result (outfit A or outfit B) you think is better.

**Text Query : Long Black Dresses outfit.**

Outfit A:

Outfit B:

We aim to search for fashion outfits using text. Below are the outfit results based on the provided text. Please choose which recommended result (outfit A or outfit B) you think is better.

**Text Query : plus size swimsuit.**

Outfit A:

Outfit B: