

# Max-Min Selection Strategy Example

Let's use a simplified example with just 5 images (A, B, C, D, E) where we want to select 3 diverse images. We'll use made-up cosine similarity values.

## Initial State

Similarity Matrix (higher number = more similar):

	A	B	C	D	E
A	1.0	0.7	0.3	0.8	0.2
B	0.7	1.0	0.4	0.6	0.5
C	0.3	0.4	1.0	0.2	0.9
D	0.8	0.6	0.2	1.0	0.1
E	0.2	0.5	0.9	0.1	1.0

## Selection Process

### Step 1: Random Initial Selection

- Randomly select first image, let's say we get 'A'
- Selected: [A]
- Remaining: [B, C, D, E]

### Step 2: Find Second Image

1. Calculate similarities to A:
  - B->A: 0.7
  - C->A: 0.3
  - D->A: 0.8
  - E->A: 0.2
2. Choose image with lowest similarity to A:
  - E has lowest similarity (0.2)
  - Select E

Current state:

- Selected: [A, E]
- Remaining: [B, C, D]

### Step 3: Find Third Image

1. Calculate similarities to BOTH A and E for each remaining image:
  - B: [0.7 to A, 0.5 to E] -> min = 0.5

- C: [0.3 to A, 0.9 to E] -> min = 0.3
- D: [0.8 to A, 0.1 to E] -> min = 0.1

2. Choose image with lowest MAXIMUM similarity:

- D has lowest maximum similarity (0.1 to E)
- Select D

Final Selection: [A, E, D]

## Key Points

1. We don't "pair up" images - instead, each new selection considers its relationship to ALL previously selected images
2. For each candidate image, we:
  - Find its similarity to all selected images
  - Take the MINIMUM similarity (representing its "best case" for being different)
  - Compare these minimums across all candidates
  - Select the one with the lowest similarity
3. This ensures each new selection is maximally different from ALL previously selected images, not just the last one
4. The process is deterministic after the first random selection, but the initial random selection means you might get different results each run

In the real script with 100 images:

- Starts with one random image
- Each subsequent selection looks at its similarity to ALL previously selected images
- Chooses the image that's most different from its most similar already-selected image
- This continues until we have 100 images that are maximally diverse from each other