SIGIR 2025 Full Papers Track

Disentangling and Generating Modalities for Recommendation in Missing Modality Scenarios

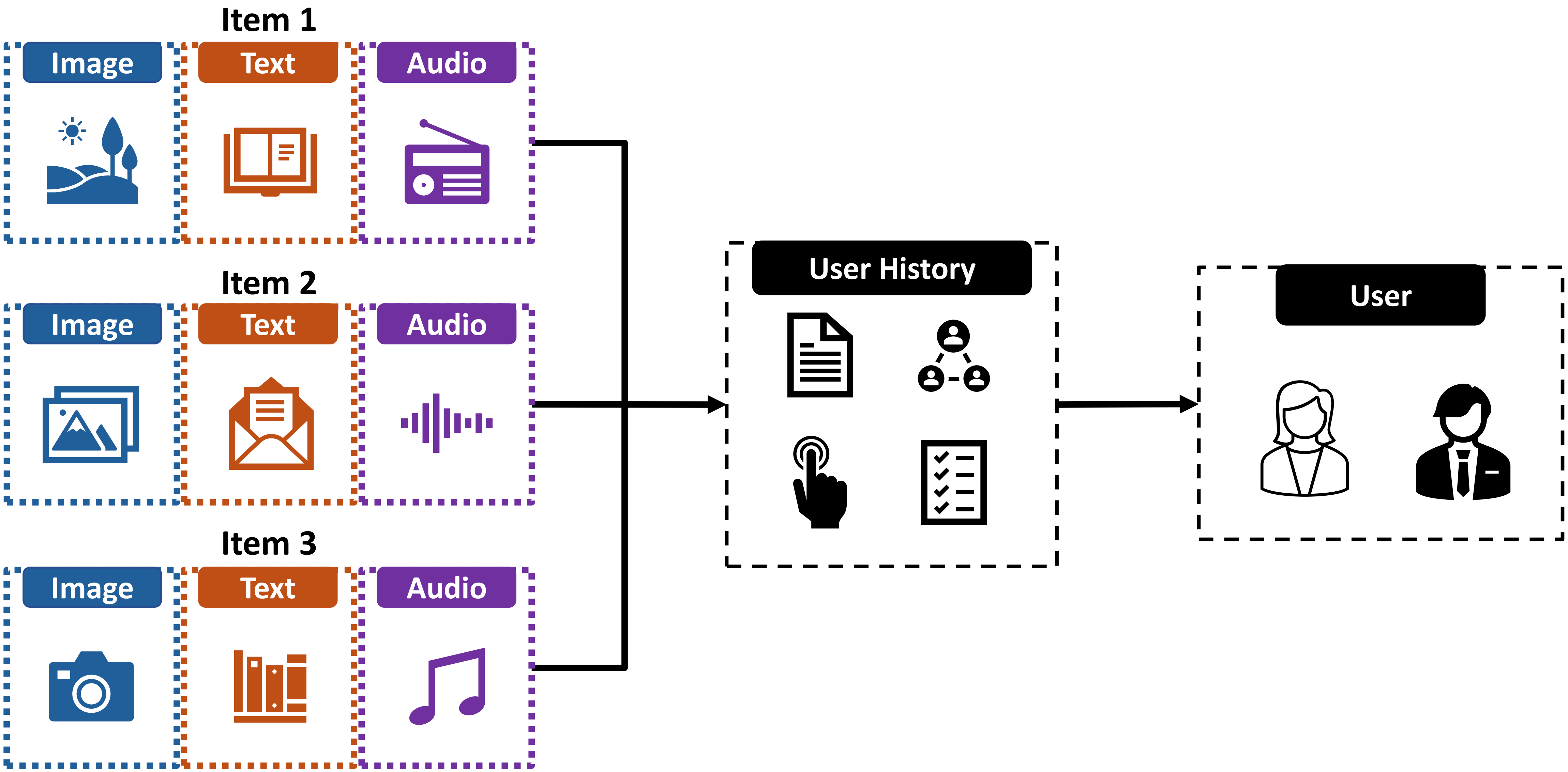
Jiwan Kim, Hongseok Kang, Sein Kim,
Kibum Kim, Chanyoung Park

Korean Advanced Institute of Science and Technology (KAIST)

DSAIL Data Science & Artificial Intelligence

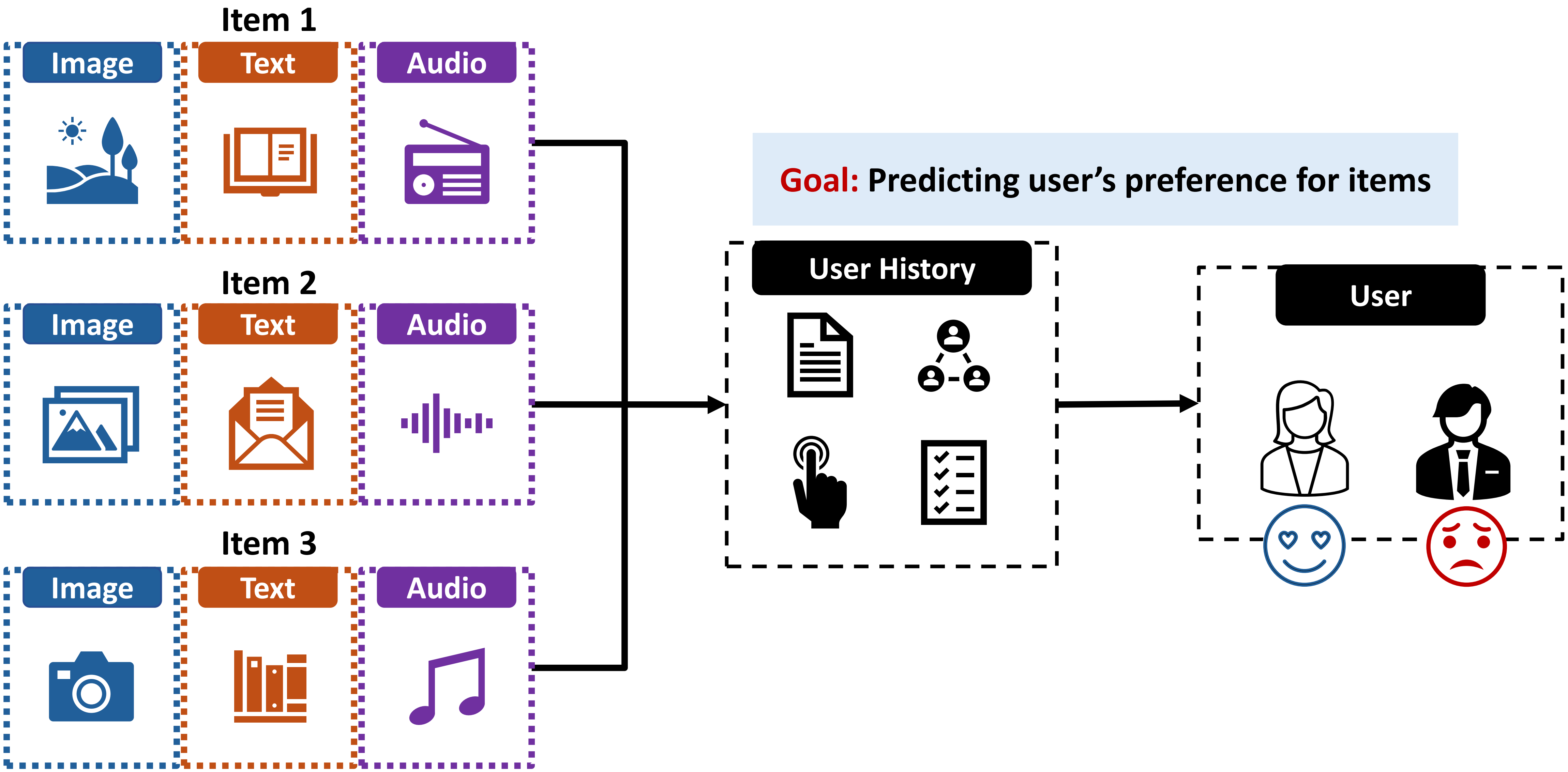
MOTIVATION

- Multi-modal Recommendation System



MOTIVATION

- Multi-modal Recommendation System

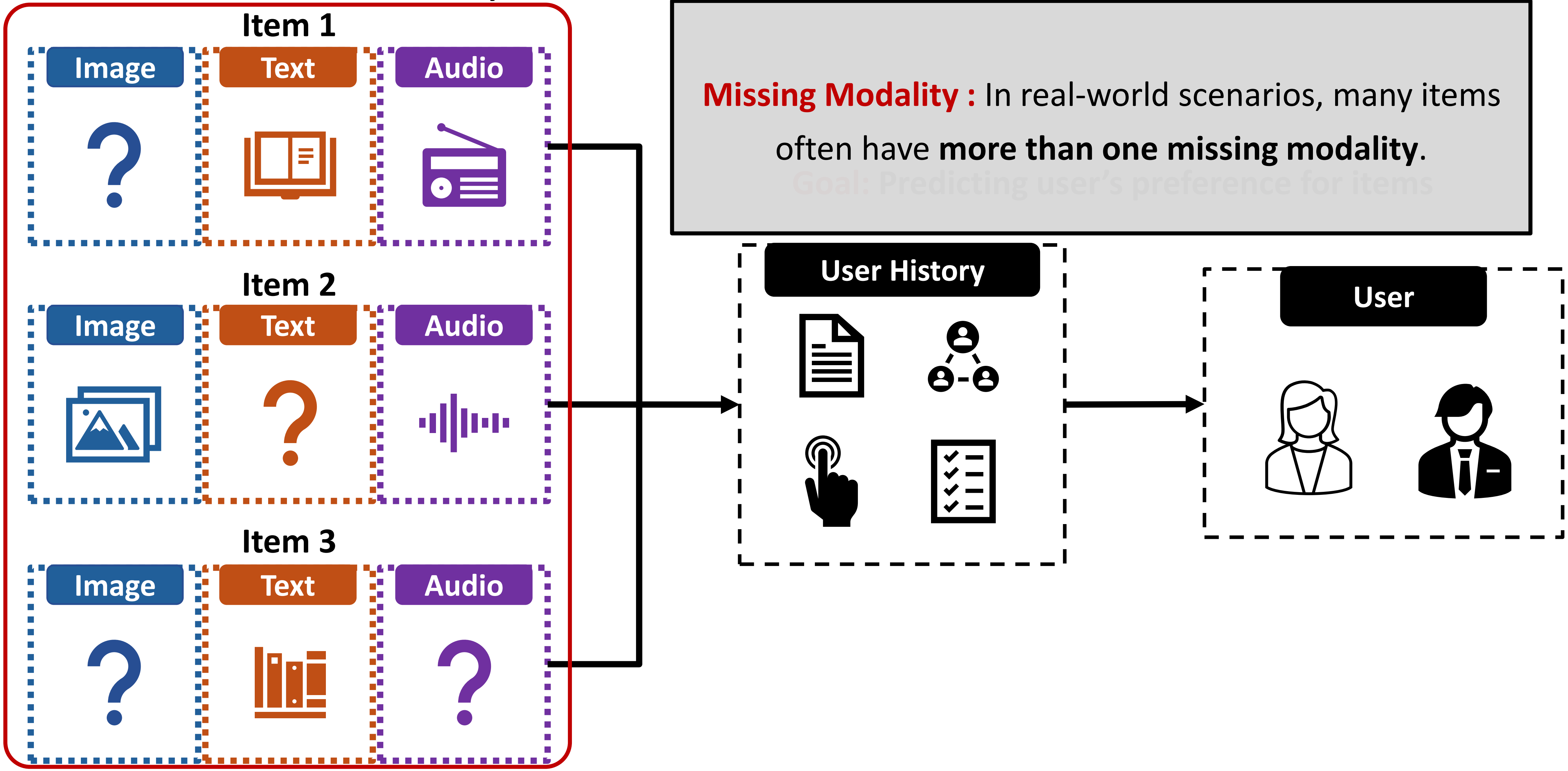


MOTIVATION

- Multi-modal Recommendation System

Missing Modality : In real-world scenarios, many items often have **more than one missing modality**.

Goal: Predicting user's preference for items

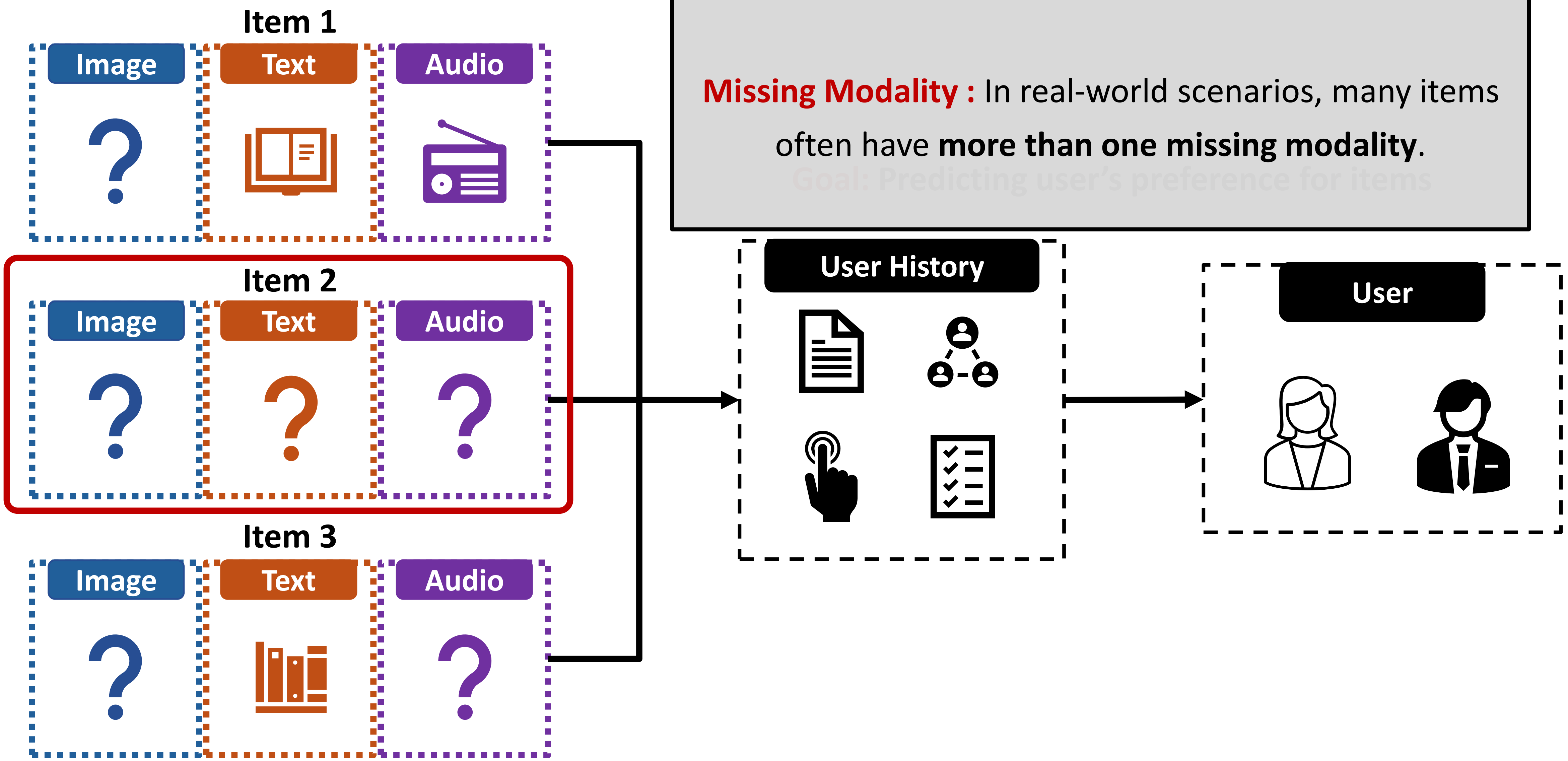


MOTIVATION

- Multi-modal Recommendation System

Missing Modality : In real-world scenarios, many items often have **more than one missing modality.**

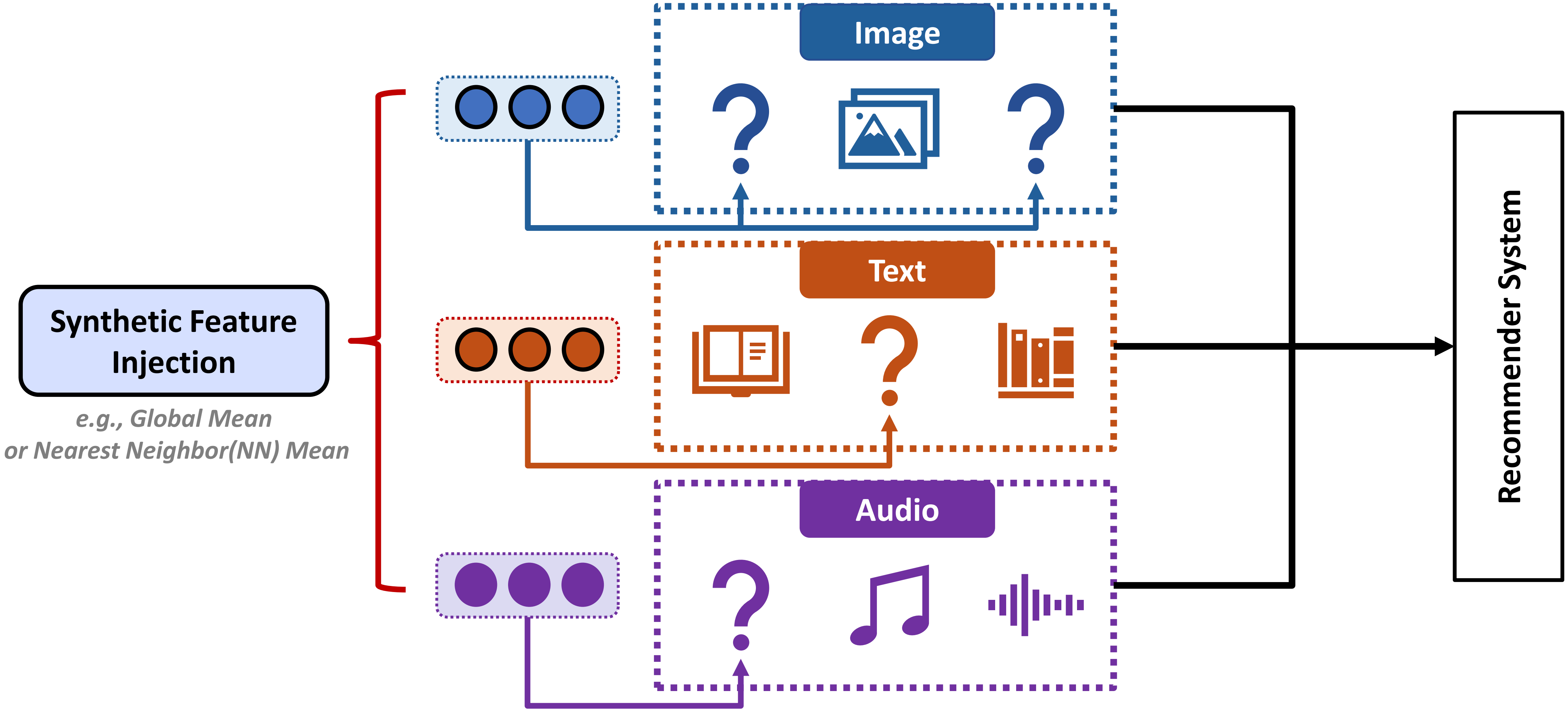
Goal: Predicting user's preference for items



MOTIVATION

- **Multi-modal Recommendation System**

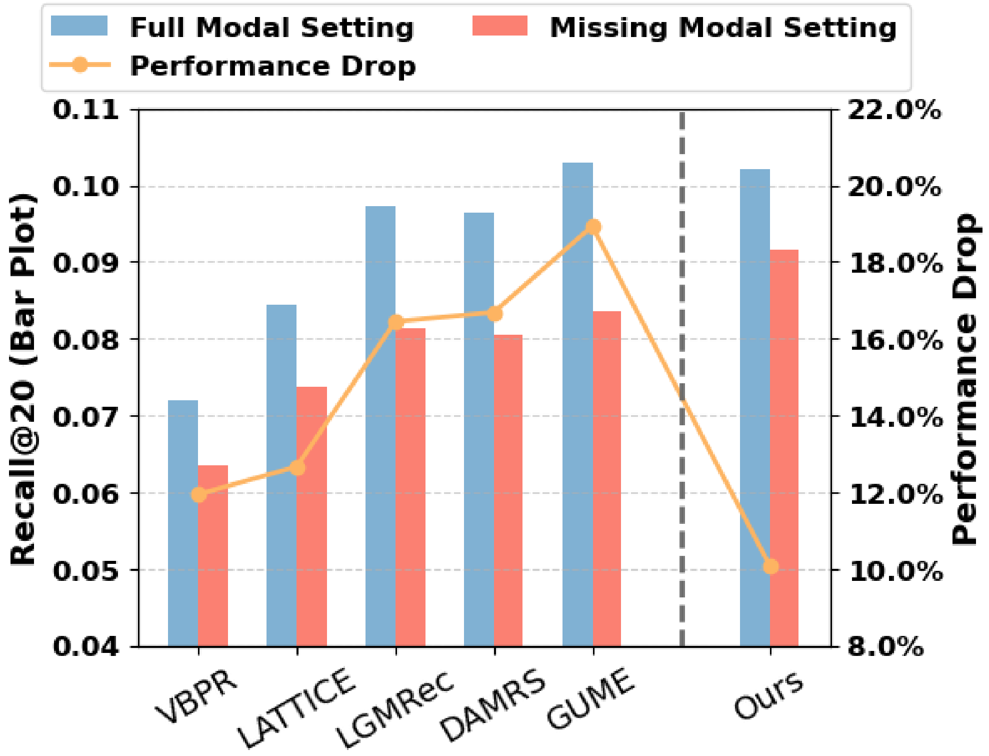
  Missing Modality : In real-world scenarios, many items often have **more than one missing modality**.



MOTIVATION

- **Multi-modal Recommendation System**

  Missing Modality : In real-world scenarios, many items often have **more than one missing modality**.



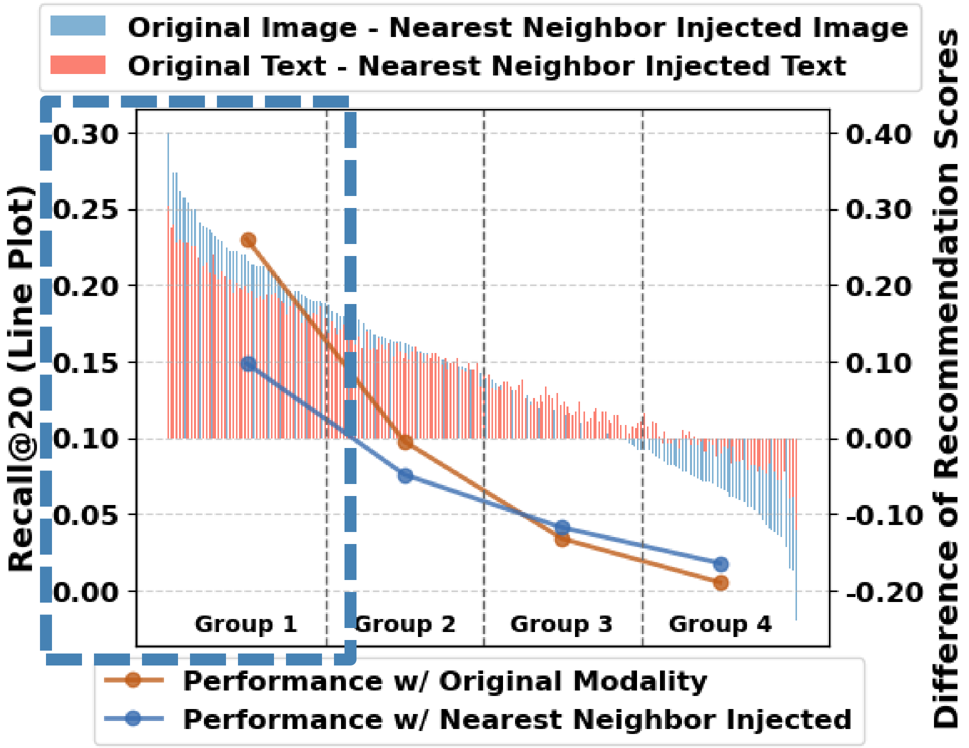
Missing Modality leads to following problems:

1) **Can not be used in training**
2) **Significant performance decrease in inference**

MOTIVATION

- **Multi-modal Recommendation System**

Missing Modality : In real-world scenarios, many items often have **more than one missing modality**.

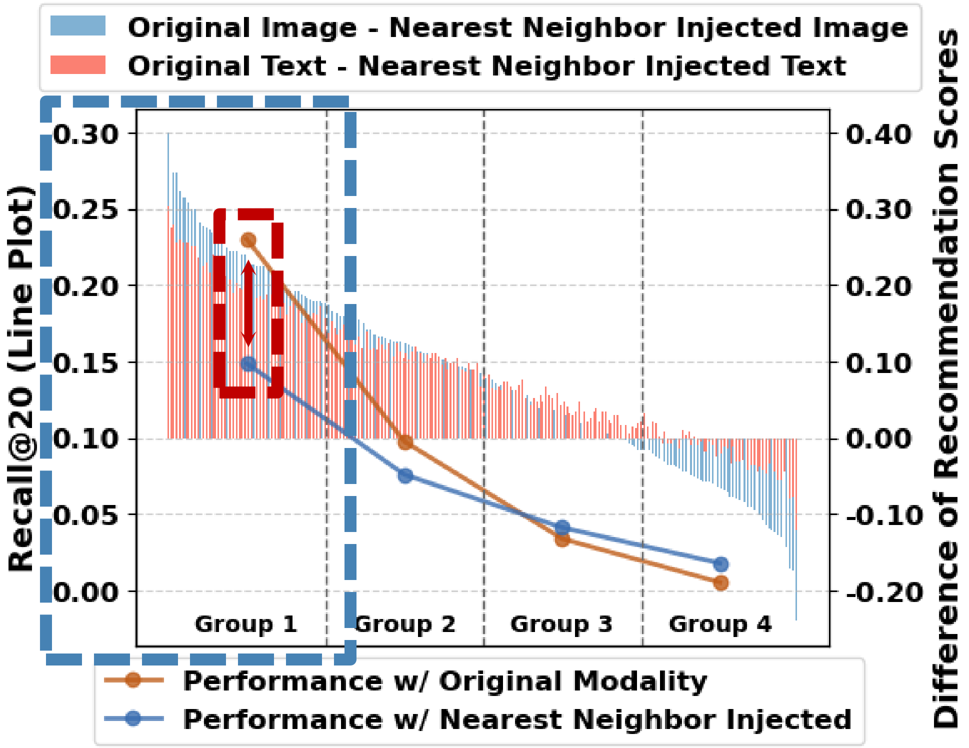


1. **Typically, synthetic features deviate from the original features of items.**

MOTIVATION

- **Multi-modal Recommendation System**

Missing Modality : In real-world scenarios, many items often have **more than one missing modality**.



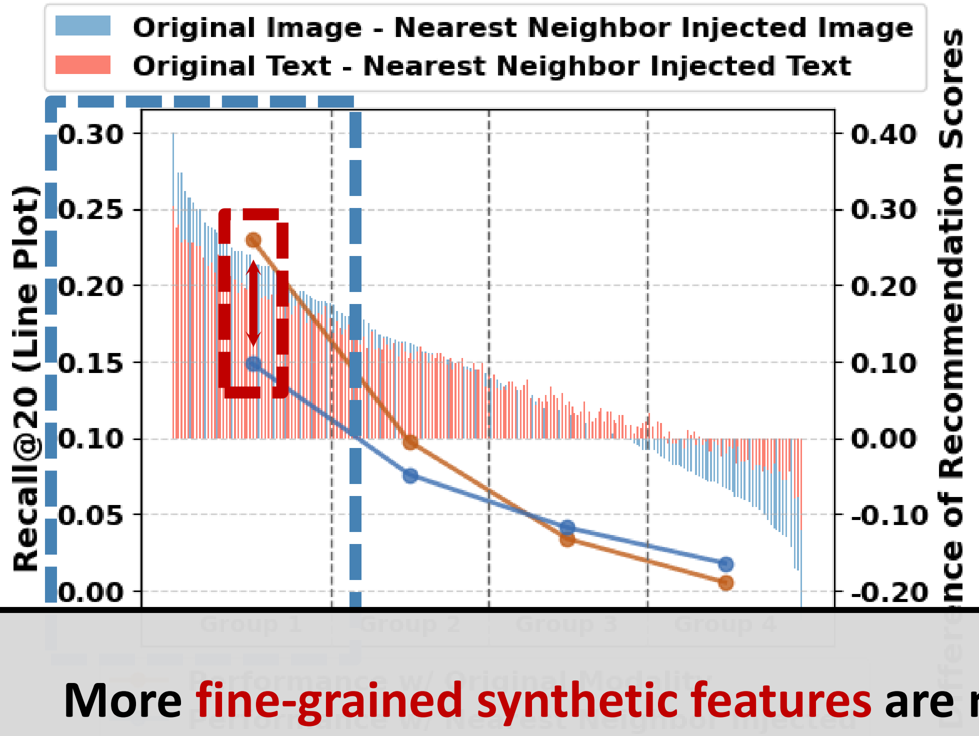
1. **Typically, synthetic features deviate from the original features of items.**

2. **Consequently, this leads to a more significant recommendation performance gap.**

MOTIVATION

- **Multi-modal Recommendation System**

  Missing Modality : In real-world scenarios, many items often have **more than one missing modality**.



1. **Typically, synthetic features deviate from the original features of items.**

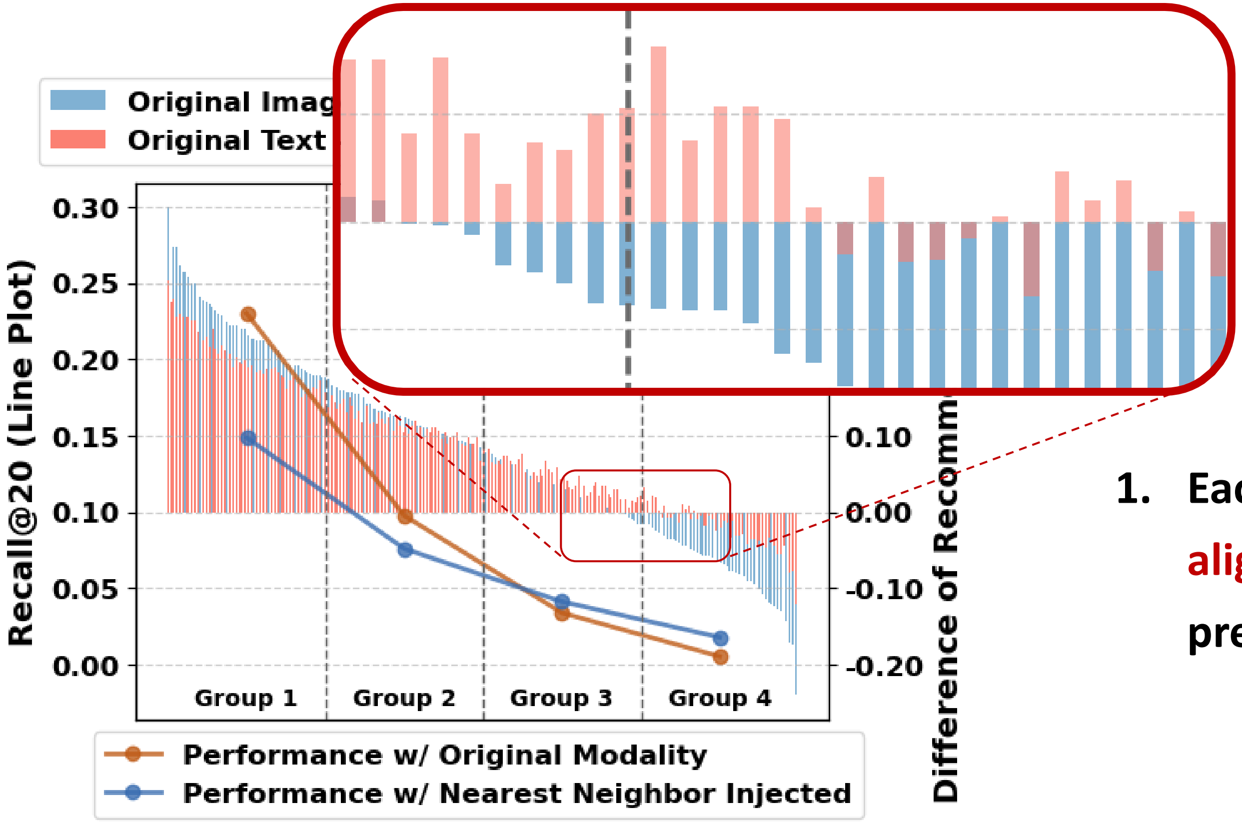
2. **Consequently, this leads to a more significant recommendation performance gap.**

More fine-grained synthetic features are needed to effectively substitute original features.

MOTIVATION

- **Multi-modal Recommendation System**

  Unique Characteristic of Modality : modalities possess unique characteristics not fully shared across others.

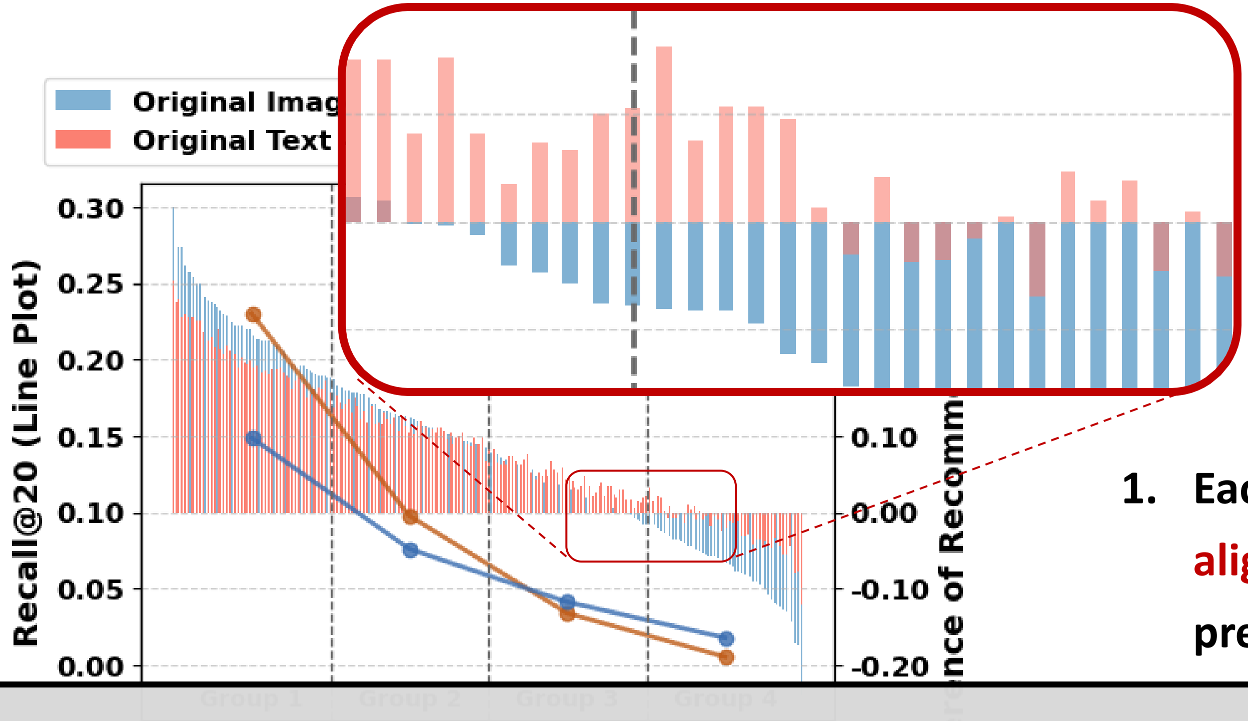


1. Each modality has **unique information which is not align with other modalities** when predicting user preferences.

MOTIVATION

- **Multi-modal Recommendation System**

Unique Characteristic of Modality : modalities possess unique characteristics not fully shared across others.



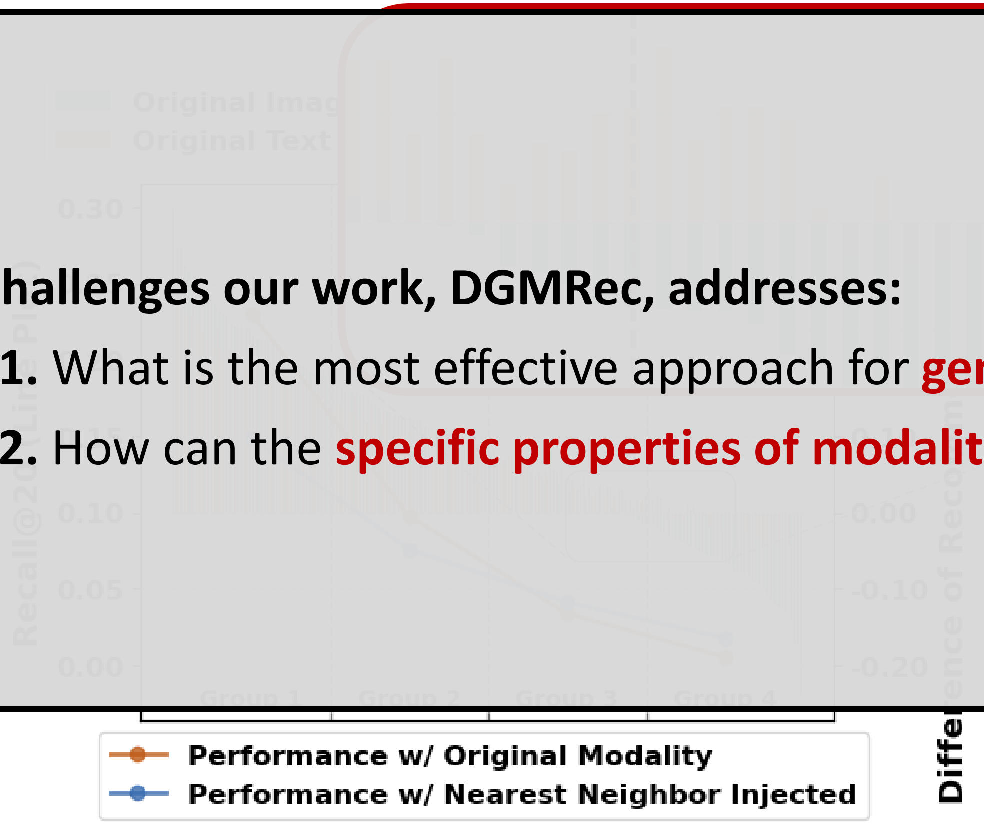
1. **Each modality has unique information which is not align with other modalities when predicting user preferences.**

Generating synthetic features requires consideration of each modality's specific properties.

MAIN CHALLENGES

- **Multi-modal Recommendation System**

Unique Characteristic of Modality : **modalities possess unique characteristics not fully shared across others.**



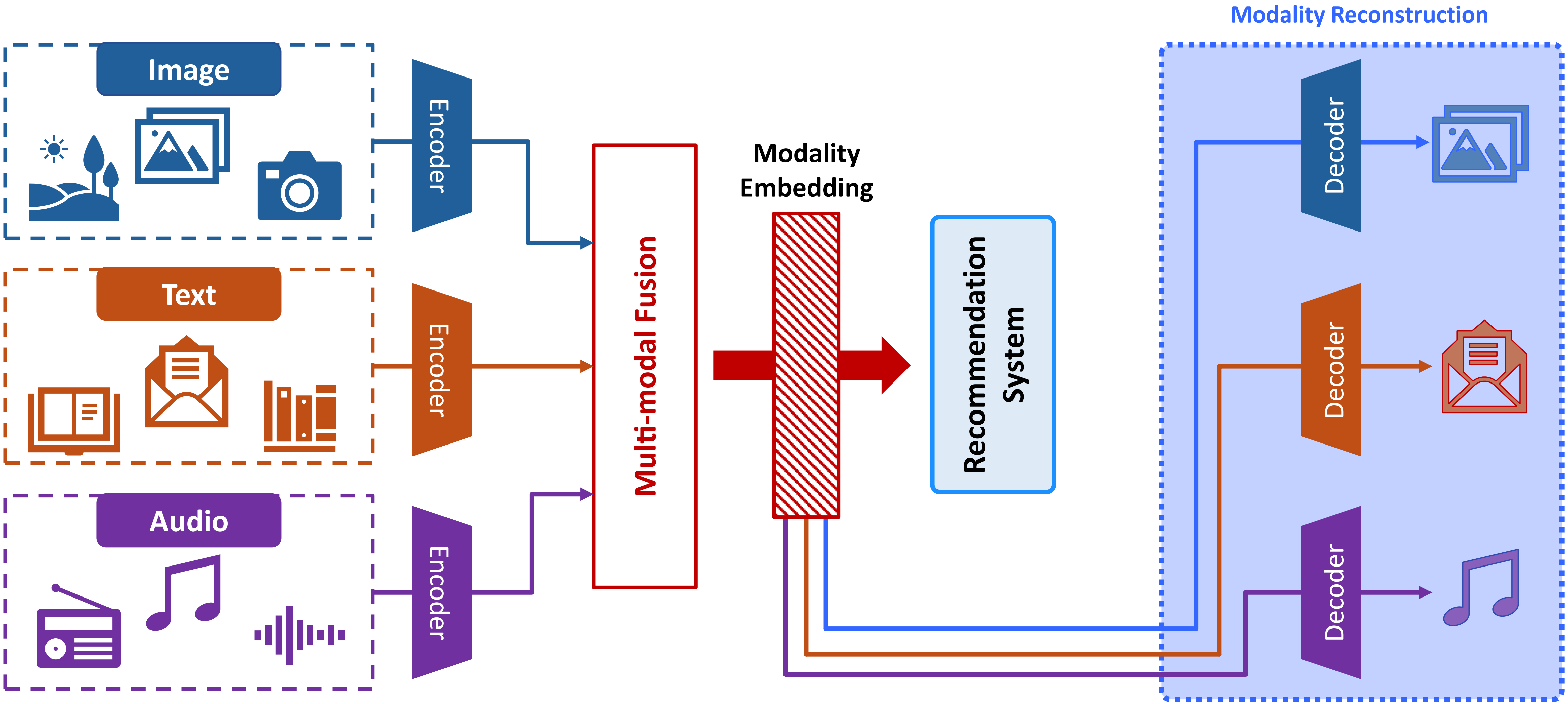
Main challenges our work, DGMRec, addresses:

- **C1.** What is the most effective approach for **generating synthetic features to address missing modalities?**
- **C2.** How can the **specific properties of modalities** be incorporated for fine-grained representations?

1. Each modality has unique information which is not align with other modalities when predicting user preferences.
2. Therefore, <u>generating synthetic features requires consideration of each modality's specific properties.</u>

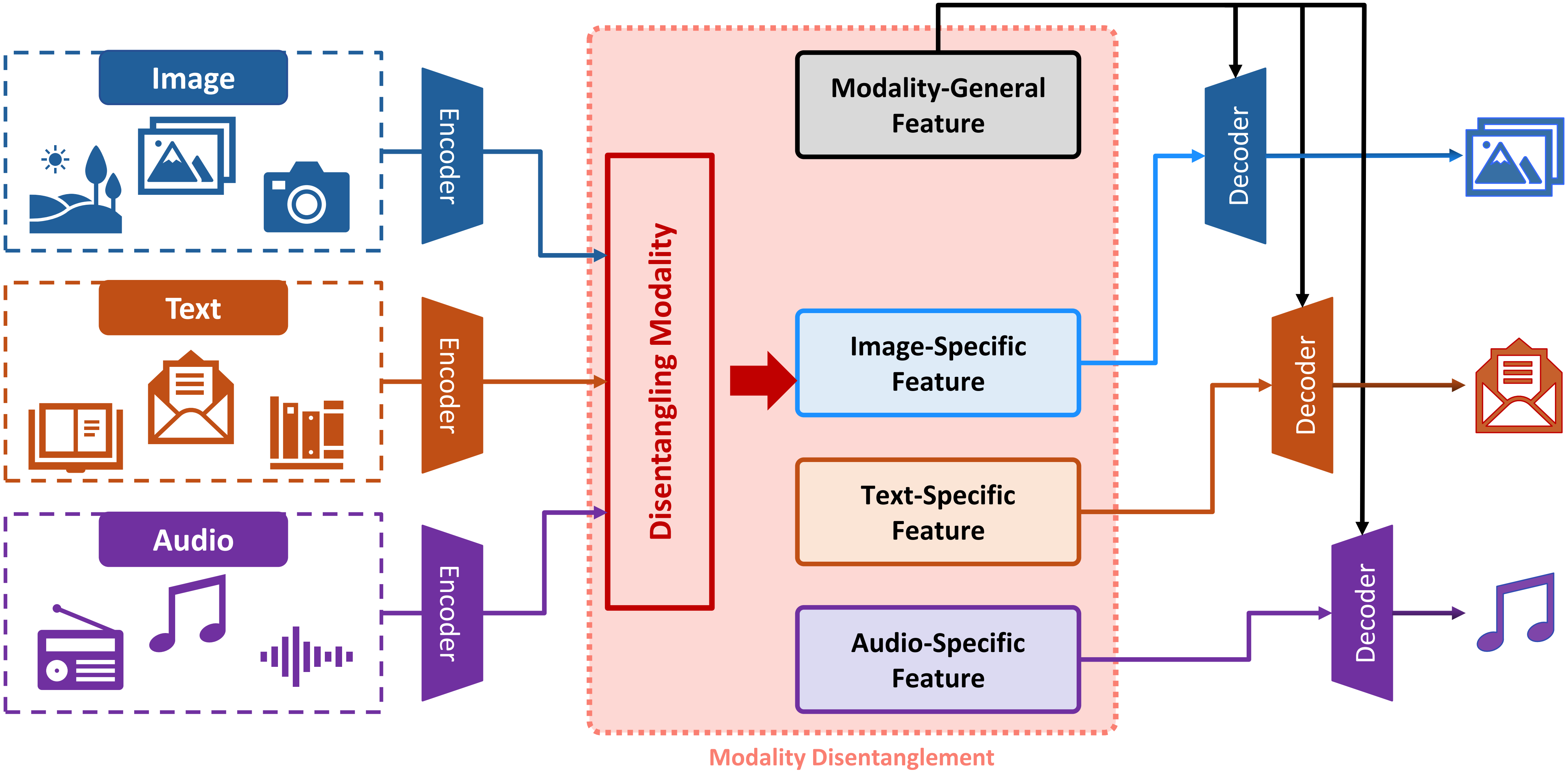
C1. HOW TO GENERATE SYNTHETIC FEATURES F0R MISSING MODALITIES?

- **Encoder-Decoder Structure for Reconstructing Modality**



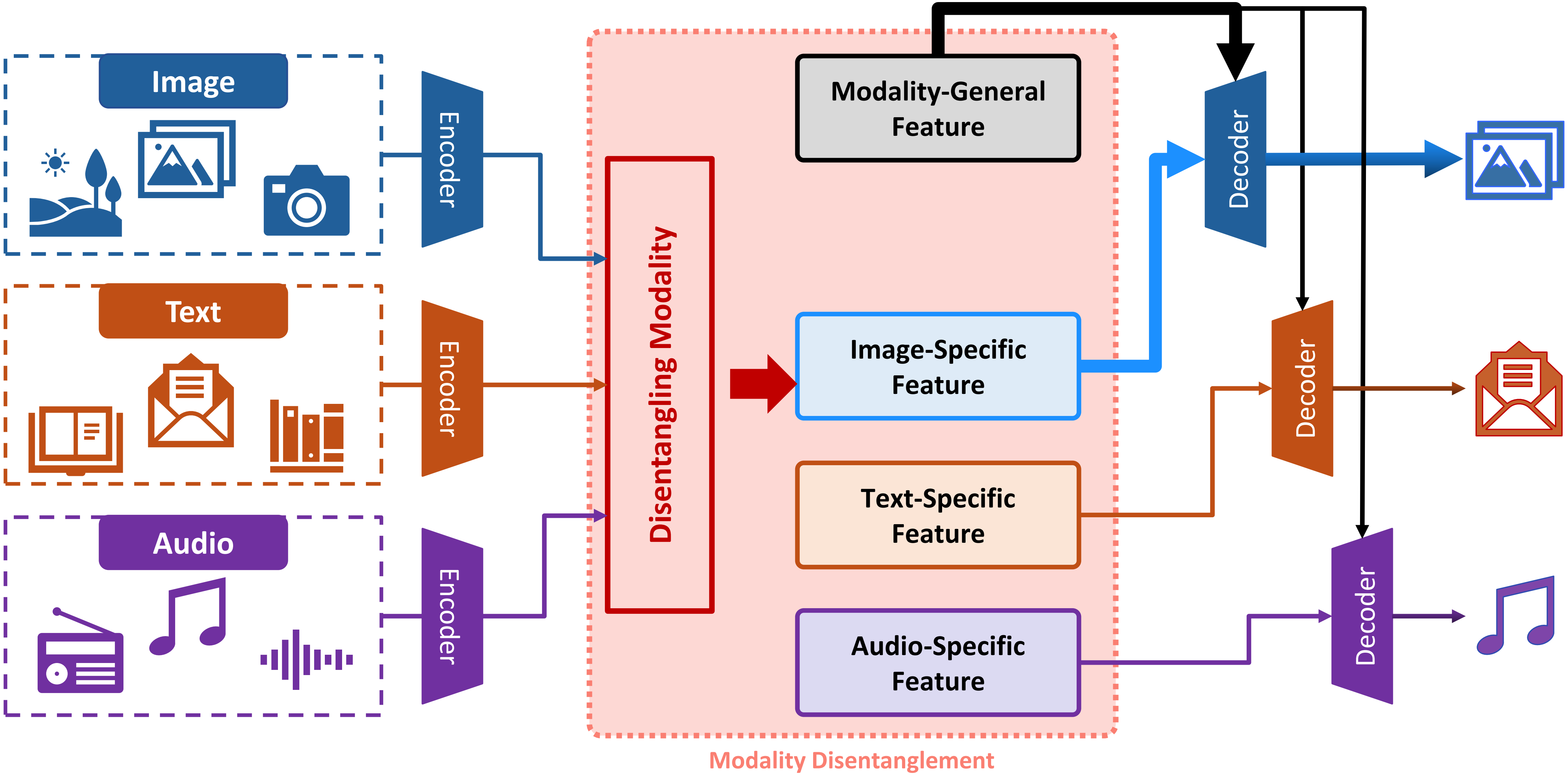
C2. HOW CAN SPECIFIC MODALITY FEATURE BE INCORPORATED?

- **Disentangling Modality Features for General and Specific Features**



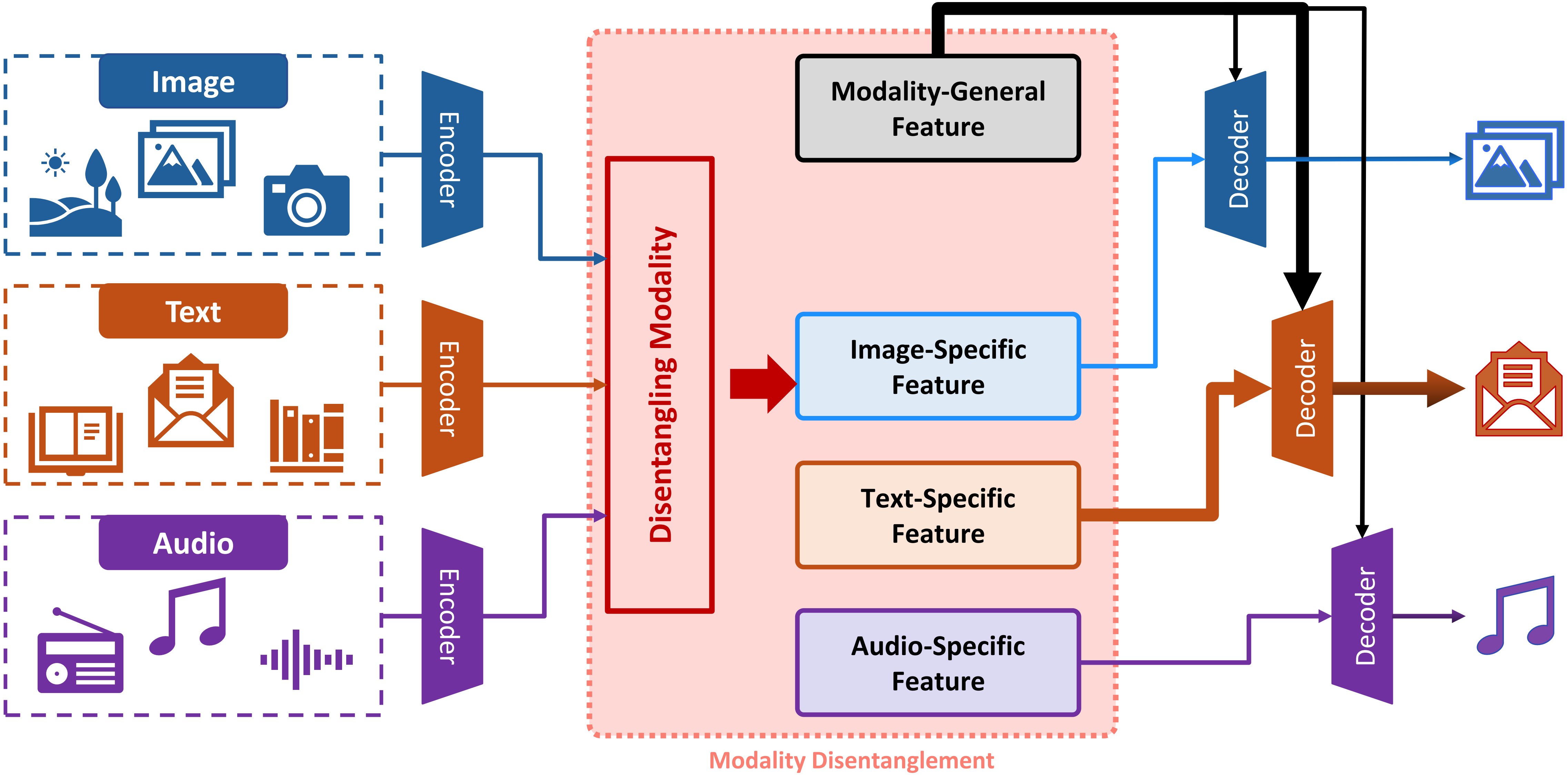
C2. HOW CAN SPECIFIC MODALITY FEATURE BE INCORPORATED?

- **Disentangling Modality Features for General and Specific Features**



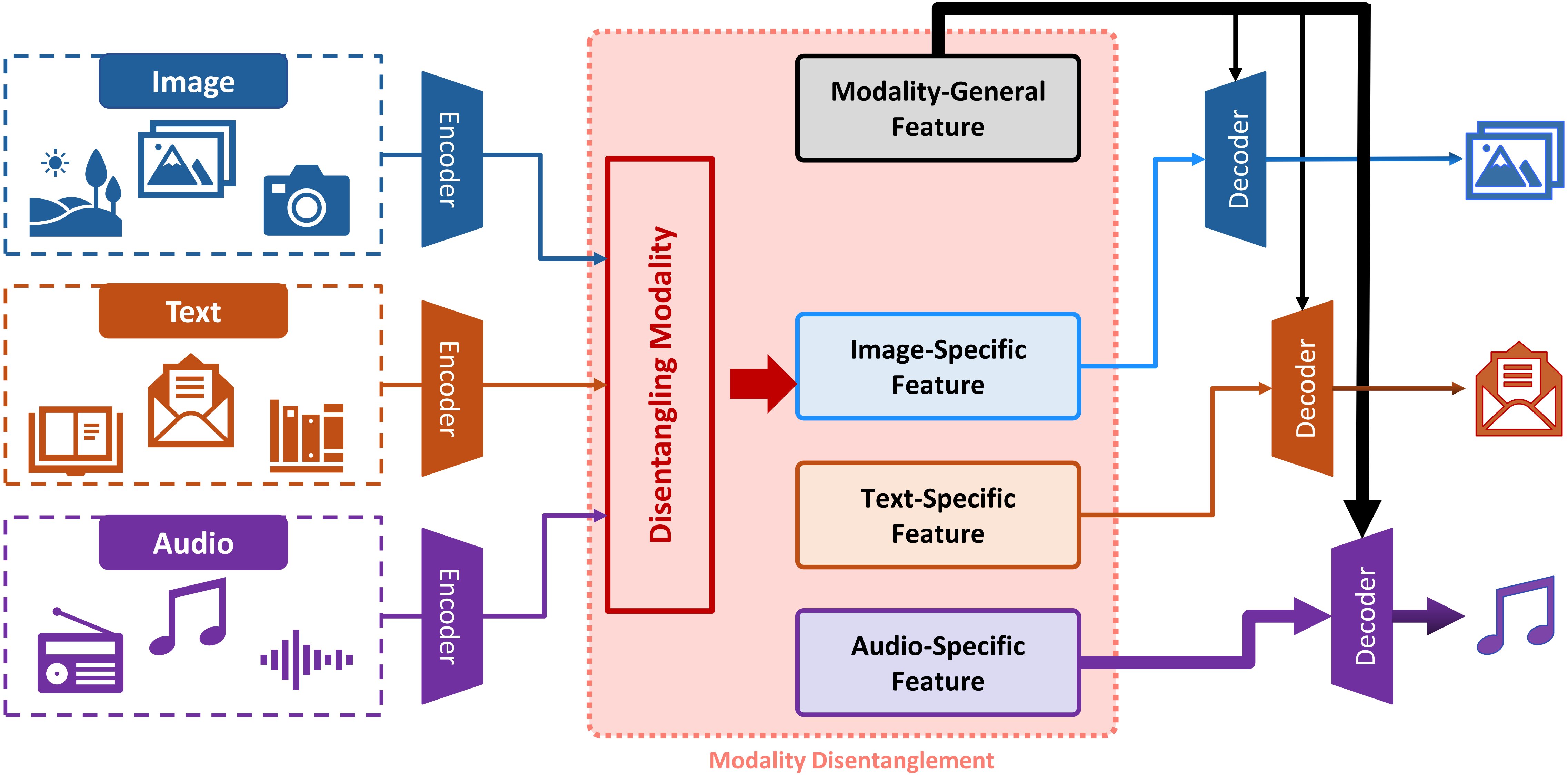
C2. HOW CAN SPECIFIC MODALITY FEATURE BE INCORPORATED?

- **Disentangling Modality Features for General and Specific Features**



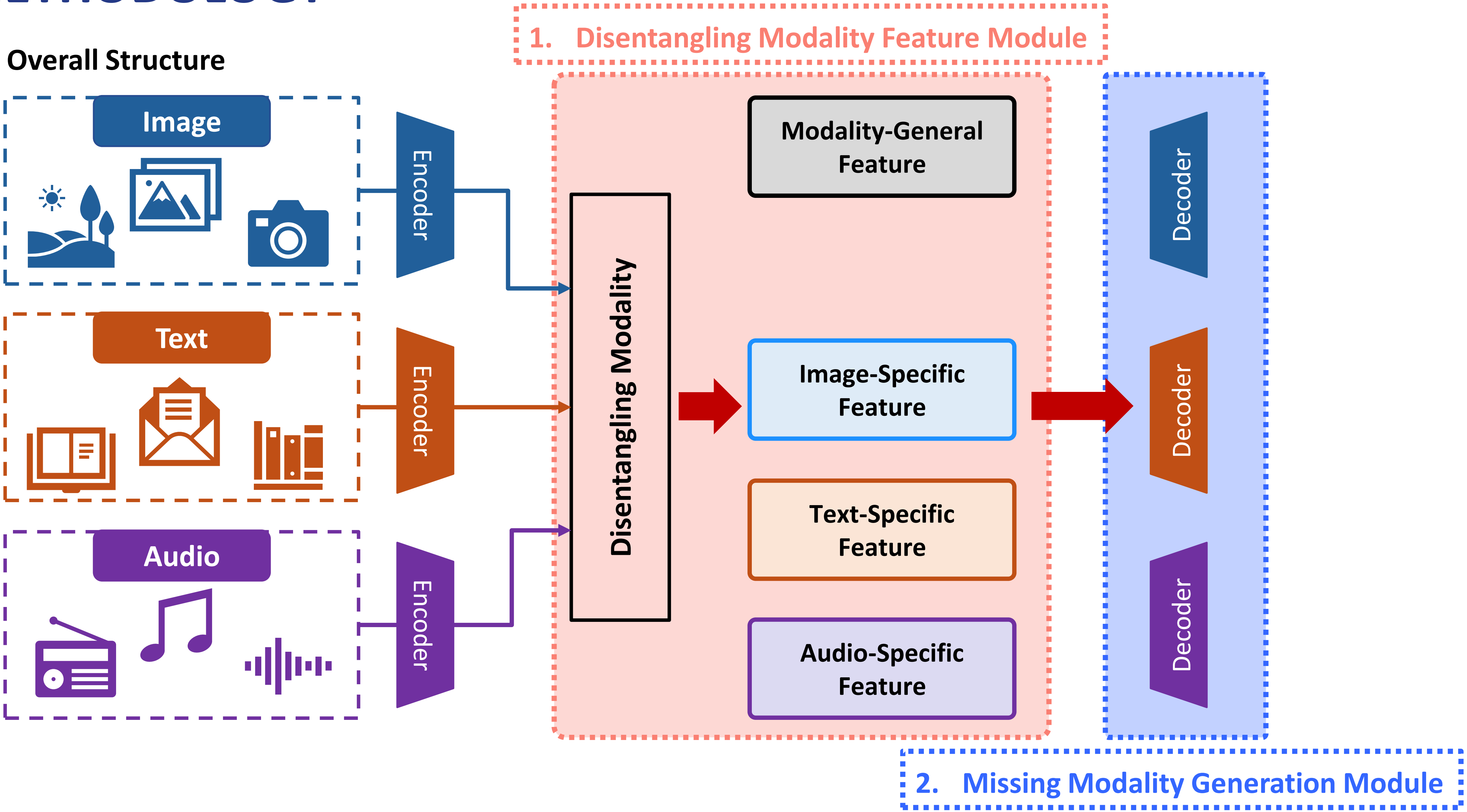
C2. HOW CAN SPECIFIC MODALITY FEATURE BE INCORPORATED?

- Disentangling Modality Features for General and Specific Features



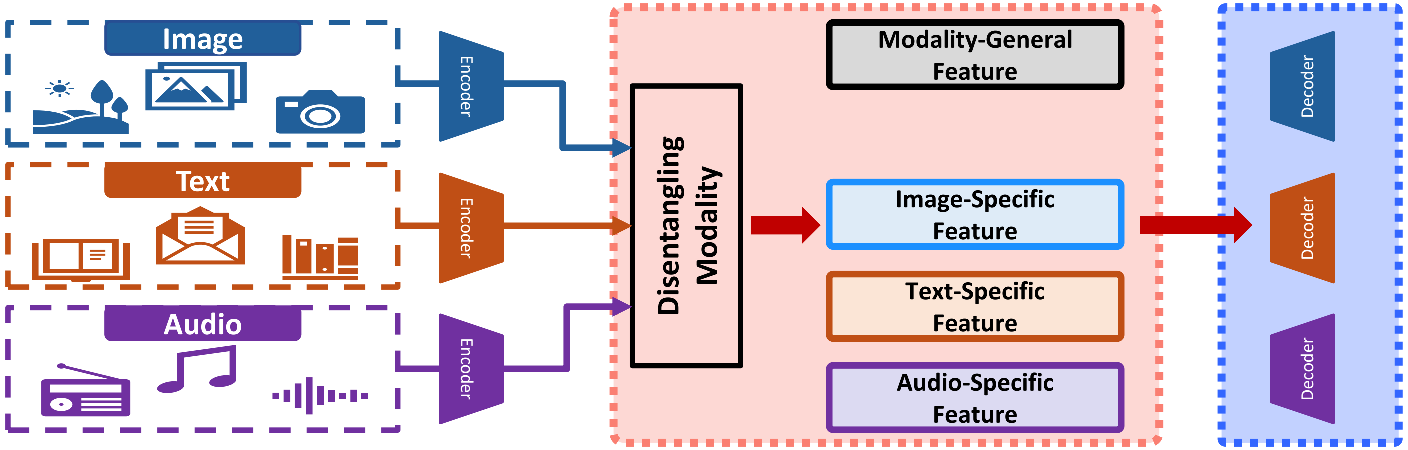
METHODOLOGY

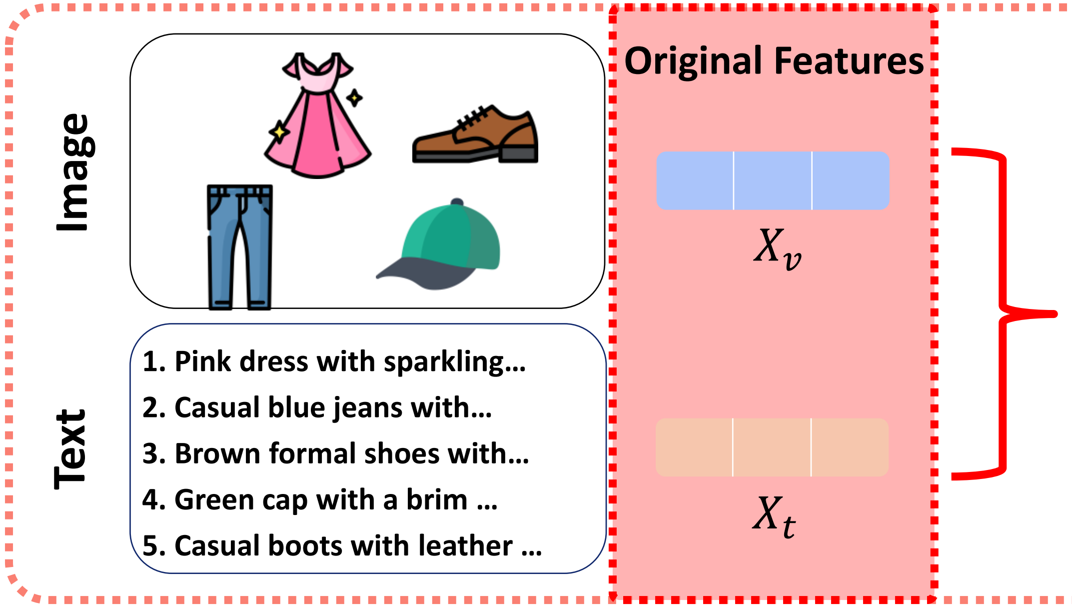
- Overall Structure



METHODOLOGY

- **1. Disentangling Modality Feature Module**



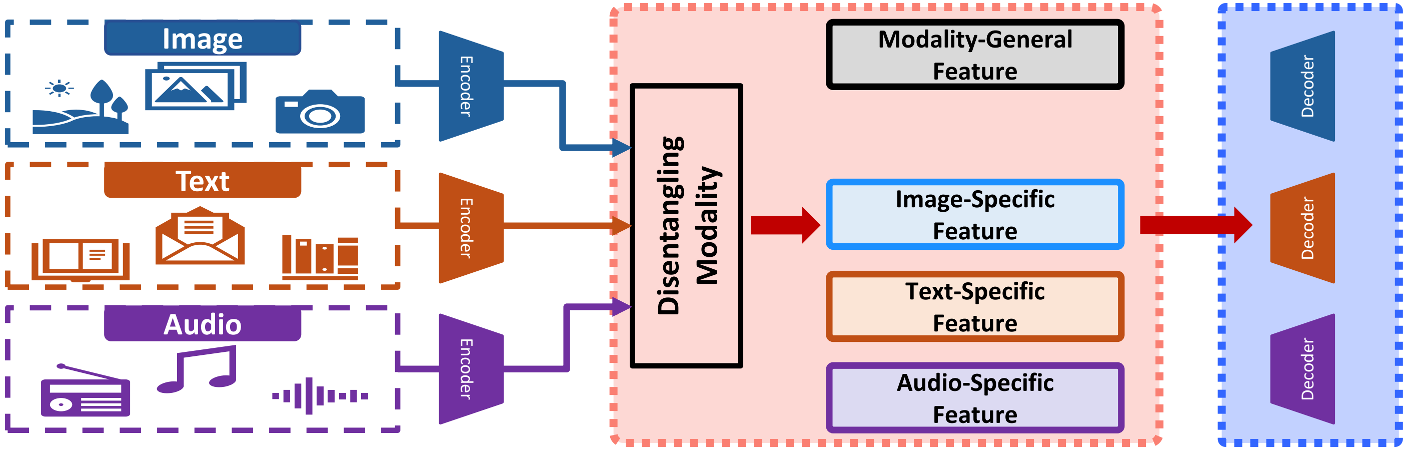


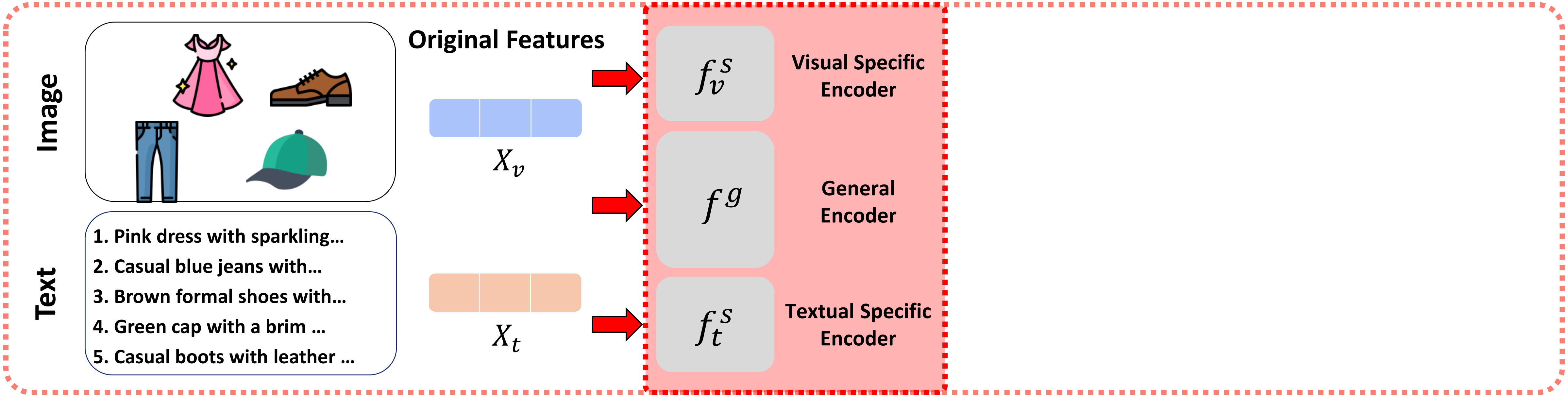
Each modality's features are extracted from general-purpose encoder

e.g., CNN for visual, SBERT for textual

METHODOLOGY

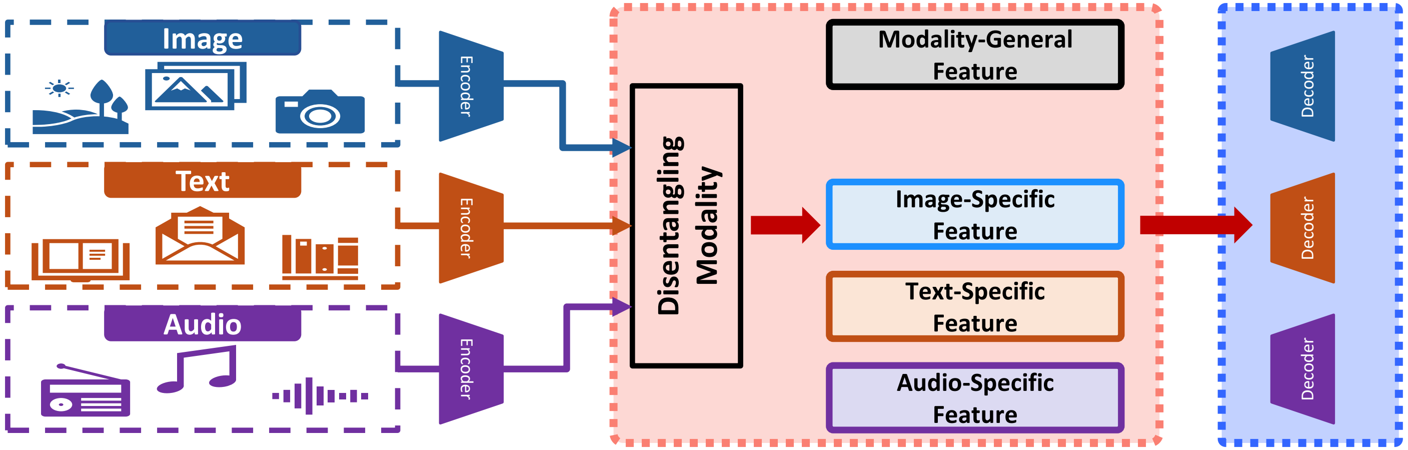
- 1. Disentangling Modality Feature Module

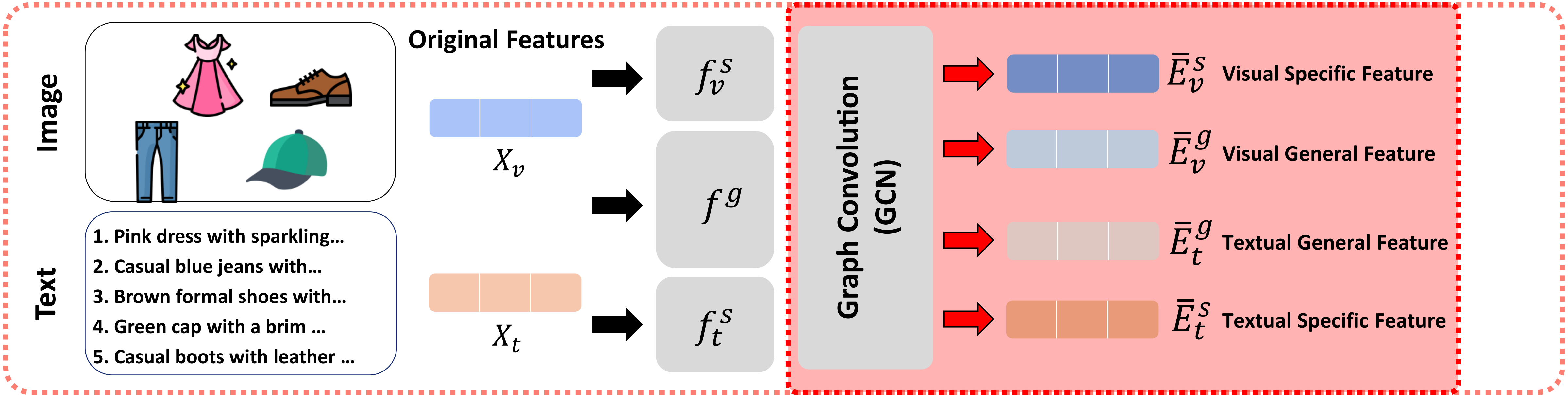




METHODOLOGY

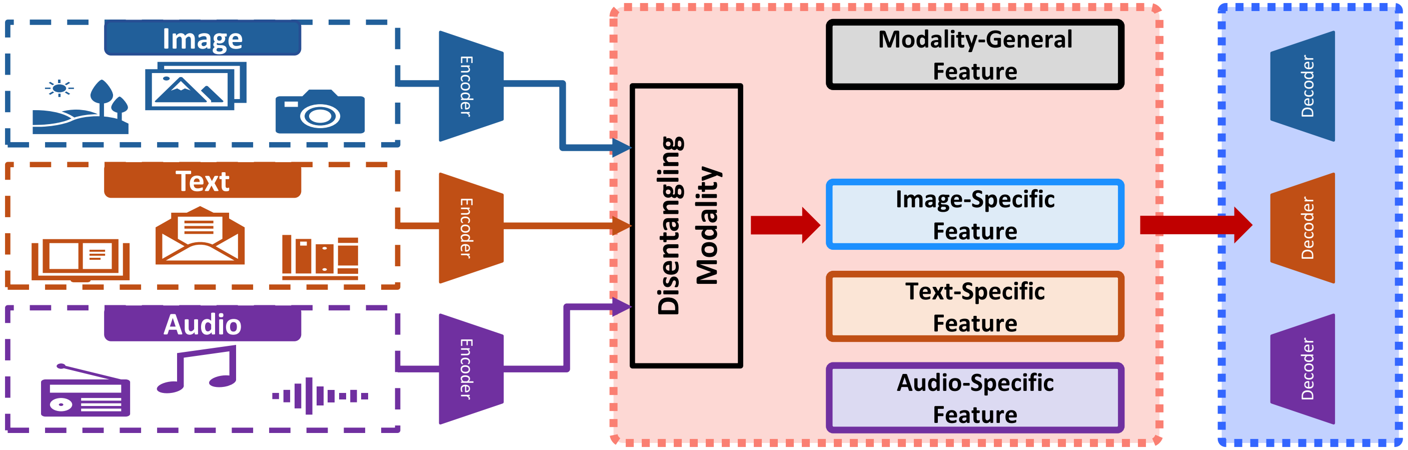
- 1. Disentangling Modality Feature Module

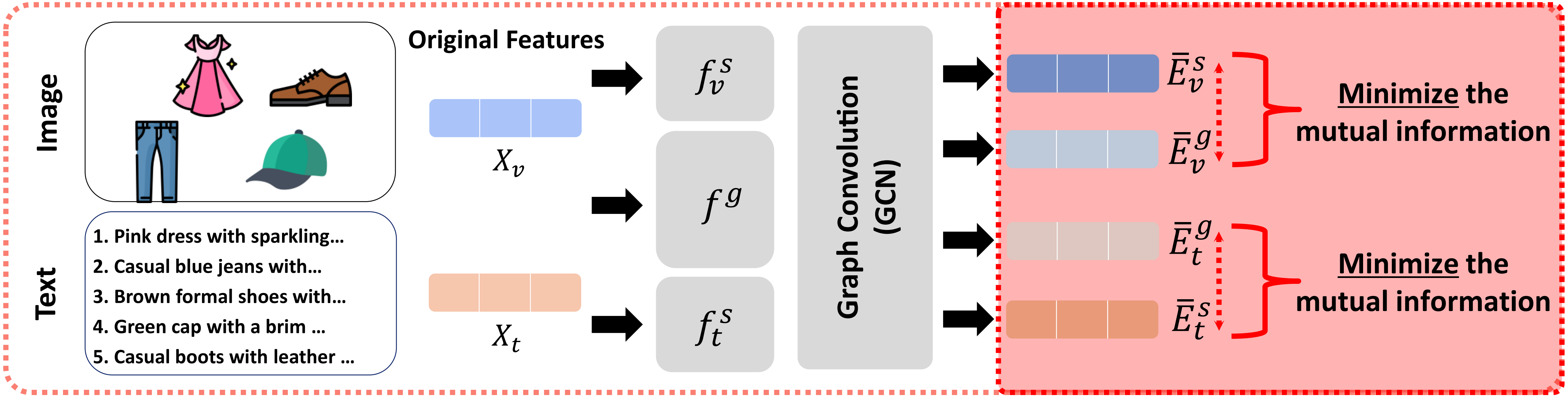




METHODOLOGY

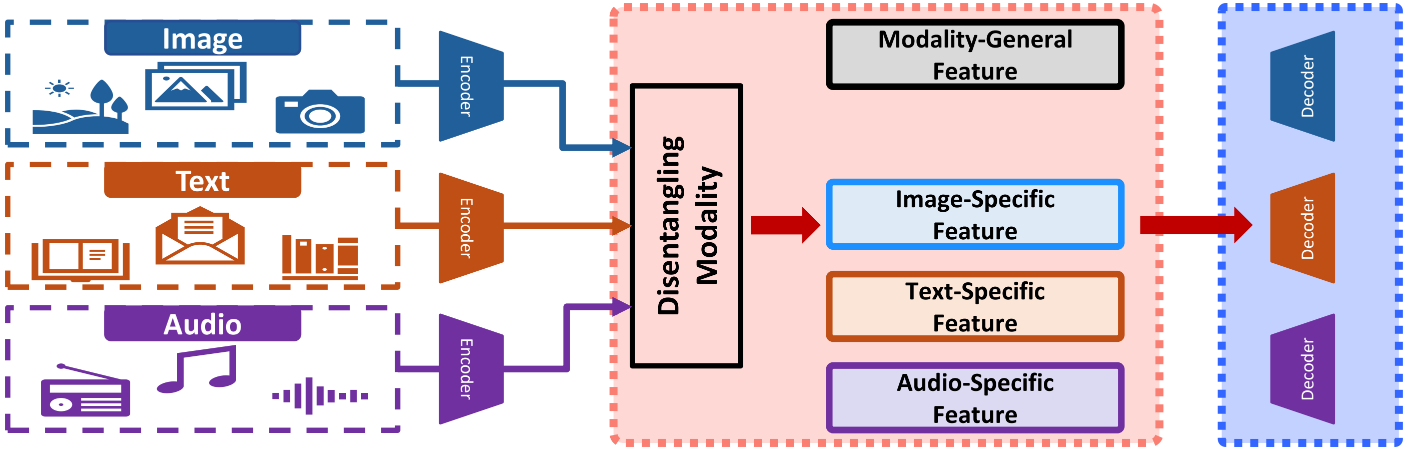
- 1. Disentangling Modality Feature Module

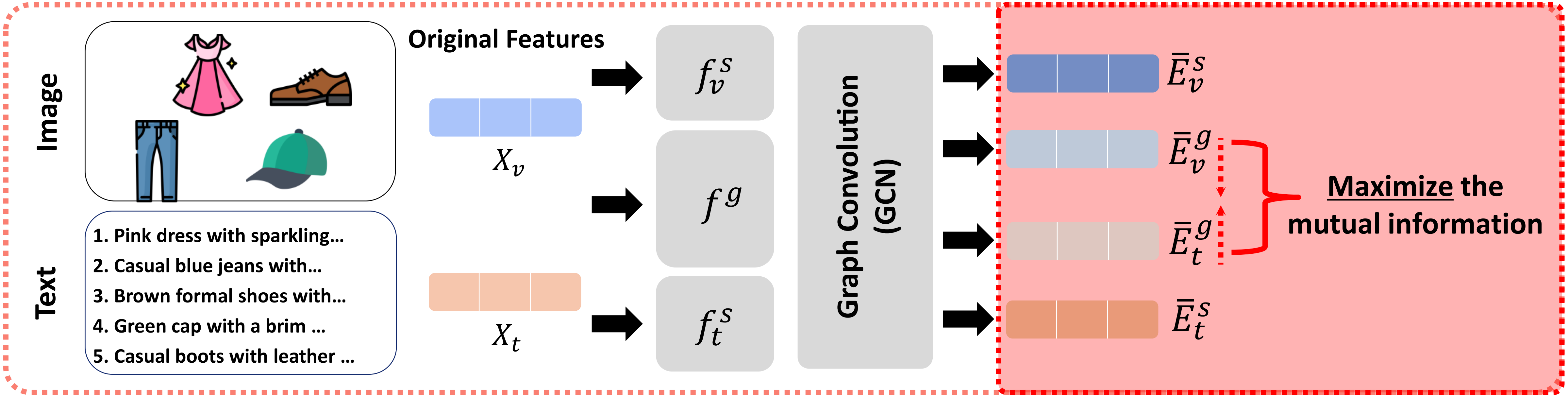




METHODOLOGY

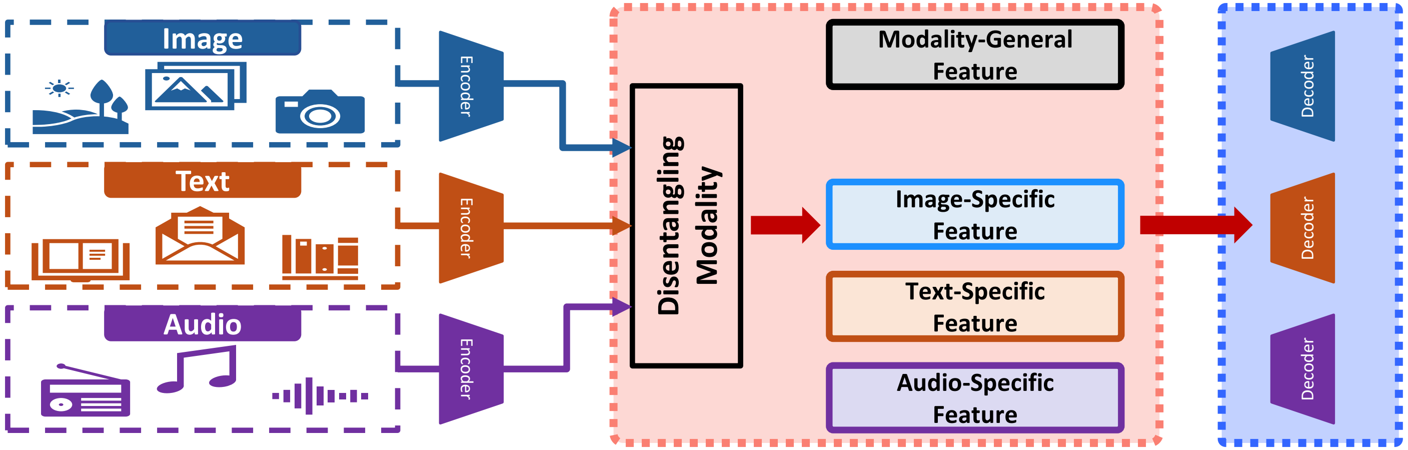
- 1. Disentangling Modality Feature Module

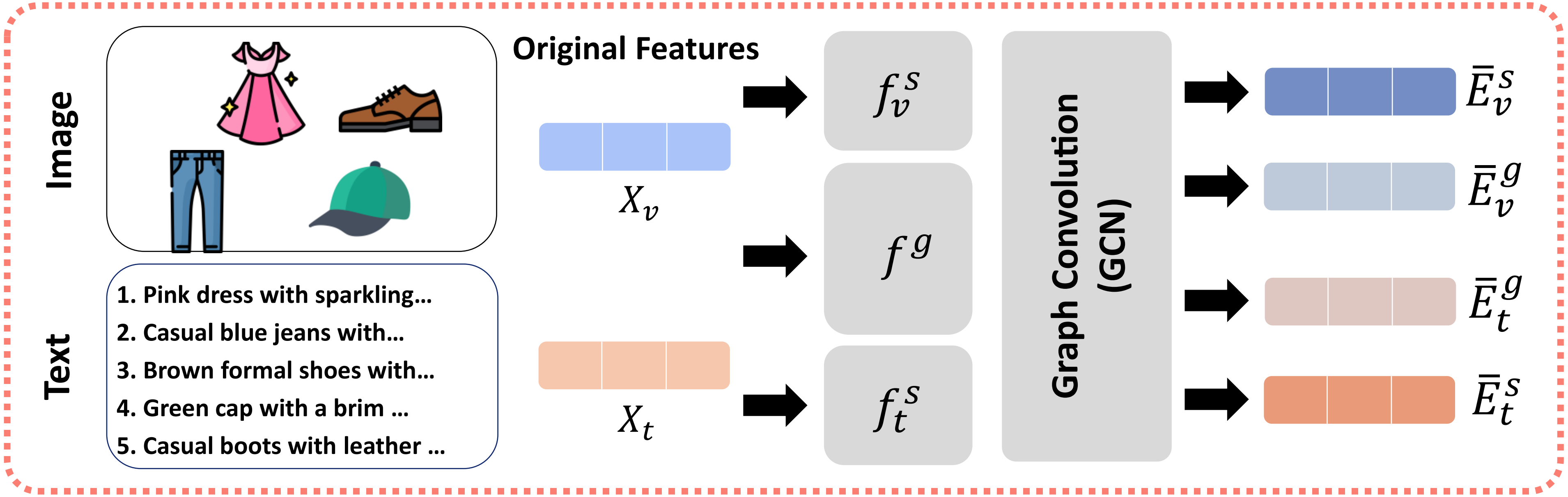




METHODOLOGY

- 1. Disentangling Modality Feature Module

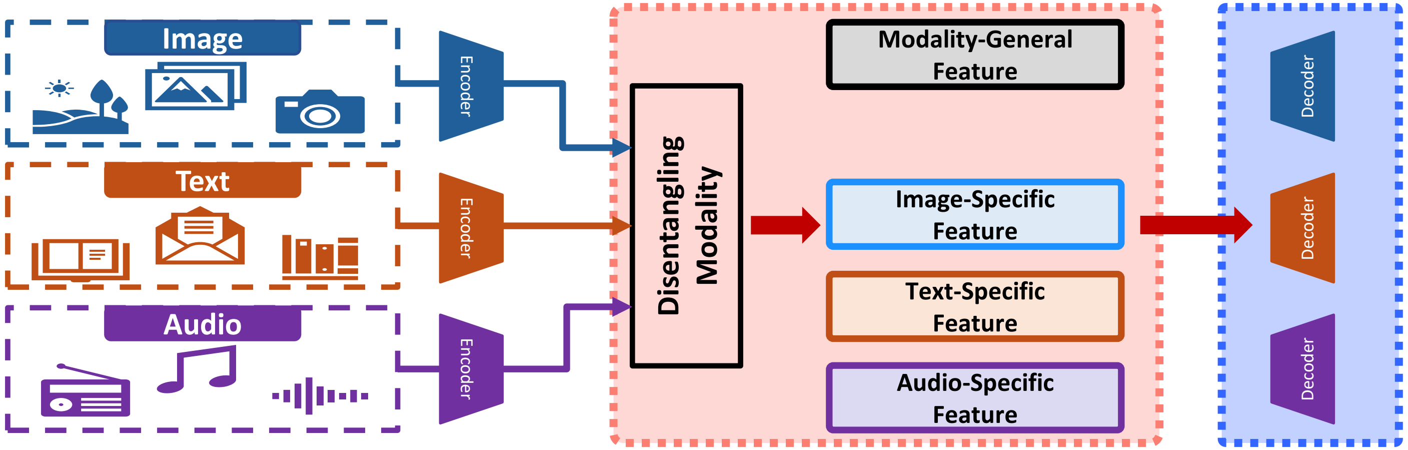


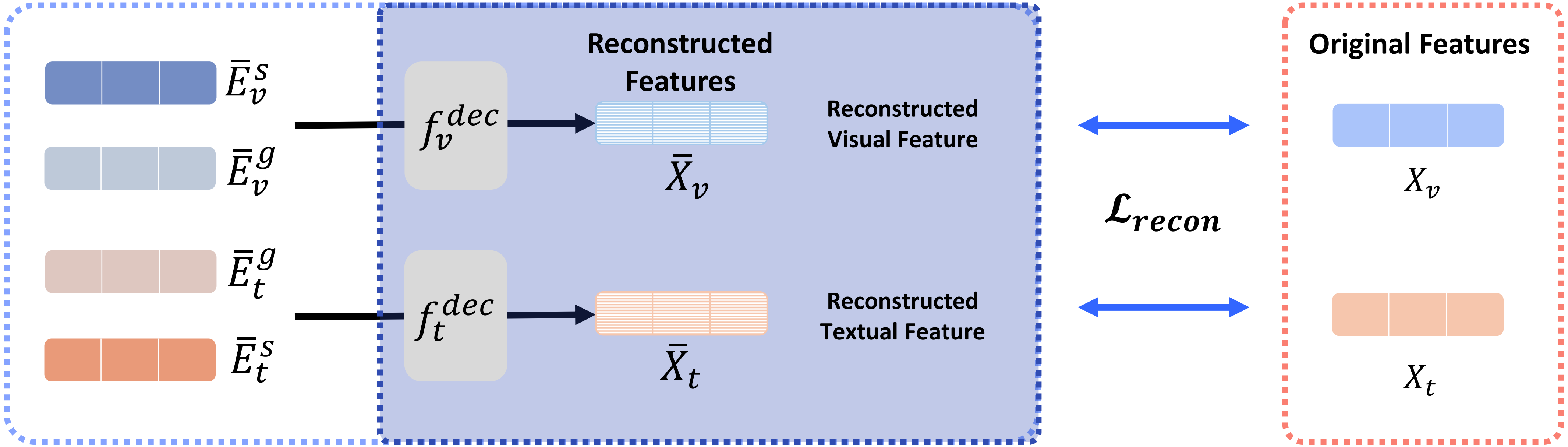


General features share **common information** across modalities.
In contrast, **specific features** capture the **distinct characteristics** of each modality.

METHODOLOGY

- **2. Missing Modality Generation Module**



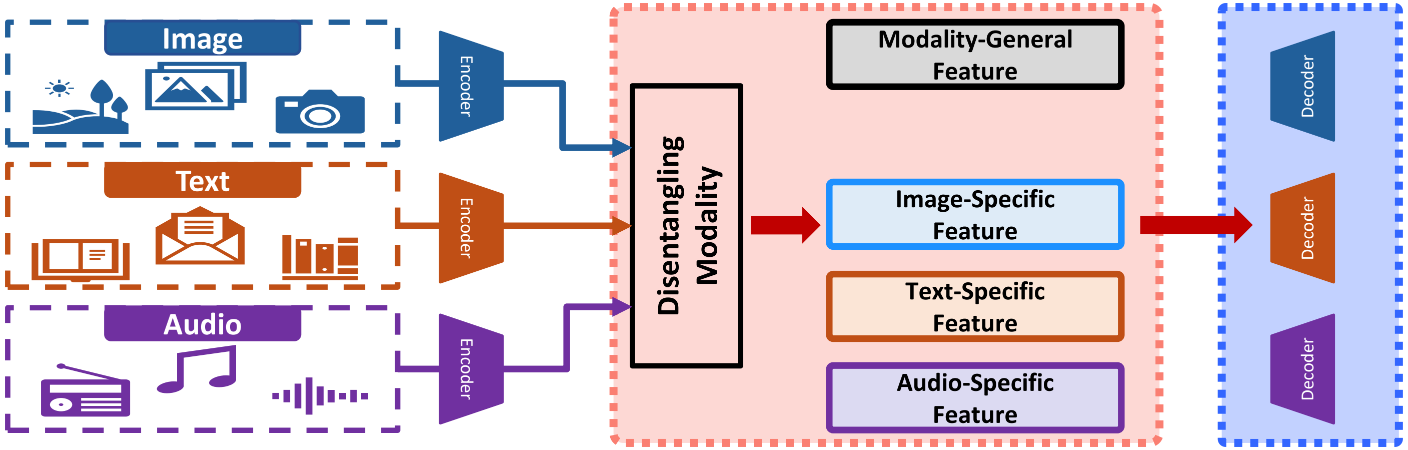


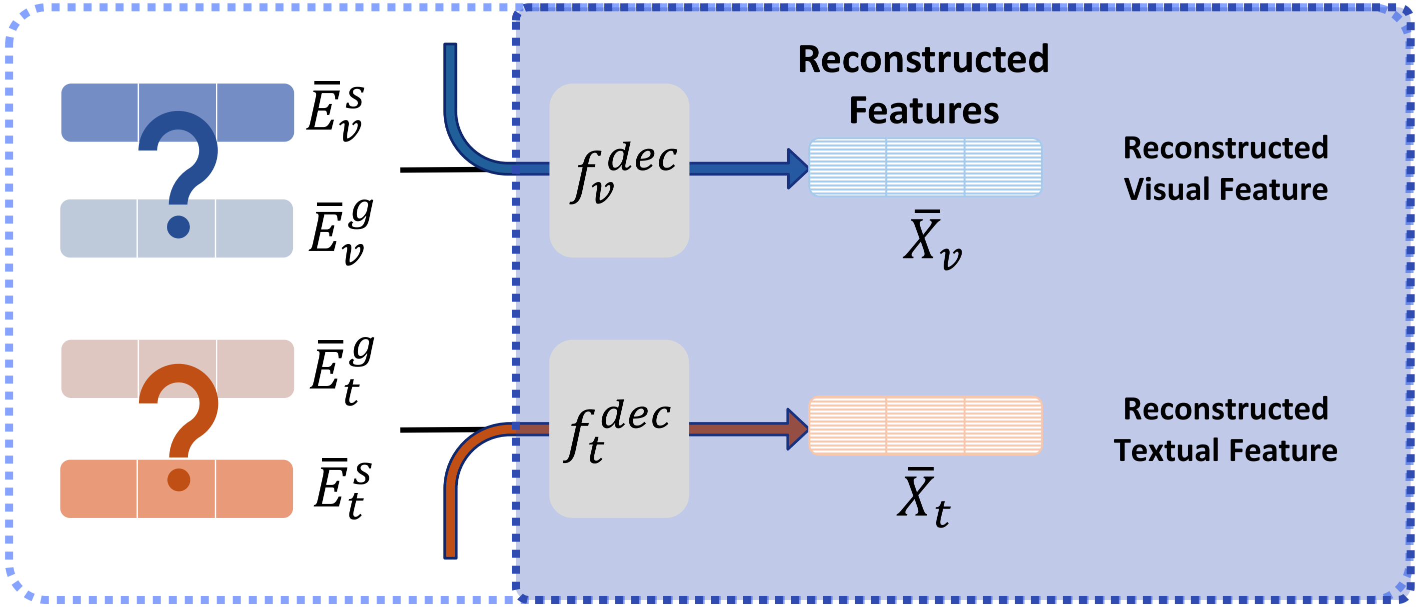
- *1. Disentangling Modality Feature Module*

METHODOLOGY

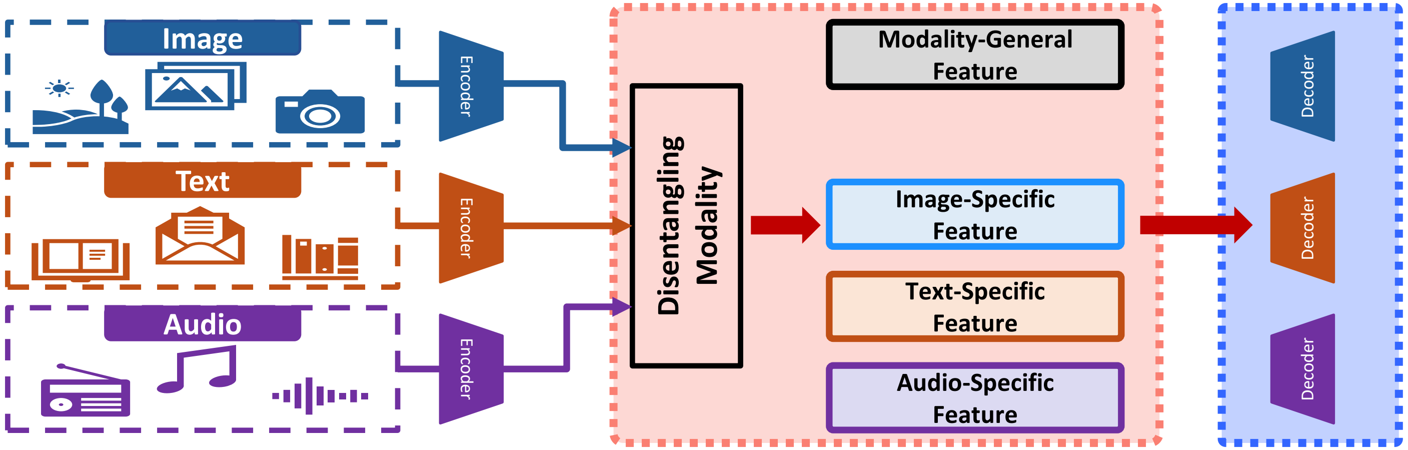
- **2. Missing Modality Generation Module**

However, Items with missing modalities don't have modality features.





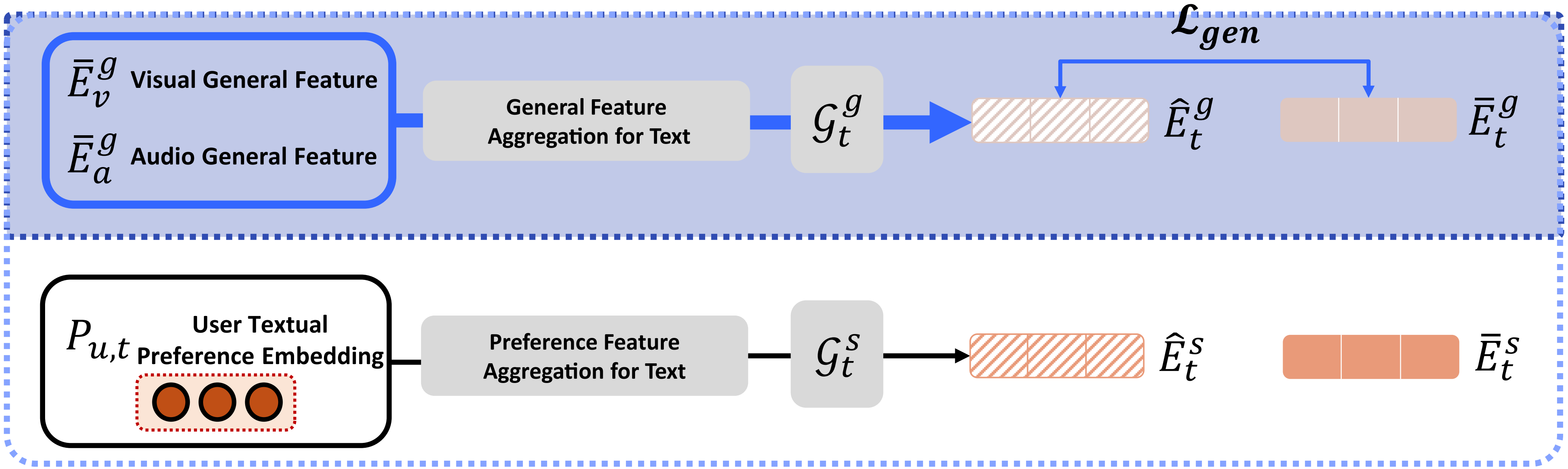
METHODOLOGY



- **2. Missing Modality Generation Module**

Train modality **general** / specific generator.

Assume textual modality generator training.

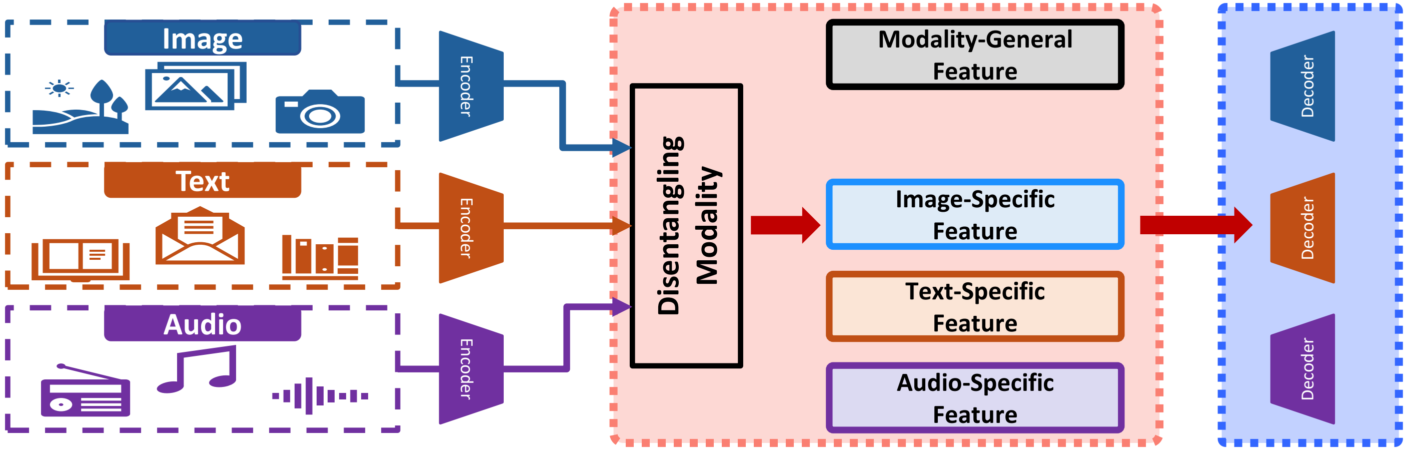


For modality **general** feature, **DGMRec** utilizes **other modalities' general features** which **share common information.**

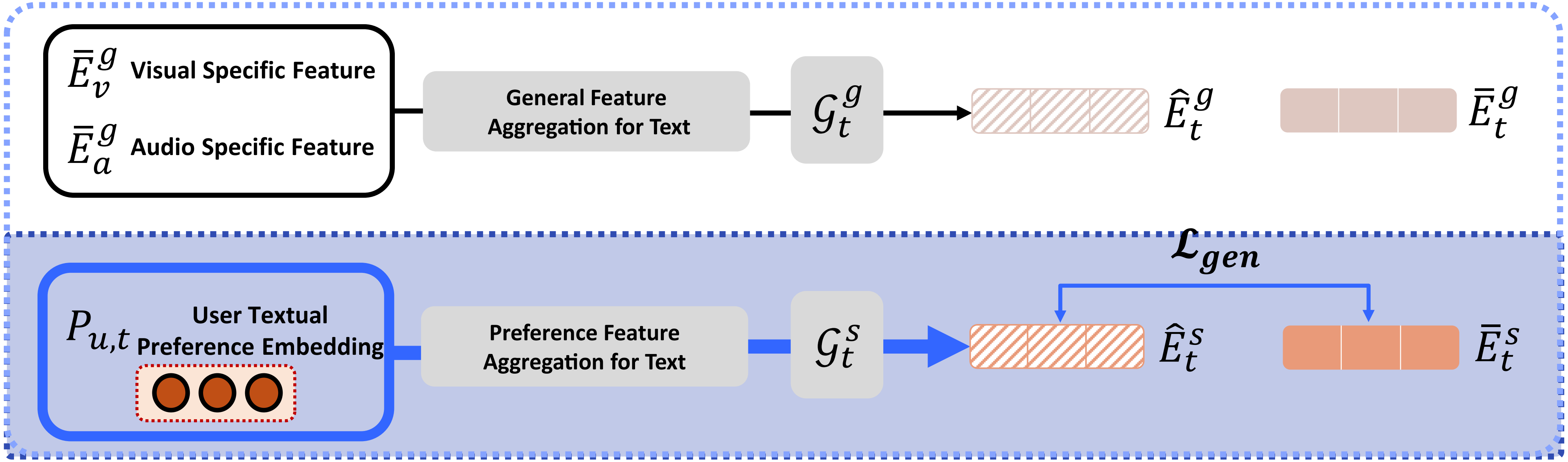
METHODOLOGY

- **2. Missing Modality Generation Module**

Train modality general / **specific** generator.



Assume textual modality generator training.

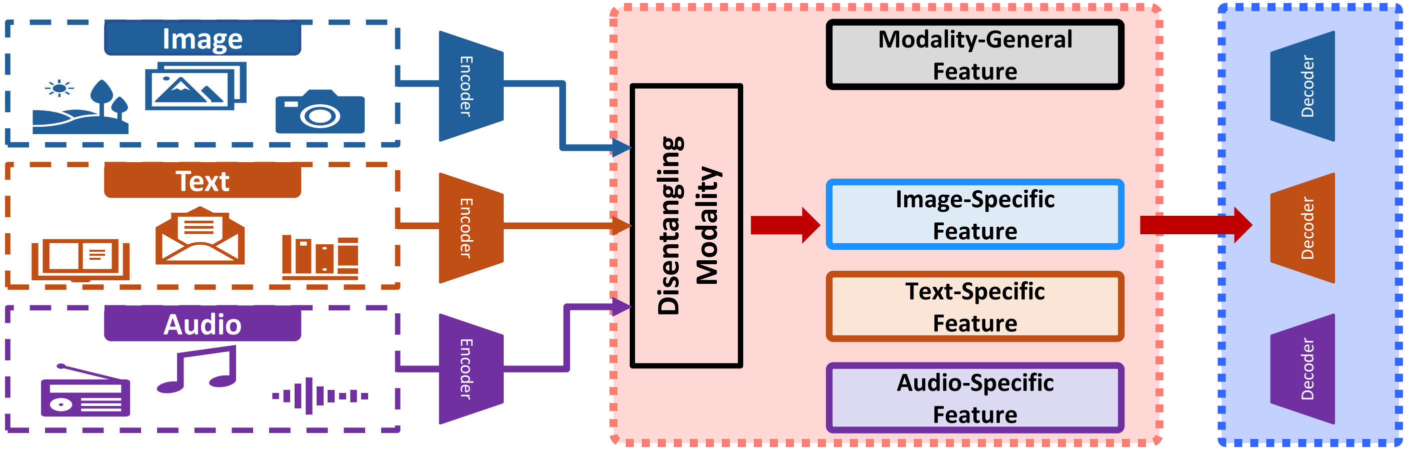


For modality **specific** feature, **DGMRec** utilizes **user's modality preference embedding** which **modality specific information.**

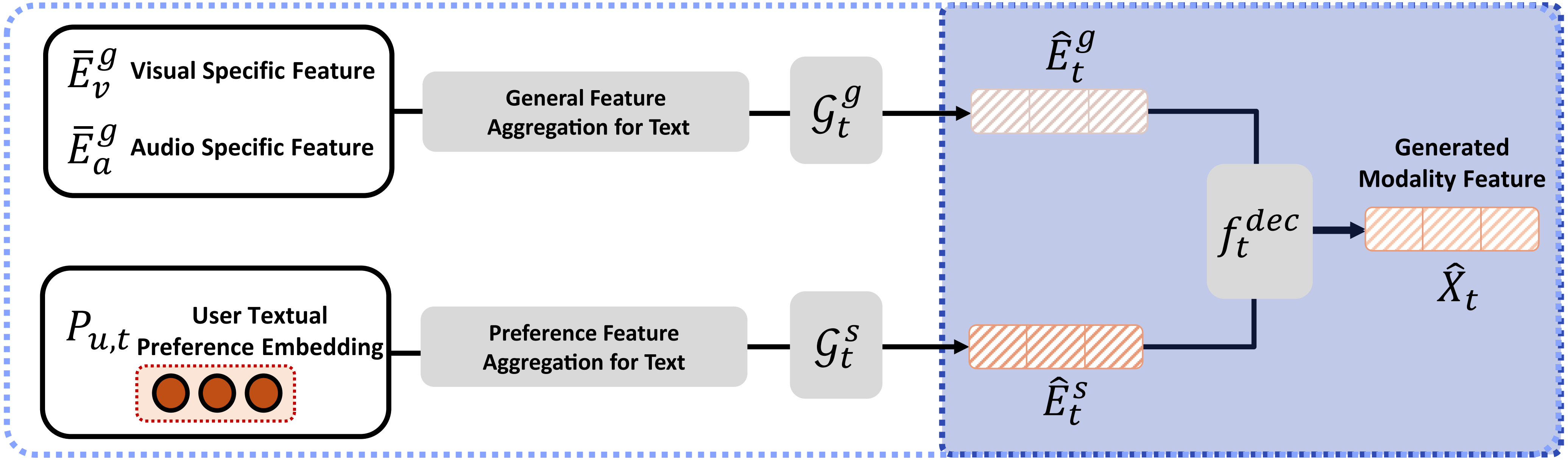
METHODOLOGY

- **2. Missing Modality Generation Module**

For items with missing modalities,



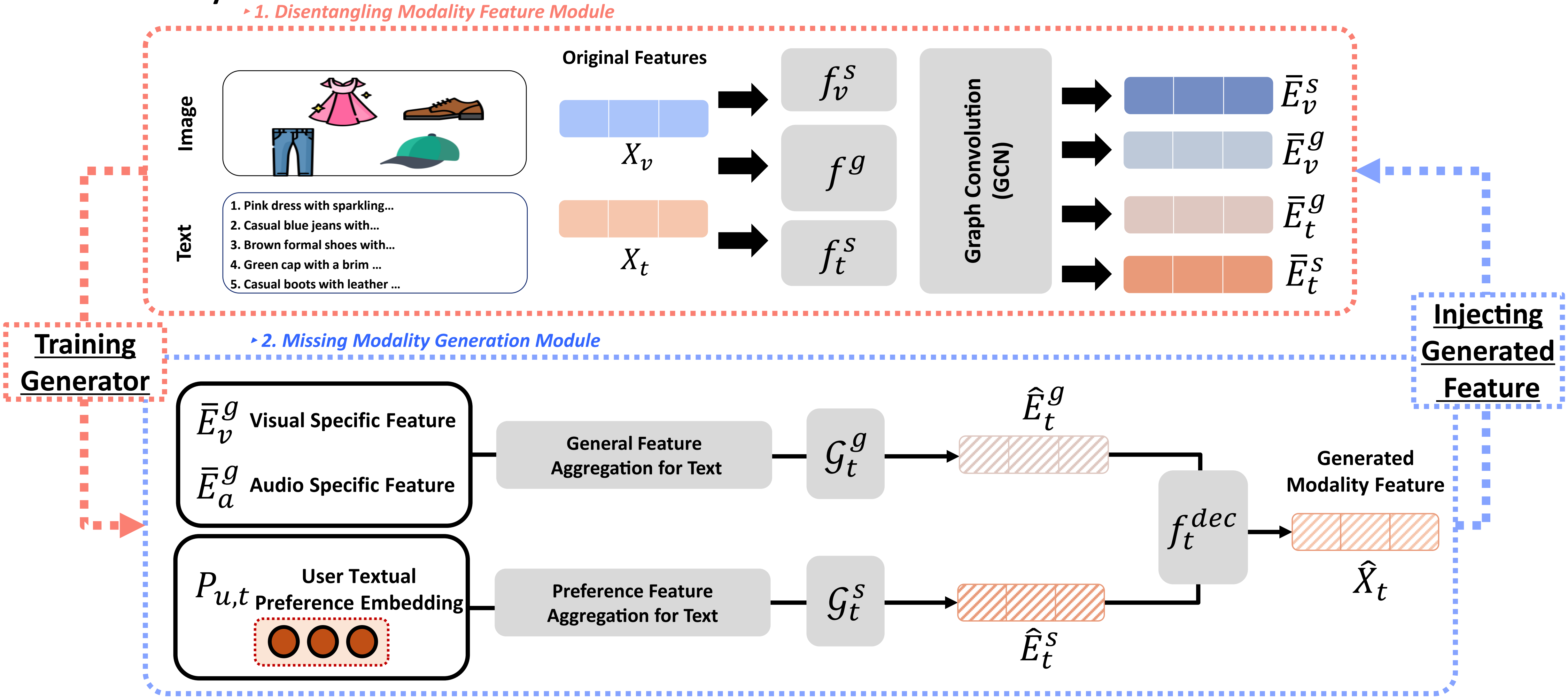
When textual modality is missing,



DGMRec can **generate modality feature** which captures general and specific information.

METHODOLOGY

- Summary



EXPERIMENTS

- **Model Performance with Real-world Scenario**

1. Missing Modality Setting : Item's modalities is randomly missing

The __number__ and __type__ of missing modalities are randomly determined.

Missing Modality Setting																	
Dataset		Baby				Sports				Clothing				TikTok			
Metric		R@20	R@50	N@20	N@50	R@20	R@50	N@20	N@50	R@20	R@50	N@20	N@50	R@20	R@50	N@20	N@50
Multi-Modal Recommenders	VBPR	0.0514	0.0937	0.0213	0.0299	0.0741	0.1229	0.0328	0.0427	0.0462	0.0737	0.0207	0.0226	0.0410	0.0699	0.0172	0.0229
	MMGCN	0.0519	0.0991	0.0215	0.0310	0.0509	0.0913	0.0215	0.0297	0.0289	0.0530	0.0120	0.0168	0.0883	0.1431	0.0372	0.0484
	GRCN	0.0644	0.1151	0.0274	0.0377	0.0681	0.1157	0.0300	0.0397	0.0381	0.0644	0.0161	0.0214	0.0716	0.1257	0.0283	0.0389
	SLMRec	0.0753	0.1254	0.0340	0.0422	0.0914	0.1462	0.0415	0.0526	0.0624	0.0979	0.0281	0.0351	0.0932	0.1523	0.0364	0.0480
	BM3	0.0683	0.1235	0.0296	0.0408	0.0908	0.1466	0.0400	0.0513	0.0591	0.0920	0.0268	0.0334	0.0768	0.1215	0.0322	0.0409
	LATTICE	0.0738	0.1297	0.0319	0.0432	0.0867	0.1401	0.0306	0.0384	0.0581	0.0929	0.0262	0.0332	0.0824	0.1353	0.0372	0.0477
	MGCN	0.0833	0.1389	0.0366	0.0481	0.0941	0.1525	0.0425	0.0544	0.0665	0.1052	0.0300	0.0377	0.0870	0.1395	0.0356	0.0460
	LGMRec	0.0813	0.1410	0.0352	0.0471	0.0906	0.1496	0.0403	0.0522	0.0624	0.1015	0.0277	0.0355	0.0791	0.1376	0.0335	0.0450
	DAMRS	0.0804	0.1390	0.0355	0.0474	0.0941	0.1526	0.0416	0.0534	0.0670	0.1066	0.0301	0.0380	0.1044	0.1638	0.0452	0.0569
	GUME	0.0835	0.1429	0.0369	0.0489	0.0947	0.1554	0.0424	0.0546	0.0639	0.1016	0.0291	0.0366	0.0968	0.1645	0.0389	0.0524
MMA RSs	CI2MG	0.0720	0.1285	0.0305	0.0420	0.0717	0.1179	0.0331	0.0425	0.0523	0.0845	0.0237	0.0301	0.0772	0.1284	0.0327	0.0429
	MILK	0.0427	0.0763	0.0182	0.0250	0.0362	0.0626	0.0155	0.0209	0.0226	0.0376	0.0094	0.0124	0.0404	0.0640	0.0184	0.0230
	SIBRAR	0.0480	0.0888	0.0207	0.0289	0.0434	0.0758	0.0190	0.0255	0.0264	0.0453	0.0110	0.0148	0.0548	0.0854	0.0220	0.0280
	DGMRec	**0.0897**	**0.1531**	**0.0404**	**0.0528**	**0.1024**	**0.1625**	**0.0462**	**0.0584**	**0.0725**	**0.1134**	**0.0324**	**0.0406**	**0.1093**	**0.1773**	**0.0476**	**0.0611**
	Improv.	7.43%	7.14%	9.49%	7.98%	8.13%	4.57%	8.71%	6.96%	8.21%	6.00%	7.64%	6.84%	4.69%	7.78%	5.31%	7.38%

DGMRec successfully outperforms existing approaches in **missing modality scenarios.**

EXPERIMENTS

- **Model Performance with Real-world Scenario**

  2. Missing Modality + New Items Setting : new items unseen during training appear in the test set.

  *A more **realistic** and **challenging** scenario where item modality features are more crucial.*

Missing Modality + New Items Setting

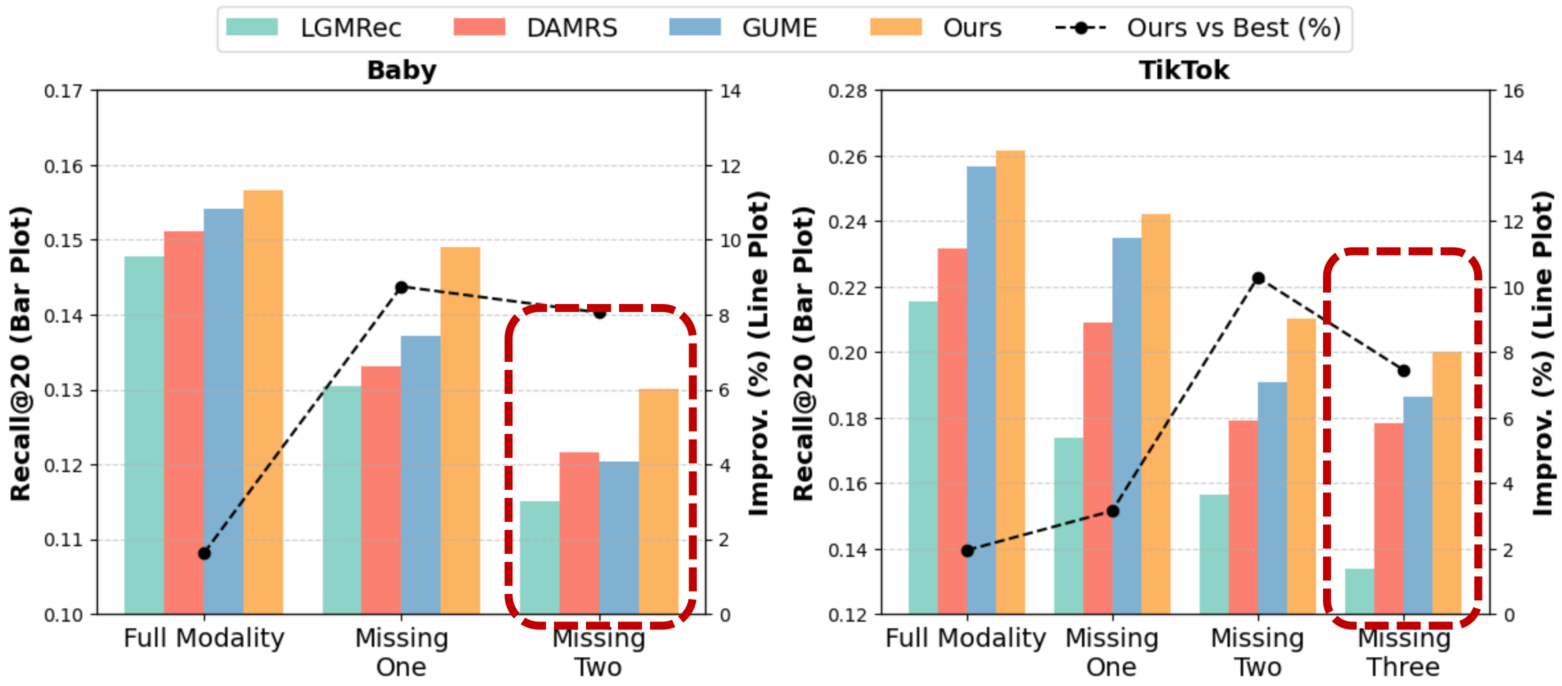
	Dataset	Baby				Sports				Clothing				TikTok			
	Metric	R@20	R@50	N@20	N@50	R@20	R@50	N@20	N@50	R@20	R@50	N@20	N@50	R@20	R@50	N@20	N@50
Multi-Modal Recommenders	VBPR	0.0347	0.0640	0.0177	0.0244	0.0393	0.0641	0.0200	0.0257	0.0265	0.0414	0.0133	0.0166	0.0244	0.0417	0.0118	0.0221
	MMGCN	0.0326	0.0596	0.0157	0.0218	0.0274	0.0489	0.0133	0.0182	0.0170	0.0308	0.0079	0.0110	0.0439	0.0599	0.0186	0.0218
	GRCN	0.0347	0.0621	0.0170	0.0233	0.0368	0.0606	0.0185	0.0239	0.0226	0.0379	0.0109	0.0143	0.0378	0.0661	0.0166	0.0223
	SLMRec	0.0434	0.0702	0.0223	0.0284	0.0477	0.0755	0.0245	0.0308	0.0344	0.0526	0.0176	0.0217	0.0548	0.0775	0.0247	0.0293
	BM3	0.0407	0.0717	0.0204	0.0274	0.0496	0.0796	0.0255	0.0324	0.0317	0.0496	0.0163	0.0202	0.0588	0.0869	0.0262	0.0318
	LATTICE	0.0423	0.0730	0.0213	0.0283	0.0432	0.0713	0.0218	0.0283	0.0344	0.0539	0.0173	0.0216	0.0444	0.0742	0.0209	0.0269
	MGCN	0.0446	0.0802	0.0230	0.0302	0.0478	0.0775	0.0240	0.0316	0.0358	0.0562	0.0182	0.0228	0.0357	0.0694	0.0140	0.0208
	LGMRec	0.0450	0.0772	0.0230	0.0303	0.0462	0.0742	0.0236	0.0300	0.0353	0.0557	0.0175	0.0221	0.0388	0.0632	0.0142	0.0191
	DAMRS	0.0455	0.0779	0.0229	0.0304	0.0480	0.0784	0.0248	0.0317	0.0380	0.0583	0.0192	0.0237	0.0598	0.0872	0.0267	0.0333
	GUME	0.0447	0.0795	0.0225	0.0304	0.0476	0.0776	0.0244	0.0313	0.0357	0.0563	0.0179	0.0224	0.0567	0.0918	0.0217	0.0289
MMA RSs	CI2MG	0.0415	0.0716	0.0210	0.0279	0.0437	0.0718	0.0226	0.0290	0.0294	0.0461	0.0149	0.0186	0.0427	0.0660	0.0188	0.0235
	MILK	0.0247	0.0429	0.0120	0.0162	0.0192	0.0323	0.0093	0.0123	0.0133	0.0226	0.0064	0.0085	0.0212	0.0332	0.0105	0.0129
	SIBRAR	0.0280	0.0495	0.0138	0.0188	0.0257	0.0435	0.0128	0.0169	0.0153	0.0259	0.0070	0.0094	0.0351	0.0527	0.0154	0.0190
	DGMRec	**0.0519**	**0.0876**	**0.0257**	**0.0336**	**0.0532**	**0.0845**	**0.0276**	**0.0348**	**0.0413**	**0.0631**	**0.0211**	**0.0260**	**0.0639**	**0.0973**	**0.0285**	**0.0353**
	Improv.	14.06%	9.22%	11.73%	10.53%	7.26%	6.16%	8.67%	7.41%	8.68%	8.23%	9.89%	9.70%	6.86%	5.99%	6.74%	6.00%

DGMRec demonstrates greater performance gains in **New Items Setting**, where item modality is critical.

EXPERIMENTS

- **How Does Performance Vary Across Different Missing Modality Levels?**

  Performance comparison by group according to the number of missing modalities



- Full Modality

  All modalities are present

- Missing N

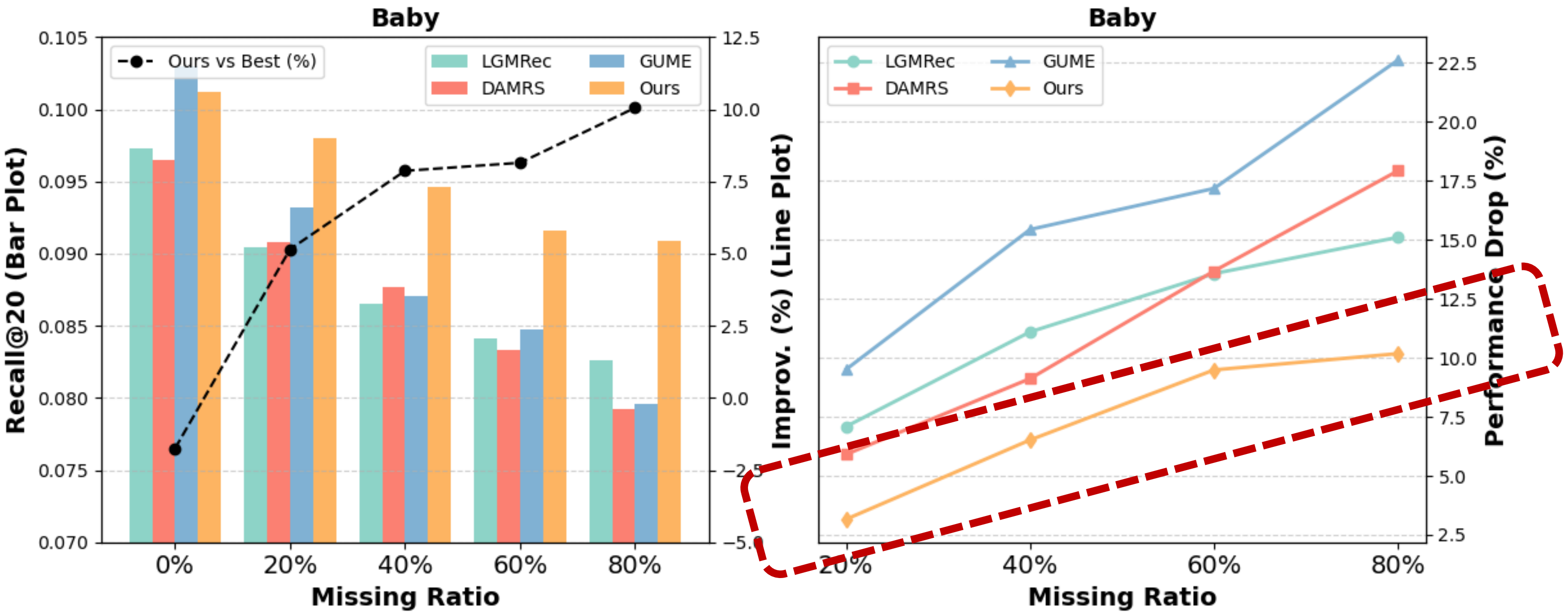
  N modalities are missing.

DGMRec demonstrates **superior performance** when missing modalities are present, even when **all modalities are absent**.

EXPERIMENTS

- **Do Varying Missing Ratios Impact Recommendation Performance?**

  Performance comparison based on missing modality ratio



- Performance Drop

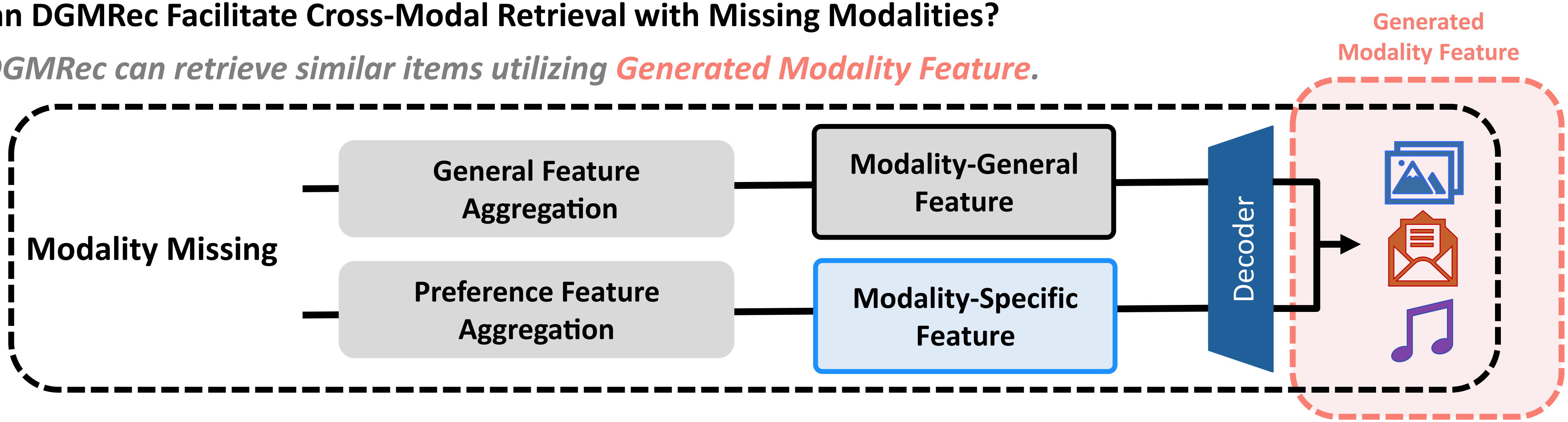
Relative performance drop compared to when the missing ratio is 0%.

DGMRec demonstrates **robustness** even with increasing missing modalities,
achieving **higher performance** than other models in most cases.

EXPERIMENTS

- **Can DGMRec Facilitate Cross-Modal Retrieval with Missing Modalities?**

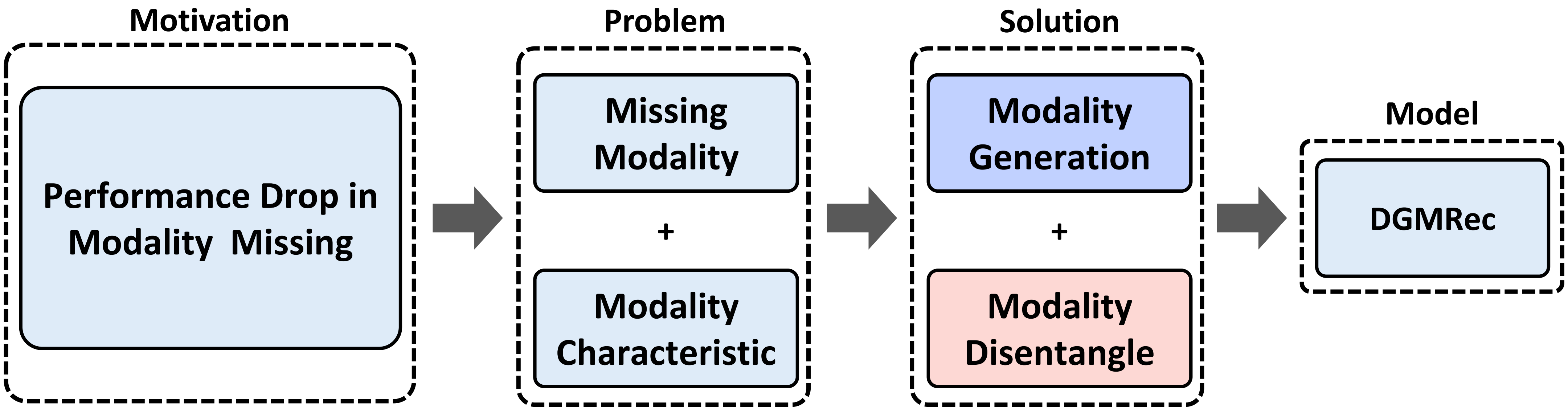
  DGMRec can retrieve similar items utilizing Generated Modality Feature.



Datasets	Baby			
	Missing 1 Modality		Missing 2 Modalities	
	NN	DGMRec	NN	DGMRec
Hit@10	0.1344	0.3577	-	0.3496
Hit@20	0.1999	0.3801	-	0.3690

DGMRec can successfully retrieve the original modality features when modalities are missing, even works when all modalities are absent.

CONCLUSIONS



Code

Disentangling and Generating Modalities for Recommendation in Missing Modality Scenarios

Code: https://github.com/ptkjw1997/DGMRec

Paper: https://arxiv.org/abs/2504.16352

Email: kim.jiwan@kaist.ac.kr

Paper