



Đại học Quốc gia Thành phố Hồ Chí Minh
Trường Đại học Công nghệ Thông tin
Khoa Khoa học và Kỹ thuật Thông tin

DS300 - HỆ KHUYẾN NGHỊ

ỨNG DỤNG ĐẶC TRƯNG ĐA PHƯƠNG
THỨC TRONG HỆ THỐNG KHUYẾN NGHỊ
SÁCH TIẾNG VIỆT

Sinh viên thực hiện:

Lê Xuân Bình 22520131

Thái Minh Lâm 22520745

Mã Kim Phát 22521071

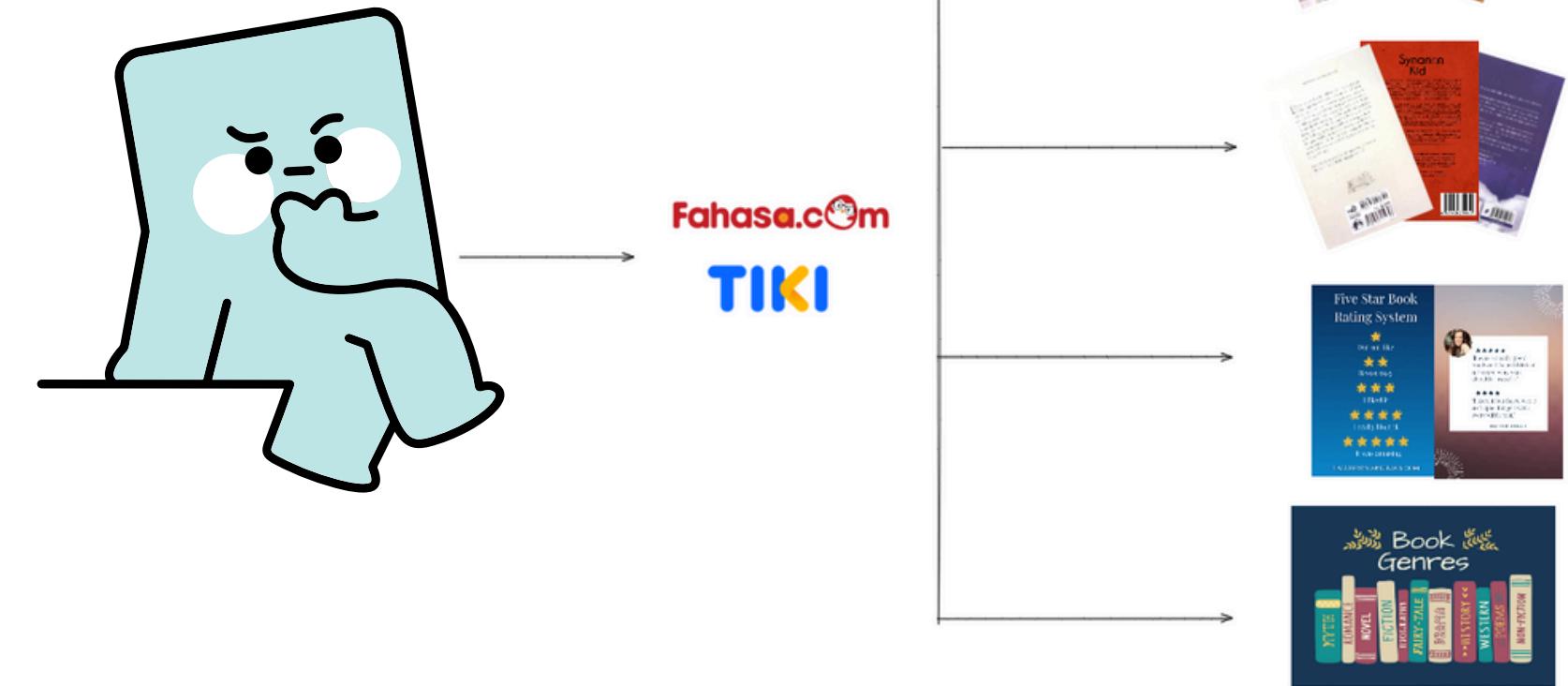
Giảng viên hướng dẫn:
ThS. Huỳnh Văn Tín

TP HCM, 29/12/2025



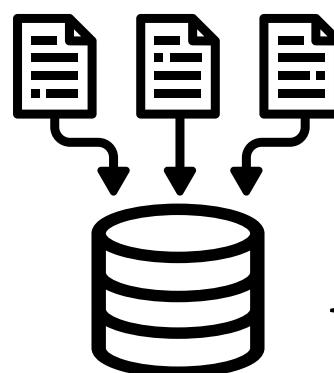
Động lực bài toán

- Sự bùng nổ của **thương mại điện tử**.
- Chưa có **bộ dữ liệu** và **nghiên cứu thực nghiệm** về *tác động đa phương thức* đến khuyến nghị sách trong tiếng Việt.
- Triển khai hệ khuyến nghị sách trên **hệ thống dữ liệu lớn**.



Khách hàng quan tâm gì khi chọn mua sách?

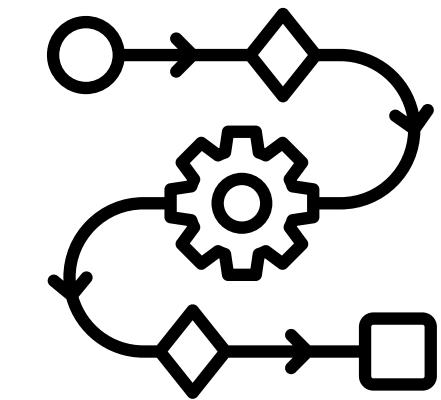
Đóng góp của nghiên cứu



Xây dựng
bộ dữ liệu

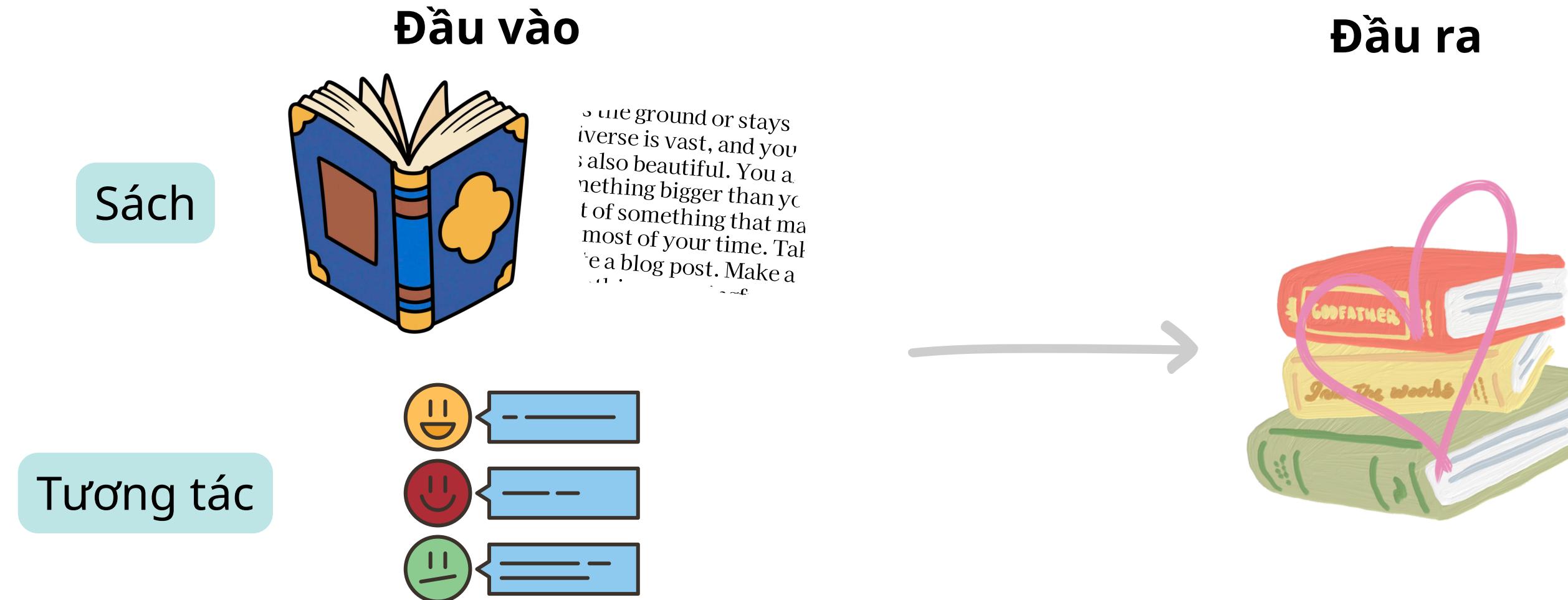


Đánh giá
thực nghiệm



Triển khai
hệ thống

Định nghĩa bài toán



Các hướng tiếp cận

- Collaborative Filtering (CF)
- Content-Based Filtering (CBF)
- Multimodal Recommend System (MRS)

Xây dựng bộ dữ liệu



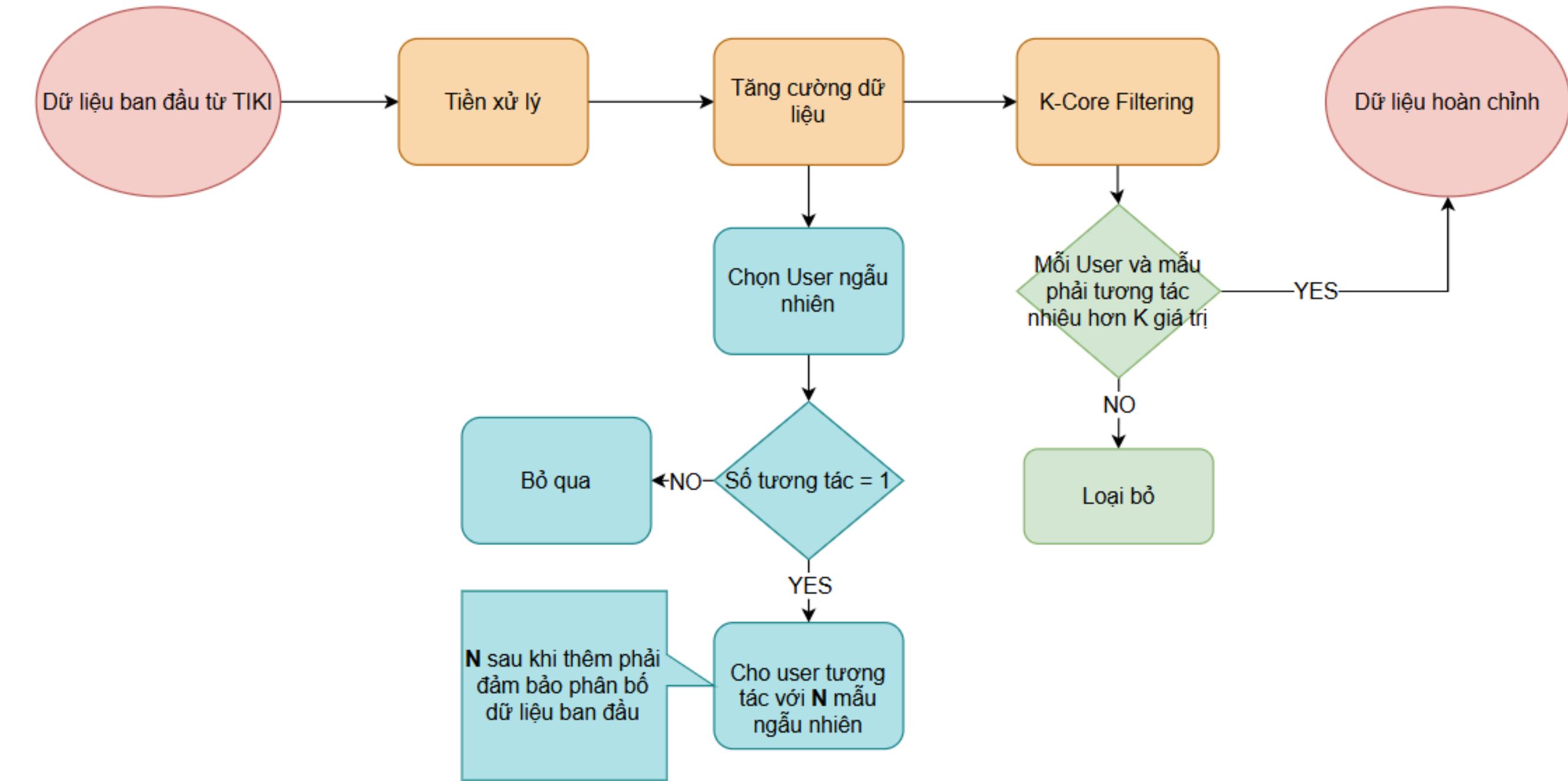
Nguồn dữ liệu



BeautifulSoup

Tool crawl data

Thu thập dữ liệu



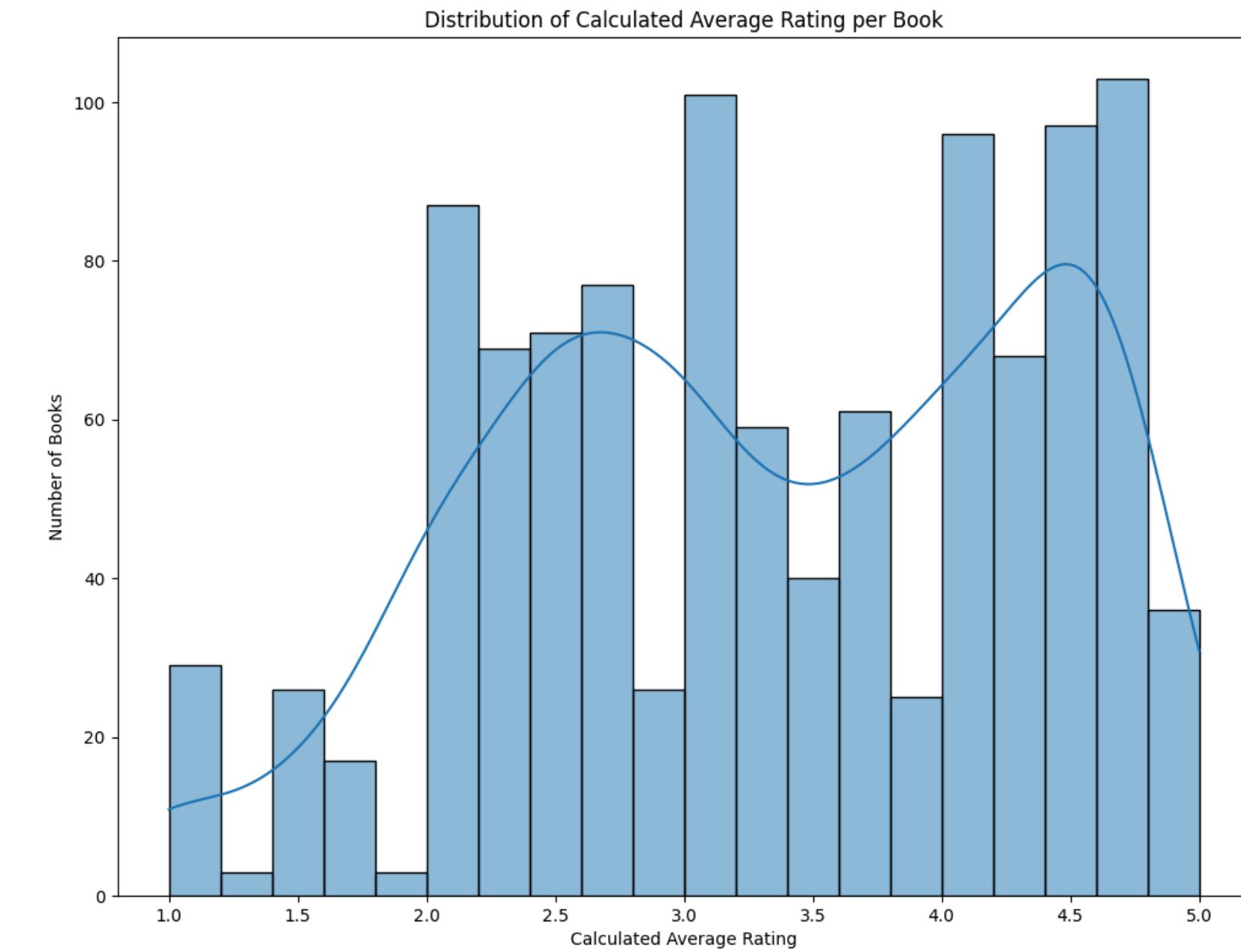
Tiền xử lý & Tăng cường dữ liệu

Phân tích bộ dữ liệu

Định lượng

Đặc trưng	Sách	Đánh giá
Số mẫu	1.115	26.050
Số thuộc tính	14	6

Định tính



Framework thực nghiệm cho MRS

Evaluation

Experiments

Trainer

Multimodal
Models

Optimizer

Data
Encapsulation

- k -core filtering
- ID alignment
- Data splitting
- Multimodal information vectorization

Raw Data

Multimodal
Information

User-Item
Interactions

Backend

Pytorch

Config

Text Encoder: ViSoBERT

Image Encoder: ViT, ResNet-152



REC

<https://github.com/enoche/MMRec>

Các mô hình thử nghiệm

MRS	CB	CF
BPR	Content-based	User-based
VBPR	Typebook-based	
LATTICE		
FREEDOM		
MMGCN		
SLMRec		
	Độ tương đồng: Cosine Pearson	

Bảng liệt kê các mô hình thực nghiệm

Thực nghiệm

Các độ đo đánh giá

NDCG@k

$$NDCG@k = \frac{DCG@k}{IDCG@k}$$

Precision@k

$$Precision@k = \frac{|R_u \cap T_u|}{k}$$

Recall@k

$$Recall@k = \frac{|R_u \cap T_u|}{|T_u|}$$

mAP

$$AP@k = \frac{1}{\min(|T_u|, k)} \sum_{i=1}^k (Precision@i \times rel(i))$$

Thực nghiệm

Kết quả thực nghiệm

	Pre@5	Rec@5	mAP@5	NDCG@5	Pre@10	Rec@10	mAP@10	NDCG@10
User-based (Cosine)	0.0011	0.0057	0.0027	0.0034	0.0016	0.0156	0.0039	0.0065
User-based (Pearson)	0.0006	0.0028	0.0008	0.0013	0.0004	0.0042	0.0011	0.0018
Content-based (Cosine)	0.0011	0.0057	0.0026	0.0034	0.0017	0.0170	0.0041	0.0070
Content-based (Pearson)	0.0011	0.0057	0.0026	0.0034	0.0017	0.0170	0.0041	0.0070
Typebook-based (Cosine)	0.0028	0.0141	0.0052	0.0074	0.0018	0.0183	0.0058	0.0088
Typebook-based (Pearson)	0.0019	0.0099	0.0022	0.0039	0.0014	0.0141	0.0027	0.0053

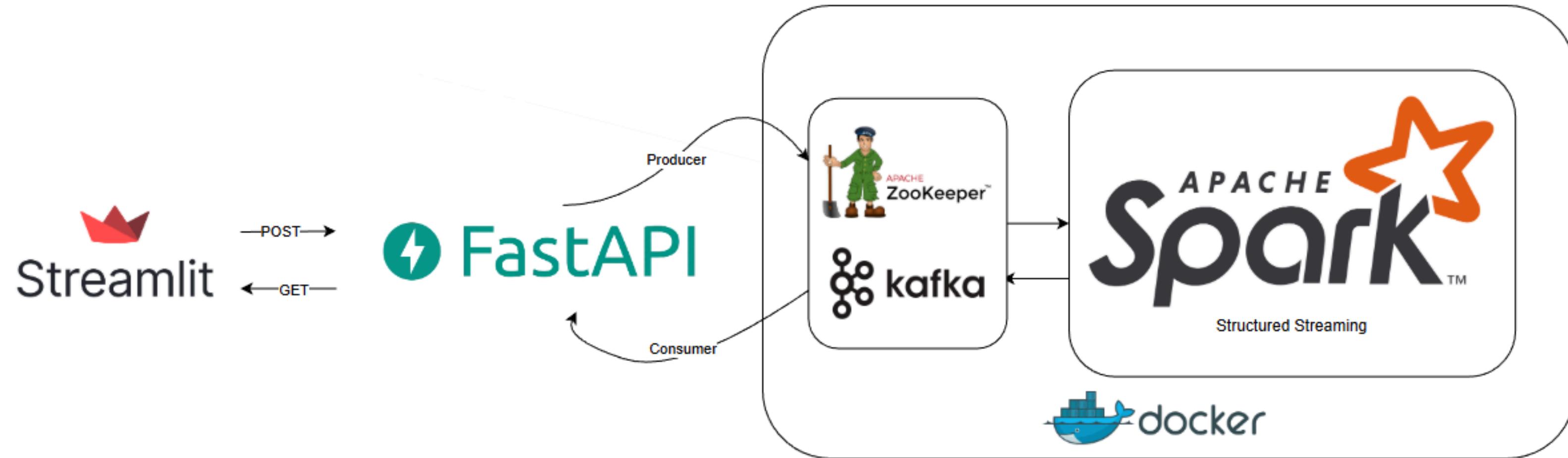
Bảng kết quả thực nghiệm của các mô hình Content-based và Collaborative Filtering

Thực nghiệm

Bảng kết quả thực nghiệm của các mô hình MRS

Model	Modality	Encoder	P@5	R@5	mAP@5	N@5	P@10	R@10	mAP@10	N@10
BPR	-	-	0.002	0.0099	0.0043	0.0057	0.0014	0.0141	0.0049	0.0071
VBPR	Text (Desc)	ViSoBERT	0.0023	0.0113	0.0035	0.0054	0.0018	0.0184	0.0046	0.0078
	Text (Type)	ViSoBERT	0.0031	0.0156	0.0058	0.0081	0.0027	0.0269	0.0072	0.0116
	Image	ViT (CLS)	0.0025	0.0127	0.0099	0.0106	0.0023	0.0226	0.0114	0.014
	Image	ResNet-152	0.0028	0.0141	0.007	0.0088	0.0021	0.0212	0.0078	0.011
	Text + Image	ViSoBERT + Res152	0.0034	0.017	0.0081	0.0102	0.0023	0.0226	0.0088	0.012
LATTICE	Text (Desc)	ViSoBERT	0.0025	0.0127	0.0031	0.0054	0.0024	0.024	0.0045	0.009
	Text (Type)	VisoBERT	0.0031	0.0156	0.0086	0.0103	0.0023	0.0226	0.0096	0.0124
	Image	ViT (CLS)	0.0028	0.0141	0.0054	0.0075	0.0028	0.0283	0.007	0.0118
	Image	ResNet-152	0.0014	0.0071	0.0031	0.0041	0.0013	0.0127	0.0038	0.0058
	Text + Image	ViSoBERT + ViT (CLS)	0.0037	0.0184	0.0067	0.0095	0.0028	0.0283	0.0078	0.0125
FREEDOM	Text (Desc)	ViSoBERT	0.0025	0.0127	0.0074	0.0086	0.0021	0.0212	0.0085	0.0114
	Text (Type)	VisoBERT	0.0054	0.0269	0.0113	0.015	0.0035	0.0354	0.0124	0.0178
	Image	ViT (CLS)	0.0028	0.0141	0.007	0.0087	0.002	0.0198	0.0076	0.0105
	Image	ResNet-152	0.0006	0.0028	0.0017	0.002	0.001	0.0099	0.0025	0.0041
	Text + Image	ViSoBERT + ViT (CLS)	0.0048	0.024	0.0111	0.0142	0.0031	0.0311	0.012	0.0165
MMGCN	Text (Desc)	ViSoBERT	0.0076	0.0382	0.023	0.0267	0.0048	0.0481	0.0243	0.0299
	Text (Type)	VisoBERT	0.0085	0.0424	0.0236	0.0282	0.0052	0.0523	0.025	0.0315
	Image	ViT (CLS)	0.0082	0.041	0.0133	0.0201	0.0057	0.0566	0.0156	0.0254
	Image	ResNet-152	0.0076	0.0382	0.0121	0.0185	0.004	0.0396	0.0124	0.019
	Text + Image	ViSoBERT + ViT (CLS)	0.0079	0.0396	0.0192	0.0243	0.0045	0.0453	0.0198	0.026
SLMRec	Text (Desc) + Image	ViSoBERT + ViT (CLS)	0.0059	0.0297	0.0128	0.0169	0.005	0.0495	0.0155	0.0234
	Text (Type) + Image	ViSoBERT + ViT (CLS)	0.0062	0.0311	0.0177	0.021	0.005	0.0495	0.0201	0.0269
	Text (Desc) + Image	ViSoBERT + ResNet-152 (CLS)	0.0071	0.0354	0.0164	0.021	0.0055	0.0552	0.0191	0.0275
	Text (Type) + Image	ViSoBERT + ResNet-152 (CLS)	0.0059	0.0297	0.0194	0.022	0.0038	0.0382	0.0205	0.0247

Triển khai hệ thống khuyến nghị



Kiến trúc pipeline xử lý dữ liệu và tích hợp mô hình khuyến nghị thời gian thực

DEMO

Kết luận

Ưu điểm

Khuyến nghị đa phương thức **giải quyết tốt vấn đề dữ liệu bị thưa**.

Khuyến nghị đa phương thức tận dụng tối đa các **thông tin ngũ cảnh** của người dùng và sản phẩm, cho phép **khai thác thông tin sâu** hơn các phương pháp truyền thống.

Hướng phát triển

Fine-tune encoder.

Làm giàu thêm nguồn dữ liệu.

Khuyết điểm

Hiệu suất của khuyến nghị đa phương thức **phụ thuộc nhiều vào các Encoder**.

Các kiến trúc mô hình phức tạp khiến **tiêu tốn thời gian và tài nguyên huấn luyện**.

Các kiến trúc mô hình có **độ phức tạp cao** tuy có độ chính xác cao nhưng **tốc độ xử lý thực tế** có thể bị ảnh hưởng.