



The RecSys Pipeline:

A case-study approach

Dr. Bereket A. Yilma



The typical RecSys Pipeline

Data
Pre-processing



Model
Training

Post
Processing

Evaluation



- Sort
- Filter
- Recommend



The RecSys pipeline: A case-study approach



Task: Design a **Personalised Visual Art Recommendation** engine for the National Gallery, London



The RecSys pipeline: A case-study approach



DISCLAIMER



The RecSys pipeline: A case-study approach

Personalised Visual Art Recommendation



Context: National Gallery, London

- $\geq 2,300$ paintings dating from the mid-13th century to 1900.
- Total floor area of 46,369 square meters, 3 floors.
- 6.2 million visitors/year (2019)

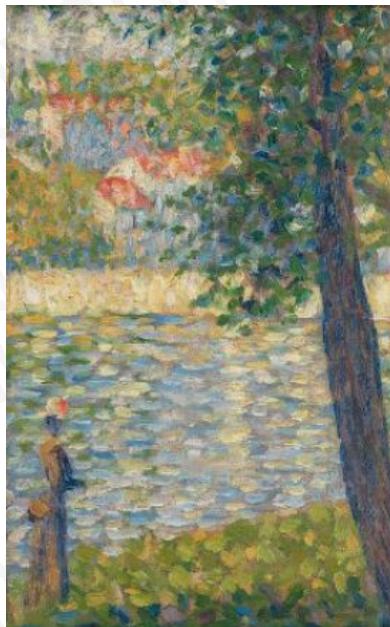
The RecSys pipeline: A case-study approach



The RecSys pipeline: A case-study approach

1. Data: Visual Art (Paintings)

- 2,368 painting



painting_id	000-018P-0000
title	The Morning Walk
artist	Georges Seurat
publication_date	19th_century
size_format	Portrait
size	Very Small
technique	oil painting
description	A woman, silhouetted against the shimmering water, strolls along a riverbank. The red roofs of houses can be made out along the opposite bank. Between 1882 and 1886 Seurat painted numerous such landscape studies on small wooden panels, some as independent works and others in preparation for his large-scale compositions. This sketch provided the starting point for a painting of 1885, 'The Seine at Courbevoie' (private collection).



The RecSys pipeline: A case-study approach



Data
Pre-processing



New Visitor



Query User (Profiling)

1. Rate few paintings
2. Popular paintings
3. Visiting style
4. Available time ...



The RecSys pipeline: A case-study approach

Data
Pre-processing



Task
Personalised
Recommendation

Model
Training

$R^{m \times m}$

1	0.77	0.57	0.54	0.37	0.46	0.45	0.44	0.46	0.37	0.63	0.66	0.59	0.54	0.59	0.52
0.77	1	0.09	0.68	0.54	0.56	0.61	0.53	0.5	0.46	0.74	0.85	0.66	0.58	0.67	0.6
0.57	0.69	1	0.92	0.34	0.39	0.45	0.4	0.43	0.45	0.48	0.59	0.57	0.62	0.67	0.59
0.54	0.68	0.92	1	0.37	0.38	0.5	0.44	0.42	0.41	0.47	0.59	0.55	0.6	0.66	0.59
0.37	0.54	0.34	0.37	1	0.62	0.55	0.51	0.48	0.52	0.55	0.62	0.42	0.38	0.39	0.39
0.46	0.56	0.39	0.38	0.62	1	0.58	0.54	0.49	0.6	0.53	0.7	0.47	0.38	0.39	0.34
0.45	0.61	0.45	0.5	0.55	0.58	1	0.7	0.39	0.45	0.57	0.65	0.58	0.47	0.54	0.45
0.44	0.53	0.4	0.44	0.51	0.54	0.7	1	0.28	0.39	0.44	0.57	0.6	0.48	0.47	0.39
0.46	0.5	0.43	0.42	0.48	0.49	0.39	0.28	1	0.65	0.49	0.52	0.47	0.37	0.4	0.37
0.37	0.46	0.45	0.41	0.52	0.6	0.45	0.39	0.65	1	0.5	0.57	0.47	0.41	0.44	0.42
0.63	0.74	0.48	0.47	0.55	0.53	0.57	0.44	0.49	0.5	1	0.82	0.53	0.43	0.58	0.51
0.66	0.85	0.59	0.59	0.62	0.7	0.65	0.57	0.52	0.57	0.82	1	0.61	0.53	0.66	0.61
0.59	0.66	0.57	0.55	0.42	0.47	0.58	0.6	0.47	0.47	0.53	0.61	1	0.7	0.53	0.45
0.54	0.58	0.62	0.6	0.38	0.38	0.47	0.48	0.37	0.41	0.43	0.53	0.7	1	0.53	0.46
0.59	0.67	0.67	0.66	0.39	0.39	0.54	0.47	0.4	0.44	0.58	0.66	0.53	0.53	1	0.9
0.52	0.6	0.59	0.59	0.39	0.34	0.45	0.39	0.37	0.42	0.51	0.61	0.45	0.46	0.9	1

- If a user likes painting A find paintings B, C, D that are similar to A.

The RecSys pipeline: A case-study approach

Data
Pre-processing



Task
Personalised
Recommendation



Image



Metadata

painting_id	000-018P-0000
title	The Morning Walk
artist	Georges Seurat
publication_date	19th century
size_format	Portrait
size	Very Small
technique	oil painting

description	A woman, silhouetted against the shimmering water, strolls along a riverbank. The red roofs of houses can be made out along the opposite bank. Between 1882 and 1886 Seurat painted numerous such landscape studies on small wooden panels, some as independent works and others in preparation for his large-scale compositions. This sketch provided the starting point for a painting of 1885, 'The Seine at Courbevoie' (private collection).
-------------	---

Textual description

The RecSys pipeline: A case-study approach

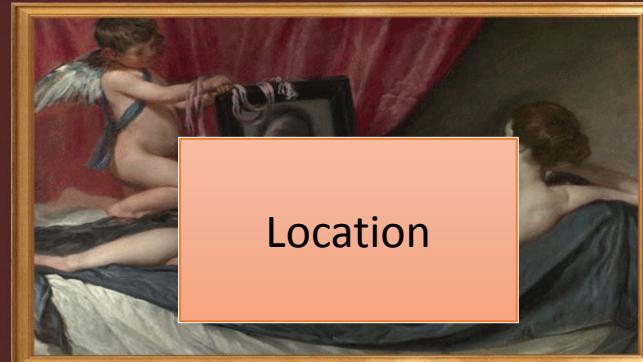
Manually curated metadata to drive recommendations.



Authorship

Art History

Style



Location

Size

Material

History



The RecSys pipeline: A case-study approach

Train a model that can learn Visual features from images of paintings

Some issues:

- Recommendations don't have direct interpretation.
- Often fail to capture complex semantics such as triggered reflections



The RecSys pipeline: A case-study approach

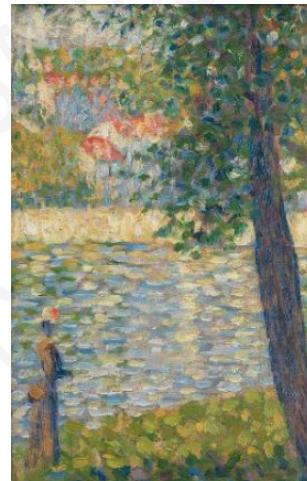
Data
Pre-processing



Task
Personalised
Recommendation

Model
Training

Textual description

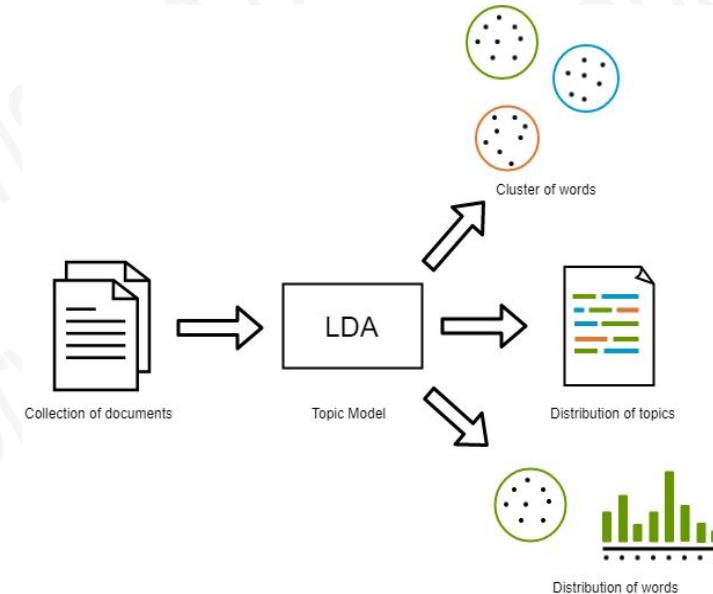


painting_id	000-018P-0000
title	The Morning Walk
artist	Georges Seurat
publication_date	19th_century
size_format	Portrait
size	Very Small
technique	oil painting
description	A woman, silhouetted against the shimmering water, strolls along a riverbank. The red roofs of houses can be made out along the opposite bank. Between 1882 and 1886 Seurat painted numerous such landscape studies on small wooden panels, some as independent works and others in preparation for his large-scale compositions. This sketch provided the starting point for a painting of 1885, 'The Seine at Courbevoie' (private collection).

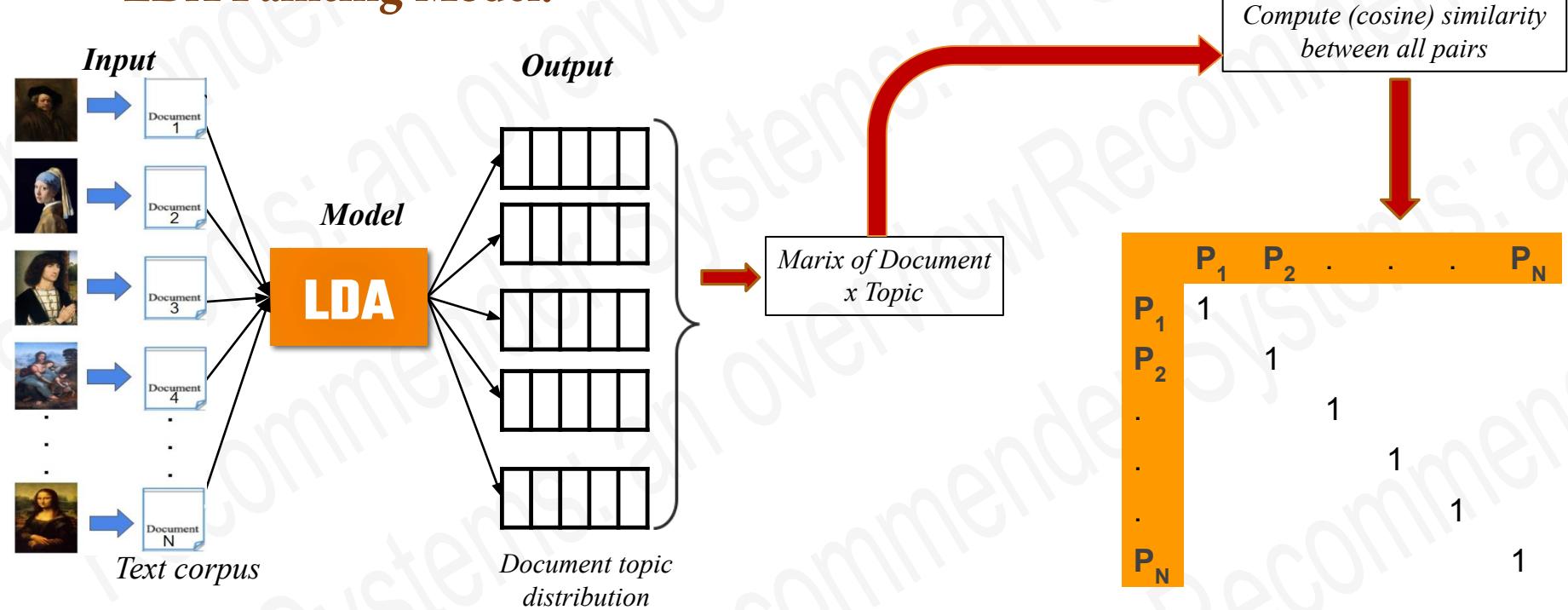
The RecSys pipeline: A case-study approach

Topic Modelling

- LDA is an unsupervised generative probabilistic model.
- Particularly it can identify similar documents by uncovering abstract topics that occur in a collection of documents.

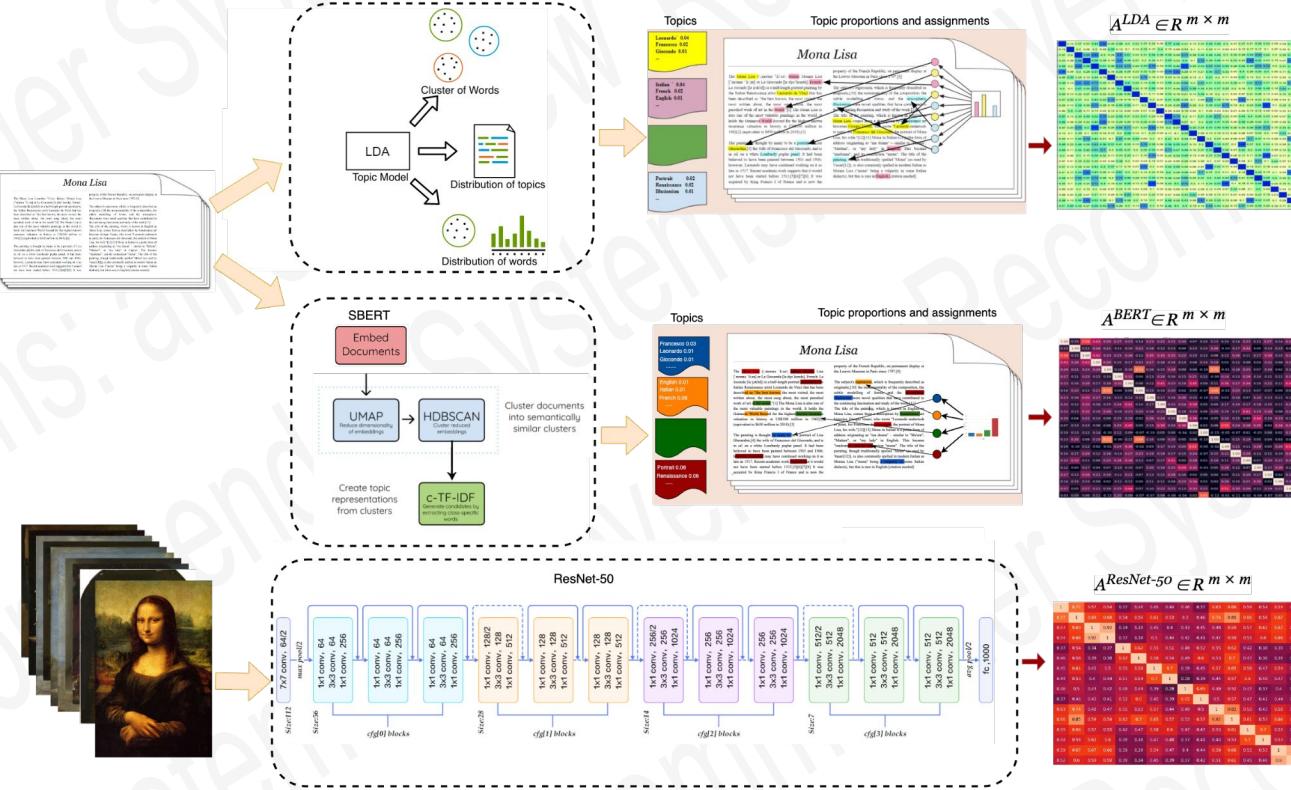


LDA Painting Model:





The RecSys pipeline: A case-study approach



Data Pre-processing



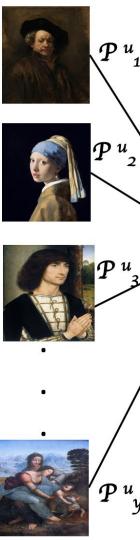
Model Training

Post Processing

Input
(User rated paintings)

Weight (W^u) (P^u)

w^{u_1}
w^{u_2}
w^{u_3}
.
w^{u_y}



$S(P, U)$

Output
Score Dataset

$S(P, U)$
$S(P_1, U)$
$S(P_2, U)$
$S(P_3, U)$
.
.
$S(P_M, U)$

- The predicted Score $S(P_i, U)$ for a novel painting P_i with respect to an active user u is based on the weighted average score from all other paintings that have been rated by the active user.

$$S(P, U) = \frac{1}{N} \sum_{j=1}^N w_j * d(p_i, p_j)$$

- $d(p_i, p_j)$ is the similarity between p_i and p_j according to our model (LDA, BERT, ResNet..)
- Top K recommendation



Active user preferences



Normalized weight
[0,1]



- Proxy for how relevant/ good are recommender systems.

1. Offline Experiment.

- Easiest to conduct
- Requires no interaction with users.

2. User Studies

- Small group of Subjects
- Controlled setting
- Qualitative/quantitative

3. Online Experiments

- The most trustworthy
- A pool of real users (unaware of the experiment)



The RecSys pipeline: A case-study approach

Evaluation

A couple of basic guidelines in general experimental studies

1. Hypothesis:

What do you want to evaluate? Form a **concise** and **restrictive** Hypothesis.

E.g. Algorithm **A** better predicts rating than Algorithm **B**

→ **Predictive Accuracy** not other factors

Algorithm **A** generates diverse recommendations than Algorithm **B**

→ **Diversity** not other factors, etc.

The RecSys pipeline: A case-study approach

A couple of basic guidelines in general experimental studies

2. Controlling variables: Variables that are not being tested should stay fixed.

E.g. Algorithm **A** better predicts rating than Algorithm **B** → **Predictive Accuracy**

Train A on NG dataset



Train B on Louvre dataset



Can not tell why $A > B$ or $B < A$

- Superior model?
- better input data?

Train both A & B on the same dataset.



The RecSys pipeline: A case-study approach



Evaluation: Offline Experiments

Performed using a pre-collected dataset of users choosing or rating of items.

- Using this data we can simulate behaviour of users that interact with the a Recommender System: (assuming users will behave the same when the Recommender System is deployed.)
 - To make reliable decision.
- **Attractive & easy:**
 - Requires no interaction with real users.
 - We can compare wide range of candidate algorithms at low cost.



The RecSys pipeline: A case-study approach



Evaluation: Offline Experiments

- **Downside:** We can not measure the recommender's influence on user behavior.

Useful for :

- Filtering out inappropriate approaches, select candidate algorithms for more costly user studies/online experiments.
- Tuning parameters of algorithms.



The RecSys pipeline: A case-study approach



Evaluation: User Studies

Conducted by recruiting a set of test subject, and asking them to perform several tasks requiring an interaction with the recommendation system.

- During interaction we observe and record their behaviour.

Quantitative measures: % of completed task, accuracy, time, etc.

Qualitative measures: pre/post interaction questions that are not directly observable.

- weather the subject enjoyed the UI
- weather the subject perceived the task easy, etc.



The RecSys pipeline: A case-study approach

Evaluation: User Studies

Downsides:

- Expensive to conduct (large set of subjects, large set of tasks)
- Subjects (volunteers or employed)
 - Motivations (intrinsic/ extrinsic)--> quality of response
 - Budget
- Testing all possible scenarios can be challenging.
- Finding subjects that represent the entire population. (Bias)

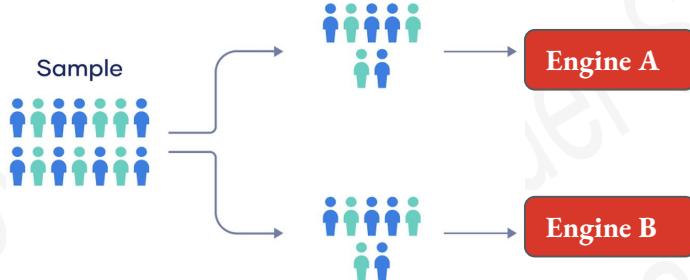
Pilot user studies: to test the systems for bugs and malfunctions.

Evaluation: User Studies

Between vs. Within Subjects

Few candidate approaches: each method must be tested on the same task.

1. Between Subjects (A-B testing)



- Easier to setup and analyse correctly
- No learning across conditions
- Test long term effects

2. Within Subjects



- More informative (superiority of methods)
- Can ask comparative questions about candidates



The RecSys pipeline: A case-study approach



Evaluation: Online Evaluation

Done with a set of real users unaware of the experiment.

- Many realistic RecSys applications wish to influence the behaviour of users.
(engagement, purchase)
- Measure the change in user behavior while interacting with recommendations.
 - Did the user follow recommendations? for how long?
 - Weather some utility gathered from users of system A exceeds the one from system B.
(which system is superior?)



The RecSys pipeline: A case-study approach

Evaluation: Online Evaluation

**Most reliable:
Real users with real needs in the context.**

- The real effect of Recommendation systems depends on several factors
 - Users' intent
 - How specific are their needs?
 - How much novelty are they looking for?
 - How much risk are they willing to take?
 - Users' context
 - What items are they already familiar with ?
 - how much they trust the system?
 - The Recommendation interface



The RecSys pipeline: A case-study approach



Evaluation: Online Evaluation

Downsides:

- May discourage users from using the real system ever again.
- Financial risk in commercial applications.

To reduce such risks it is best to run an online evaluation last

- After an extensive Offline study followed by a user study provided evidence that the candidate approaches are reasonable.



The RecSys pipeline: A case-study approach



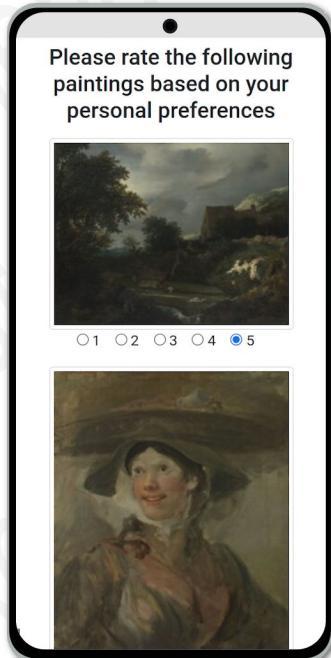
Evaluation: Drawing Reliable Conclusions

Even after a carefully conducted experiment Recsys might fail in unseen scenarios when deployed in real life.

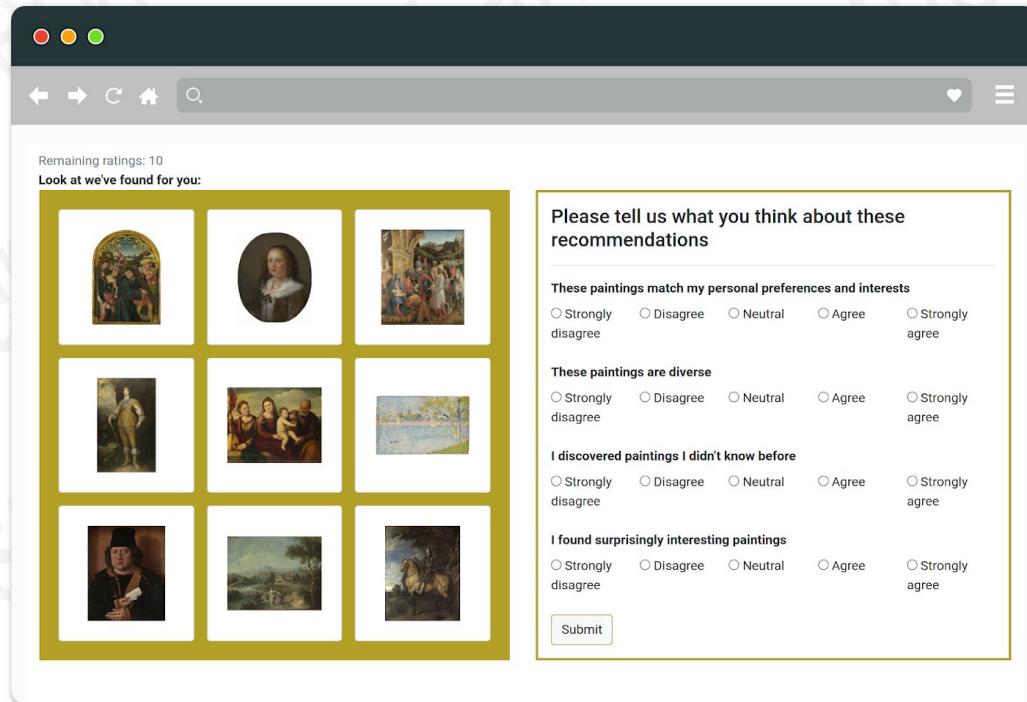
To reduce such risks it is best to perform significance testing on the results.

The RecSys pipeline: A case-study approach

Evaluation: User Study



User-centric evaluation

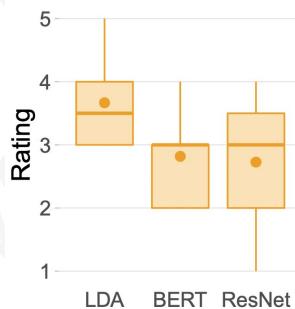


A user study interface. At the top, it says "Remaining ratings: 10" and "Look at we've found for you:". Below this is a grid of nine painting thumbnails. To the right is a large box for user feedback. The box contains the following sections:

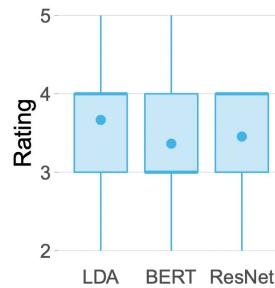
- Please tell us what you think about these recommendations**
- These paintings match my personal preferences and interests**
 Strongly disagree Disagree Neutral Agree Strongly agree
- These paintings are diverse**
 Strongly disagree Disagree Neutral Agree Strongly agree
- I discovered paintings I didn't know before**
 Strongly disagree Disagree Neutral Agree Strongly agree
- I found surprisingly interesting paintings**
 Strongly disagree Disagree Neutral Agree Strongly agree

At the bottom left of the feedback box is a "Submit" button.

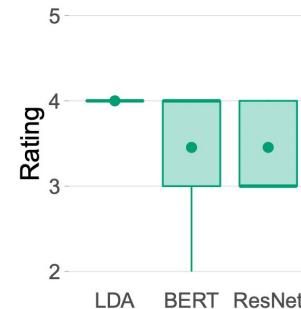
Evaluation: User Study



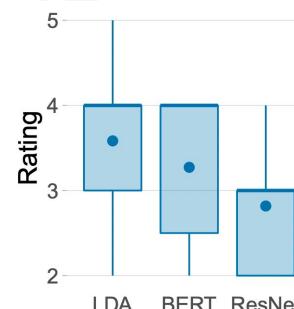
(a) Accuracy



(b) Diversity



(c) Novelty



(d) Serendipity

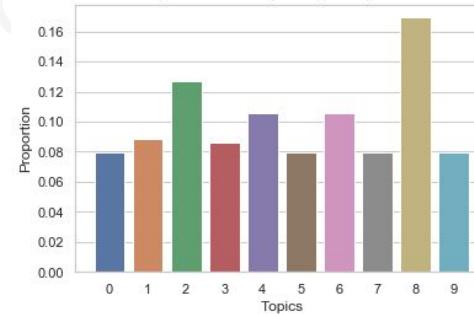
The RecSys pipeline: A case-study approach

Explaining Recommendations

Target painting



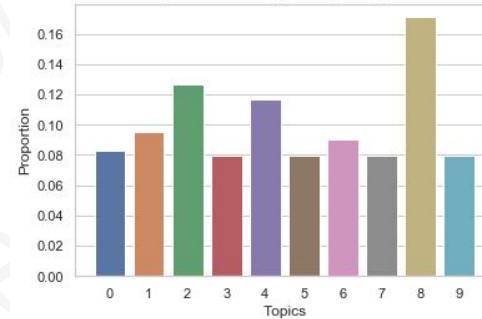
Proportions of the topics for painting n°2330



Most similar painting



Proportions of the topics for painting n°843



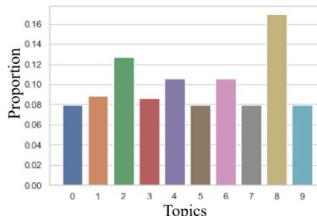
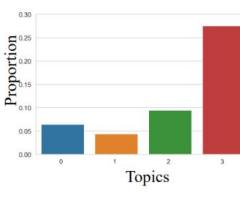
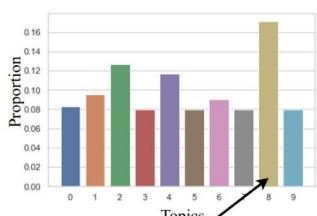
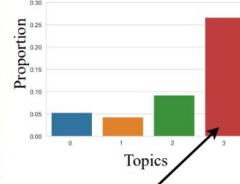
Topic 8
CHRIST
SAINT
JESUS
EVANGELIST
CROSS
CHURCH



Explainable recommendations have a positive impact on user experience.

The RecSys pipeline: A case-study approach

Explaining Recommendations

	LDA	BERT	ResNet																																
Target painting	  <table border="1"> <caption>LDA Topic Proportions for Target Painting</caption> <thead> <tr> <th>Topic</th> <th>Proportion</th> </tr> </thead> <tbody> <tr><td>0</td><td>0.08</td></tr> <tr><td>1</td><td>0.09</td></tr> <tr><td>2</td><td>0.13</td></tr> <tr><td>3</td><td>0.08</td></tr> <tr><td>4</td><td>0.10</td></tr> <tr><td>5</td><td>0.07</td></tr> <tr><td>6</td><td>0.10</td></tr> <tr><td>7</td><td>0.07</td></tr> <tr><td>8</td><td>0.16</td></tr> <tr><td>9</td><td>0.08</td></tr> </tbody> </table>	Topic	Proportion	0	0.08	1	0.09	2	0.13	3	0.08	4	0.10	5	0.07	6	0.10	7	0.07	8	0.16	9	0.08	  <table border="1"> <caption>BERT Topic Proportions for Target Painting</caption> <thead> <tr> <th>Topic</th> <th>Proportion</th> </tr> </thead> <tbody> <tr><td>0</td><td>0.04</td></tr> <tr><td>1</td><td>0.03</td></tr> <tr><td>2</td><td>0.07</td></tr> <tr><td>3</td><td>0.26</td></tr> </tbody> </table>	Topic	Proportion	0	0.04	1	0.03	2	0.07	3	0.26	
Topic	Proportion																																		
0	0.08																																		
1	0.09																																		
2	0.13																																		
3	0.08																																		
4	0.10																																		
5	0.07																																		
6	0.10																																		
7	0.07																																		
8	0.16																																		
9	0.08																																		
Topic	Proportion																																		
0	0.04																																		
1	0.03																																		
2	0.07																																		
3	0.26																																		
Most similar painting	  <table border="1"> <caption>LDA Topic Proportions for Most Similar Painting</caption> <thead> <tr> <th>Topic</th> <th>Proportion</th> </tr> </thead> <tbody> <tr><td>0</td><td>0.07</td></tr> <tr><td>1</td><td>0.09</td></tr> <tr><td>2</td><td>0.13</td></tr> <tr><td>3</td><td>0.07</td></tr> <tr><td>4</td><td>0.11</td></tr> <tr><td>5</td><td>0.07</td></tr> <tr><td>6</td><td>0.08</td></tr> <tr><td>7</td><td>0.07</td></tr> <tr><td>8</td><td>0.16</td></tr> <tr><td>9</td><td>0.07</td></tr> </tbody> </table>	Topic	Proportion	0	0.07	1	0.09	2	0.13	3	0.07	4	0.11	5	0.07	6	0.08	7	0.07	8	0.16	9	0.07	  <table border="1"> <caption>BERT Topic Proportions for Most Similar Painting</caption> <thead> <tr> <th>Topic</th> <th>Proportion</th> </tr> </thead> <tbody> <tr><td>0</td><td>0.04</td></tr> <tr><td>1</td><td>0.03</td></tr> <tr><td>2</td><td>0.07</td></tr> <tr><td>3</td><td>0.26</td></tr> </tbody> </table>	Topic	Proportion	0	0.04	1	0.03	2	0.07	3	0.26	
Topic	Proportion																																		
0	0.07																																		
1	0.09																																		
2	0.13																																		
3	0.07																																		
4	0.11																																		
5	0.07																																		
6	0.08																																		
7	0.07																																		
8	0.16																																		
9	0.07																																		
Topic	Proportion																																		
0	0.04																																		
1	0.03																																		
2	0.07																																		
3	0.26																																		

(christ, saint, altarpiece, panel,
Jesus, new testament, evangelist,
cross, church crucification)

(landscape, oil, van, anchor,
17th_century, river, view, scene,
17th_century landscape)

The RecSys pipeline: A case-study approach



Time orders Old Age to destroy Beauty
by Pompeo Girolamo Batoni
18th century

Explaining Recommendations

Beauty

Old Age

Time

Batoni intends to encourage considering the brevity of youth and the inevitable passing of time.

Semantic Gap

Mary Magdalene

precious ointment



The Donor and Saint Mary Magdalene
by Marten van Heemskerck.
16th century

The typical RecSys Pipeline

Data
Pre-processing



Model
Training

Post
Processing

Evaluation



- Sort
- Filter
- Recommend



Next session:

- **A Multi-Stakeholder aware RecSys**
- **How to formulate the RecSys Problem? →(A framework)**