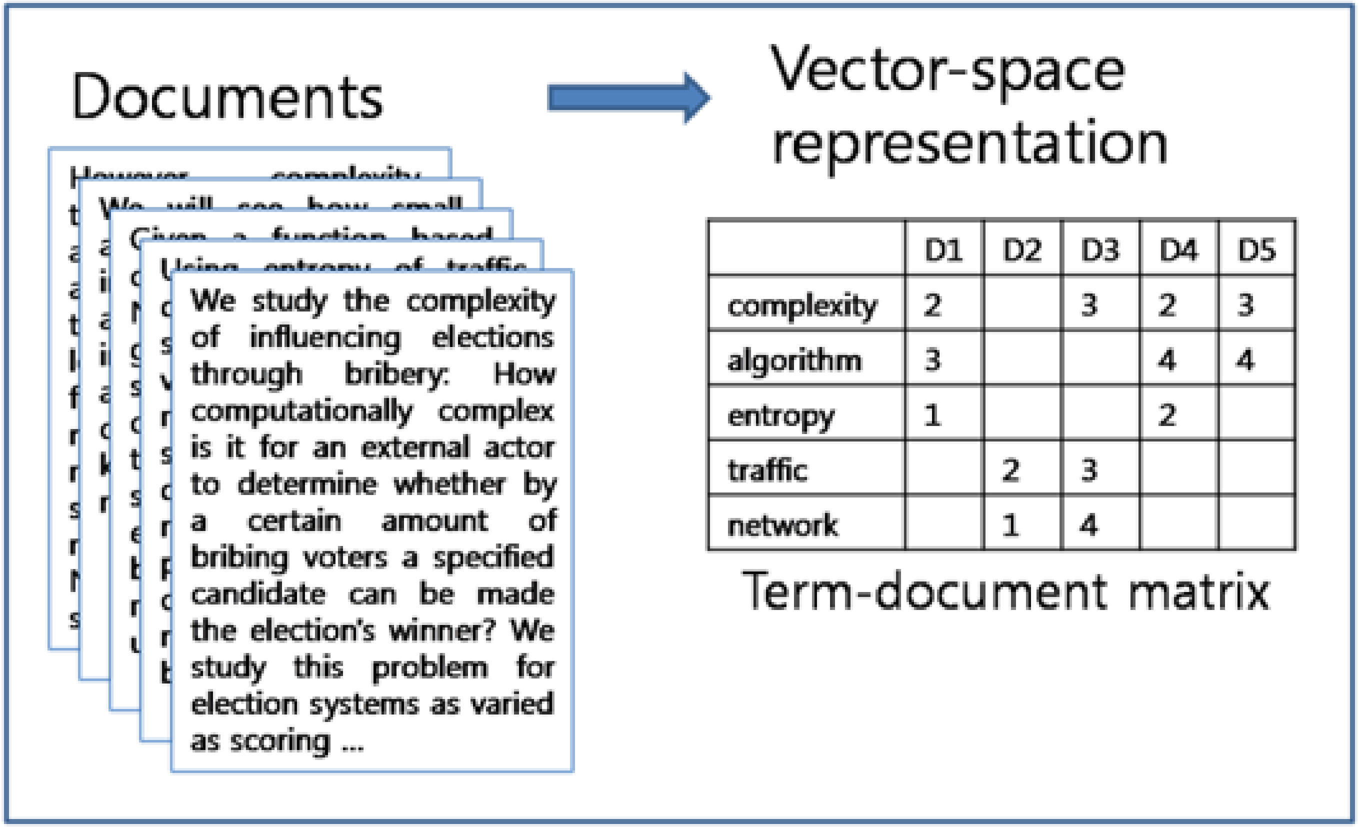
Applications of Matrix Factorizations

Diyah Utami K.P, S.Kom, M.Sc., M.Cs.
Department of Computer Science and Electronifcs
Universitas Gadjah Mada, Yogyakarta

Lab Sistem Cerdas Online Seminar Series #5
Thursday, July 30th 2020

Data Representation



Documents

We study the complexity of influencing elections through bribery: How computationally complex is it for an external actor to determine whether by a certain amount of bribing voters a specified candidate can be made the election's winner? We study this problem for election systems as varied as scoring ...

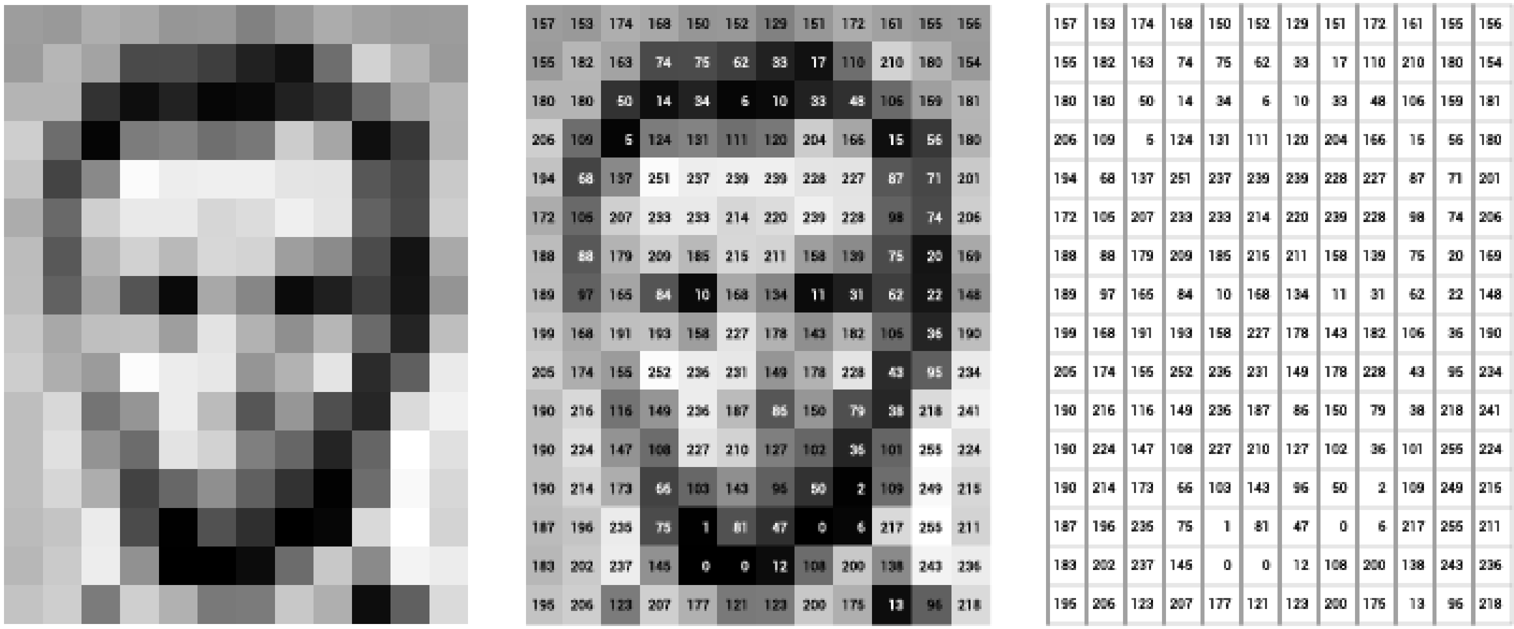
Vector-space representation

	D1	D2	D3	D4	D5
complexity	2		3	2	3
algorithm	3			4	4
entropy	1			2	
traffic		2	3		
network		1	4		

Term-document matrix

Document-Term Matrix

Data Representation



157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

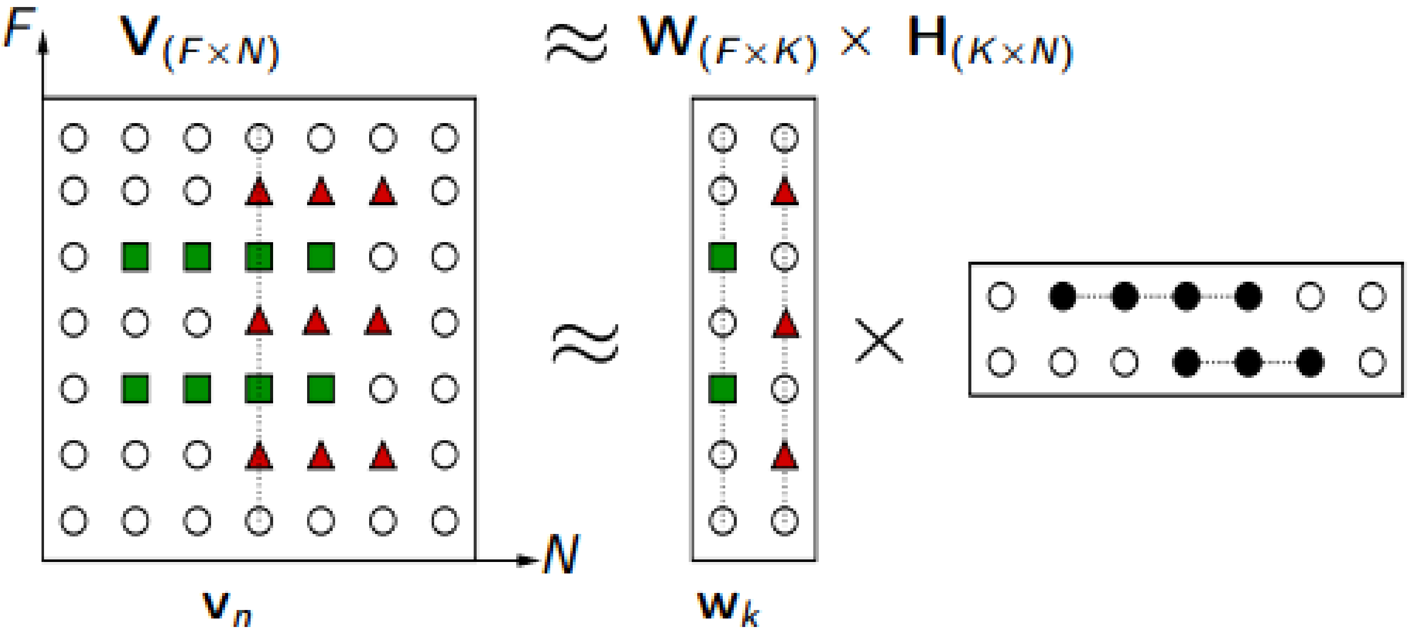
Color Image Representation in Matrix

Dimensionality Reduction

- ☐ Dimensionality reduction is an effective method to represent data by finding factorizations of a matrix.
- ☐ Matrix factorization is one of the methods for dimensionality reduction and it has been applied in many applications.
- ☐ Matrix factorization methods in feature extraction reduce a matrix into constituent parts, that make the algorithm much easier and improve its performance and less computation load.

Explaining Data by Factorization

General Formulation



Data Matrix

Basis Matrix
“explanatory variables”
“regressors”
“dictionary”
“patterns”
“topics”

Coefficient Matrix
“activation coefficients”,
“expansion coefficients

Illustration by C. Févotte

Matrix Factorization Methods

Principal Component Analysis (PCA)

Linear Discriminant Analysis (LDA)

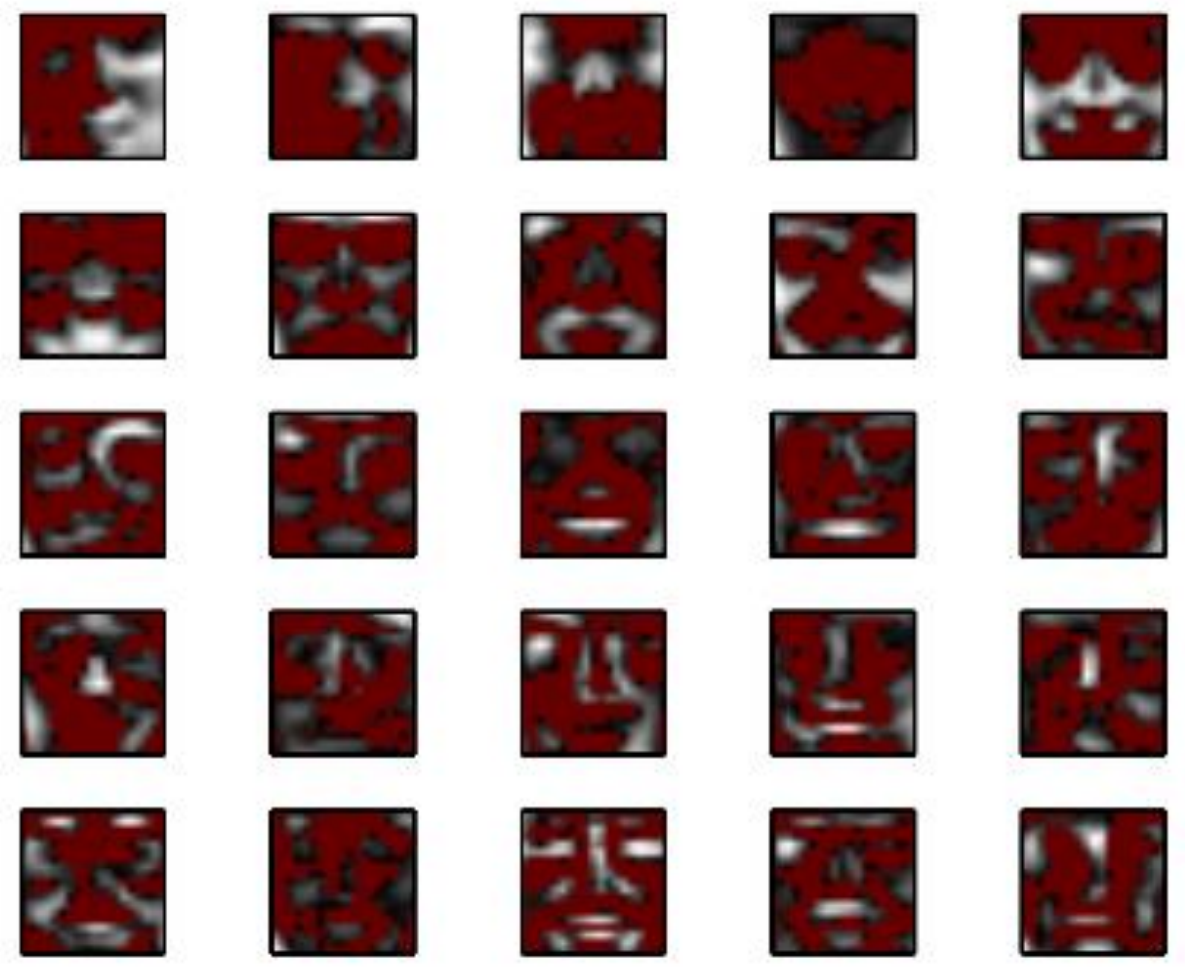
Locally Linear Embedding (LLE)

Independent Component Analysis (ICA)

Partial Least Squares (PLS)

Face Images by PCA

Eigenfaces



Red pixels indicate **negative values!** How to interpret this?

In the Real World

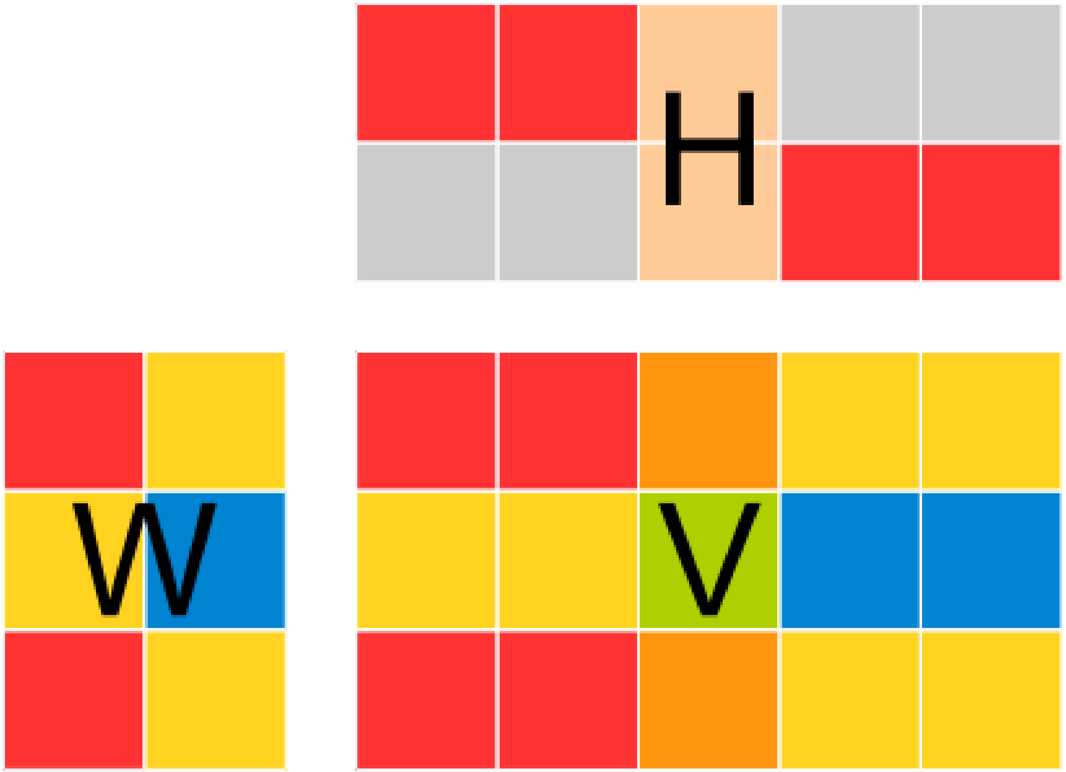
Data is always non-negative by nature

- Image pixel intensities
- Signal intensities
- Occurrence counts
- Gene expression data
- User rating scores
- Chemical compound concentration
- Stock market values
- Food or energy consumption
- ……

The **non-negativity constraints** allows the intuitive interpretation as the real understanding of the original data.

Nonnegative Matrix Factorization

NMF aims to find low dimensional approximation to the original matrix using two **nonnegative** matrices



$$V \approx WH;$$

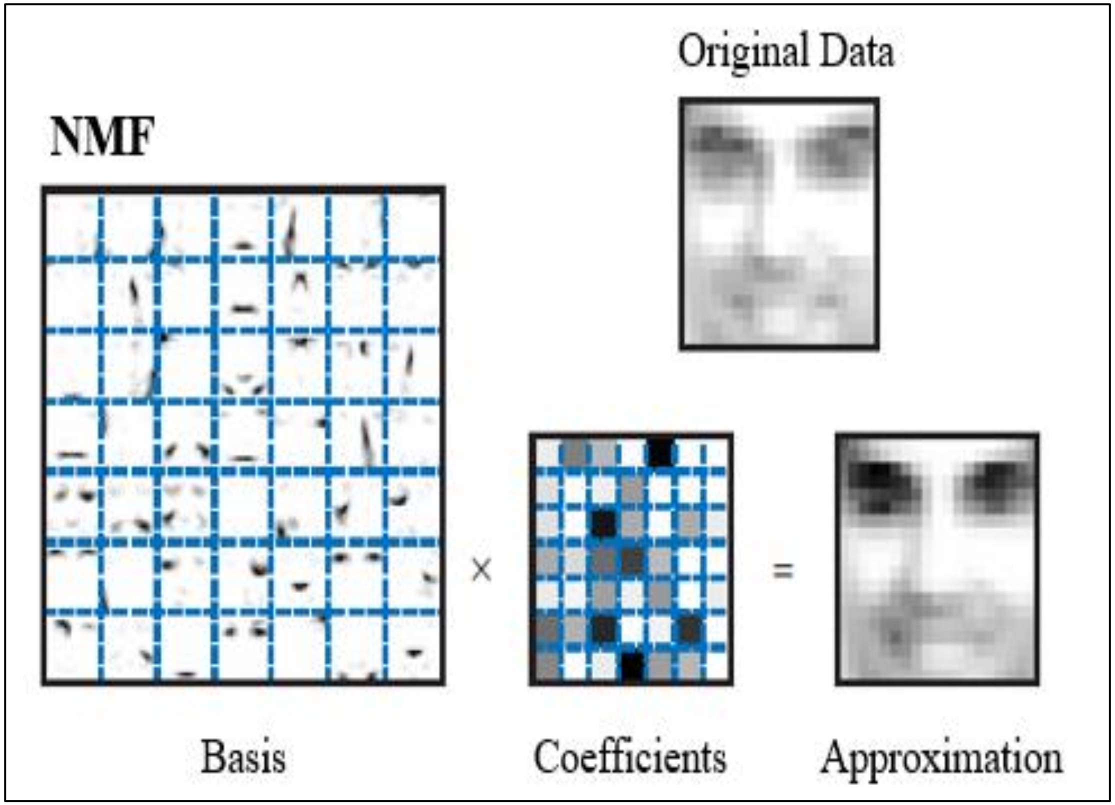
- a **basis matrix** $\mathbf{W} \in \mathbb{R}_+^{\mathbf{F \times K}}$
- a **coefficient matrix** $\mathbf{H} \in \mathbb{R}_+^{\mathbf{K \times N}}$
  - $\mathbf{W} = [w_{fk}]$ s.t. $w_{fk} \geq 0$ and
  - $\mathbf{H} = [h_{kn}]$ s.t. $h_{kn} \geq 0$.

NMF provides an unsupervised linear representation of the data

Illustration by N. Seichepine

Face Images by NMF

Eigenfaces



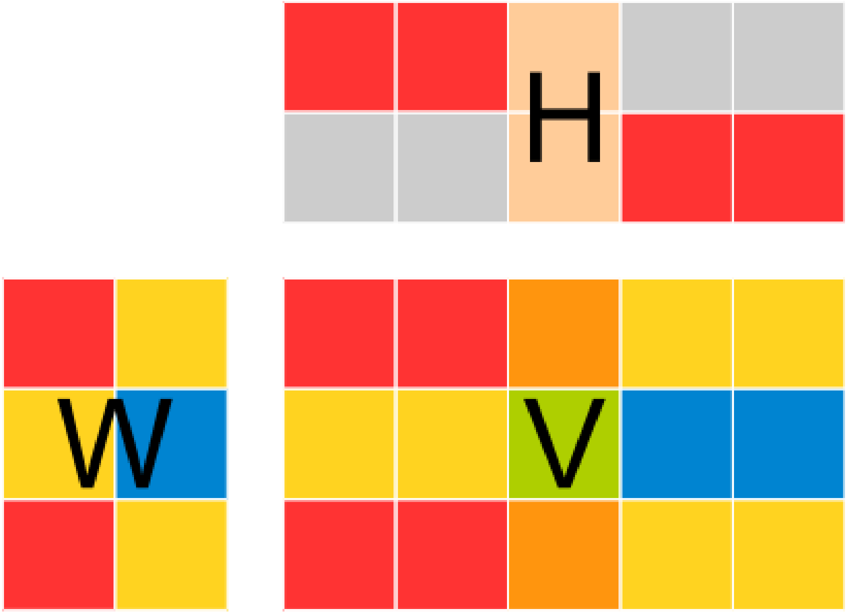
The work of **Lee and Seung (1999)** for **"learning the parts-based representation of faces"** has brought much attention to NMF in the data mining and machine learning field.

- ❑ Nonnegativity leads to **part-based decompositions** of object.
- ❑ Nonnegativity induces **sparsity**.

Figure from Lee and Seung (1999)

Nonnegative Matrix Factorization

Notation I



$V \approx WH;$

- $\mathbf{W} = [w_{fk}]$ s.t. $w_{fk} \geq 0$
  and
- $\mathbf{H} = [h_{kn}]$ s.t. $h_{kn} \geq 0.$

The standard NMF is usually formulated as an optimization:

$$\min_{\mathbf{W},\mathbf{H}} \mathbf{D}(\mathbf{V}\|\mathbf{WH}) \, \text{s.t.} \qquad \mathbf{W} \geq \mathbf{0}, \mathbf{H} \geq \mathbf{0}$$

where $\mathbf{D}(\mathbf{V}\|\mathbf{WH})$ is **divergence function** to measure the distance or error between $\mathbf{V}$ and $\mathbf{WH}$.

Nonnegative Matrix Factorization

Notation II

There are two commonly used **divergence functions** that quantify **the error of the approximation** to solve NMF model.

Euclidean distance (Frobenius norm)

$$D_{Fro}(V\|WH) = \|V - WH\|_F^2 = \sum_{i,j}\left(v_{ij} - \sum_{k=1}^{k} w_{ik}h_{kj}\right)^2$$

Kullback-Leibler divergence

$$D_{KL}(V\|WH) = \sum_{i,j}\left(v_{ij}\log\frac{v_{ij}}{y_{ij}} - v_{ij} + y_{ij}\right)$$

where $Y = [y_{ij}] = WH.$

Nonnegative Matrix Factorization

Optimization Methods

There are several algorithms that have been developed to solve the NMF problem

- Multiplicative Update Rule Algorithm
- Gradient Descent Algorithm
- Alternating Least Squares Algorithm

Optimization methods of NMF produce a sequence of iterations

Multiplicative Update Rule

Algorithm
Multiplicative Update Rule of NMF with Frobenius norm

```
Input: $V \in R_{mxn}$
Output: $W \in R_{mxk}, H \in R_{kxn}$
1: Initialize
      $W^{(0)} \in R_{mxk} = \mathrm{rand}(m, k)$
      $H^{(0)} \in R_{kxn} = \mathrm{rand}(k, n)$
2: for t = 1:maxiter
3:        Update $H^{(t)}_{kn} \leftarrow H^{(t-1)}_{kn} \frac{\left(W^{(t-1)T}V\right)_{kn}}{W^{(t-1)T}_{kn}\left(W^{(t-1)}H\right)_{kn}}$
4:        Update $W^{(t)}_{mk} \leftarrow W^{(t-1)}_{mk} \frac{\left(VH^{(t-1)T}\right)_{mk}}{\left(W^{(t-1)}H\right)_{mk}H^{(t-1)T}_{mk}}$
5:        Convergence condition testing
6:        t= t+1
7: end for
```

Gradient Descent Algorithm

Gradient descent algorithm apply the update rules using **step size** and **the partial derivatives of objective function**.

Algorithm
The Gradient Descent Algorithm of NMF

```
Input: $V \in R_{mxn}$
Output: $W \in R_{mxk}, H \in R_{kxn}$
1: Initialize
    $W^{(0)} \in R_{mxk} = rand(m, k)$
    $H^{(0)} \in R_{kxn} = rand(k, n)$
2: for t = 1: maxiter
3:     Update $H \leftarrow H - \beta_H \frac{\partial O}{\partial H}$
       (nonneg) Set all negative elements in H to 0
4:     Update $W \leftarrow W - \beta_W \frac{\partial O}{\partial W}$
       (nonneg) Set all negative elements in W to 0
5:     Convergence condition testing
6:     t= t+1
7: end for
```

The partial derivatives equation

$$O_{Fro\_NMF}(V\|WH) = O_{Fro\_NMF}(W, H)$$
$$= \|V - WH\|_F^2$$
$$= Tr(V - WH)^T(V - WH)$$
$$= Tr(V^TV - 2H^TW^TV + H^TW^TWH)$$

The **gradient** of the function $\boldsymbol{O_{Fro\_NMF}(W, H)}$

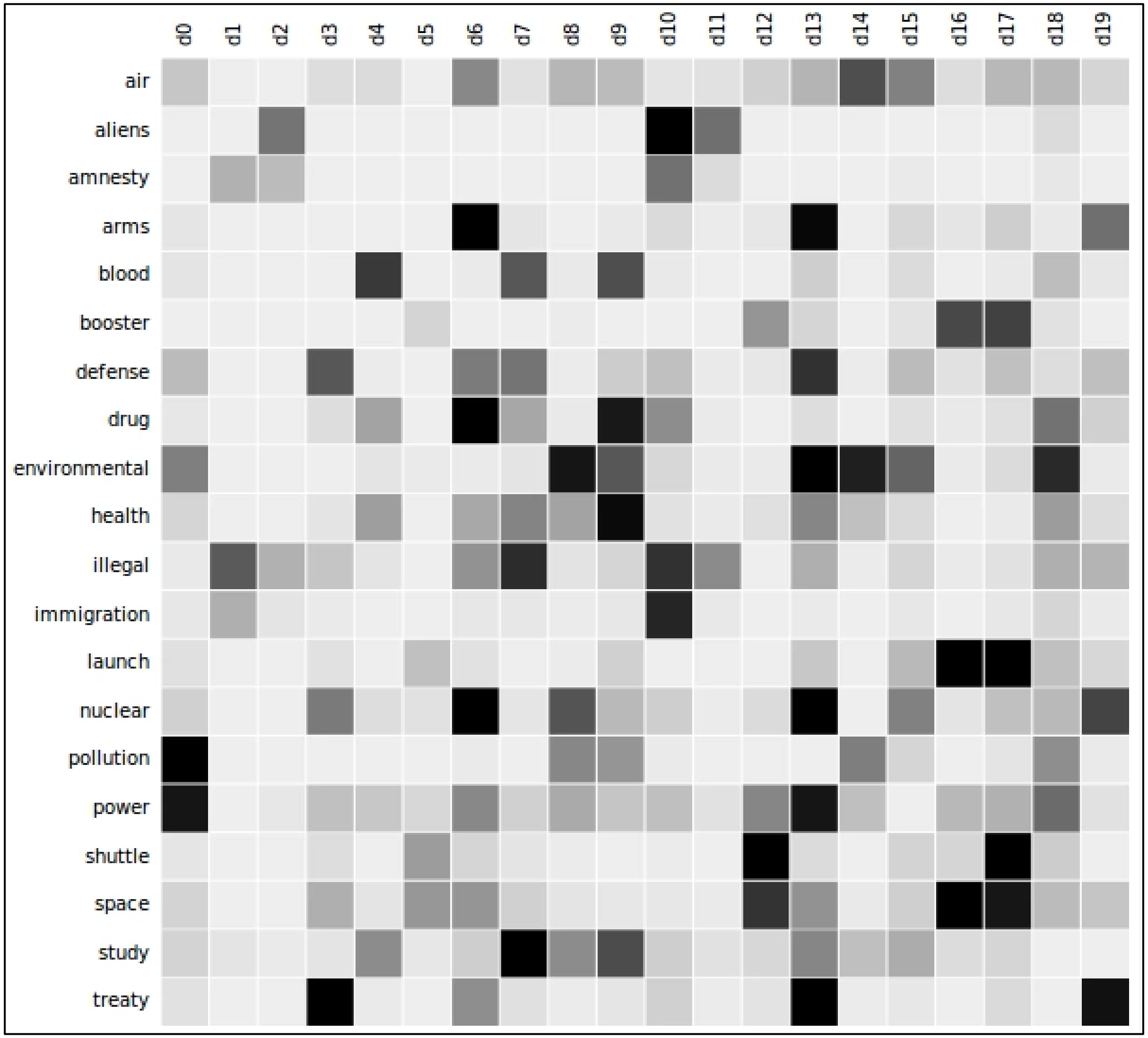
$$\nabla_W O_{Fro\_NMF}(W, H) = \frac{\partial O}{\partial W} = -VH^T + WHH^T$$
$$\nabla_H O_{Fro\_NMF}(W, H) = \frac{\partial O}{\partial H} = -W^TV + W^TWH$$

The **step size parameters** $\boldsymbol{\beta_W}$ and $\boldsymbol{\beta_H}$ vary depending on the algorithm.

UNIVERSITAS
GADJAH MADA

Application of Matrix Factorizations for Topic Modeling and Document Clustering

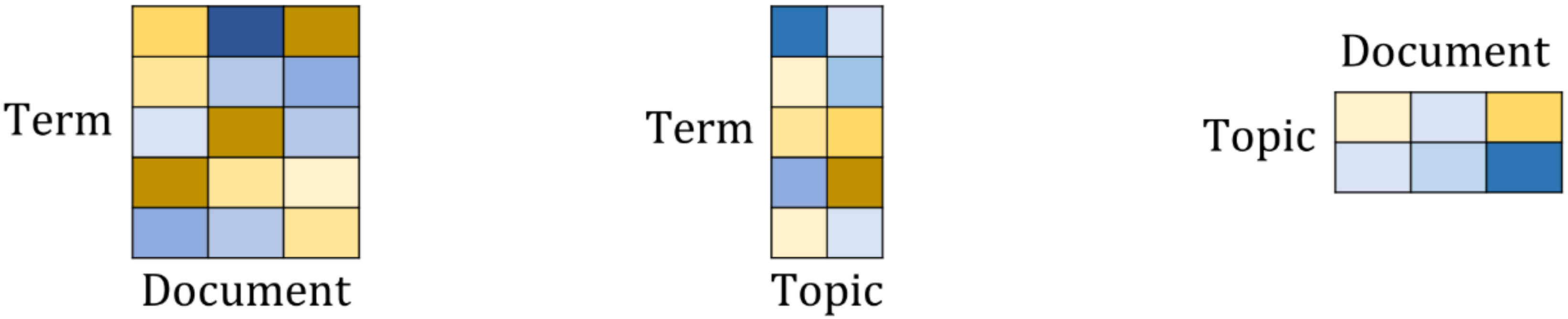
Data Representation



Document-Term Matrix
(https://en.wikipedia.org/wiki/Topic_model)

Nonnegative Matrix Factorization for Topic Modeling and Document Clustering

$$V_{F\times N} \approx W_{F\times K} \times H_{K\times N}$$



Document-Term Matrix

	d_1	d_2	d_3
term_1	1	0	0
term_2	0	1	1
term_n	2	0	1

Basis/Features Matrix

	topic_1	topic_2
term_1	0.5	0
term_2	0	0.5
term_n	1	0

Coefficient Matrix

	d_1	d_2	d_3
topic_1	1	0	0
topic_2	0	1	1

https://www.cc.gatech.edu/~hpark/papers/nmf_book_chapter.pdf

Matrix Factorization Methods for Topic Modeling and Document Clustering

Non-negative Matrix Factorization (NMF)

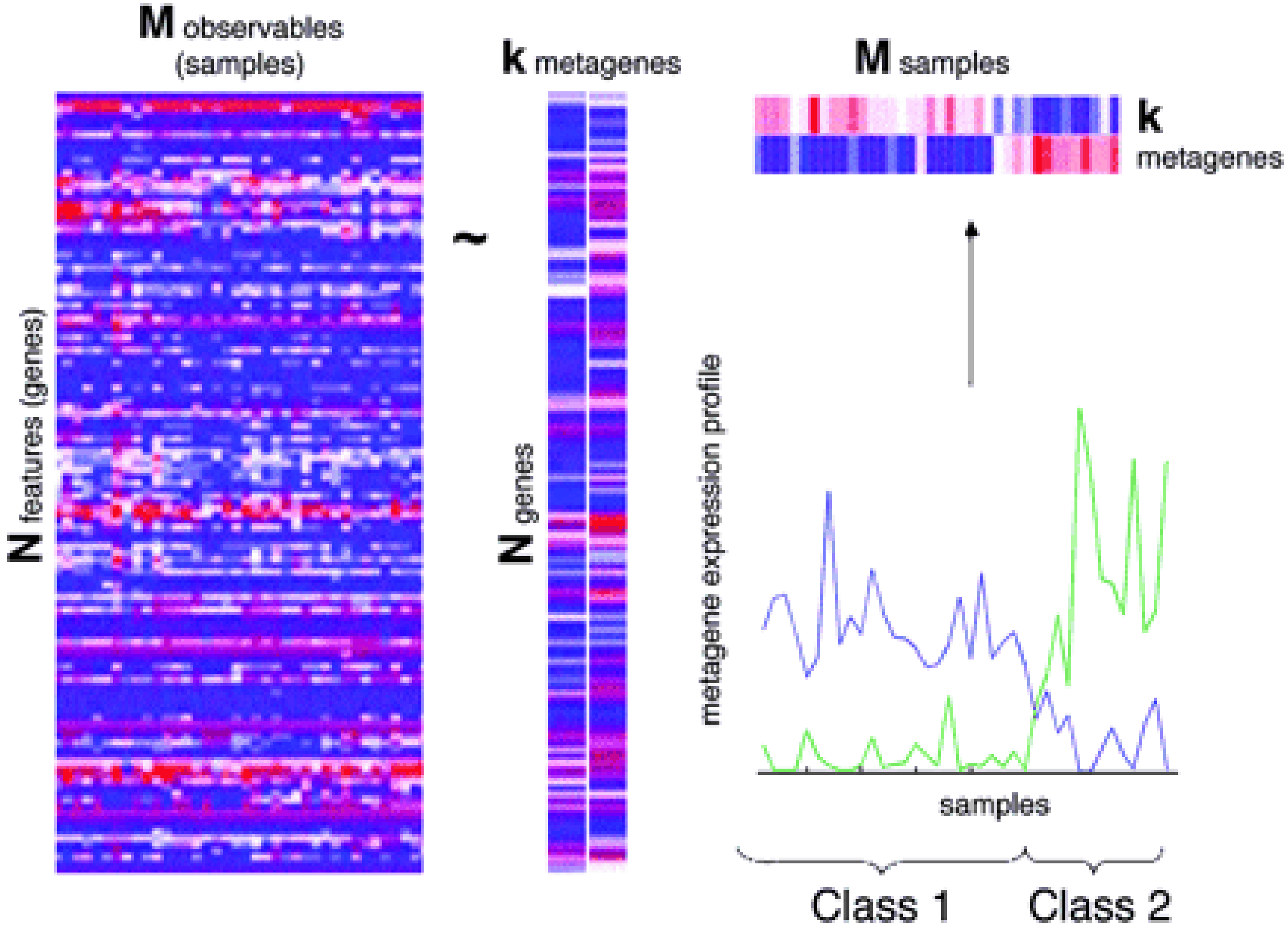
Latent Dirichlet Allocation (LDA)

Singular Value Decomposition (SVD)

Applications of Matrix Factorizations for Gene Expression Analysis

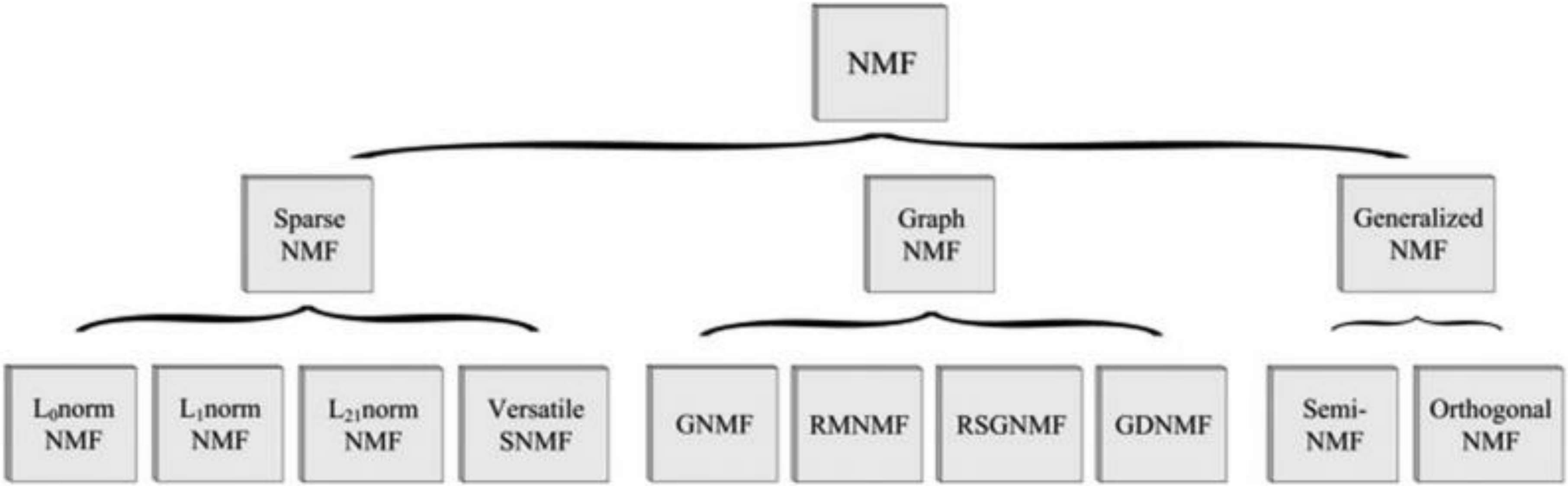
Nonnegative Matrix Factorization for Clustering Tumor Subtypes

A (rank M) ~ W H (rank k=2)



https://doi.org/10.1073/pnas.0308531101

Matrix Factorization Methods for Clustering Tumor Subtypes



NMF

Sparse NMF

Graph NMF

Generalized NMF

L_0norm NMF

L_1norm NMF

L_{21}norm NMF

Versatile SNMF

GNMF

RMNMF

RSGNMF

GDNMF

Semi-NMF

Orthogonal NMF

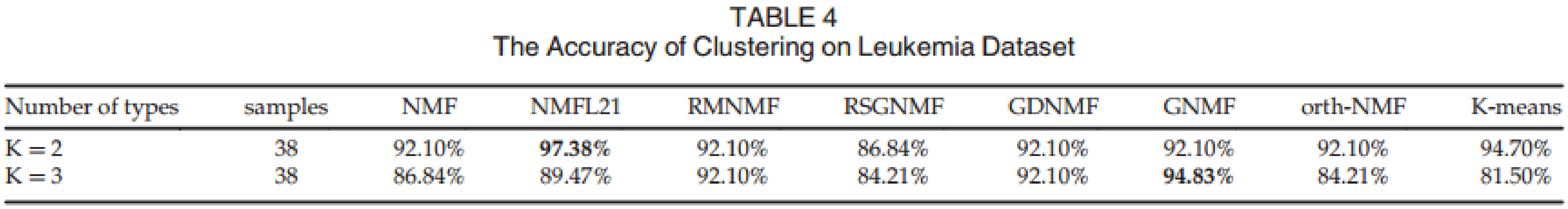


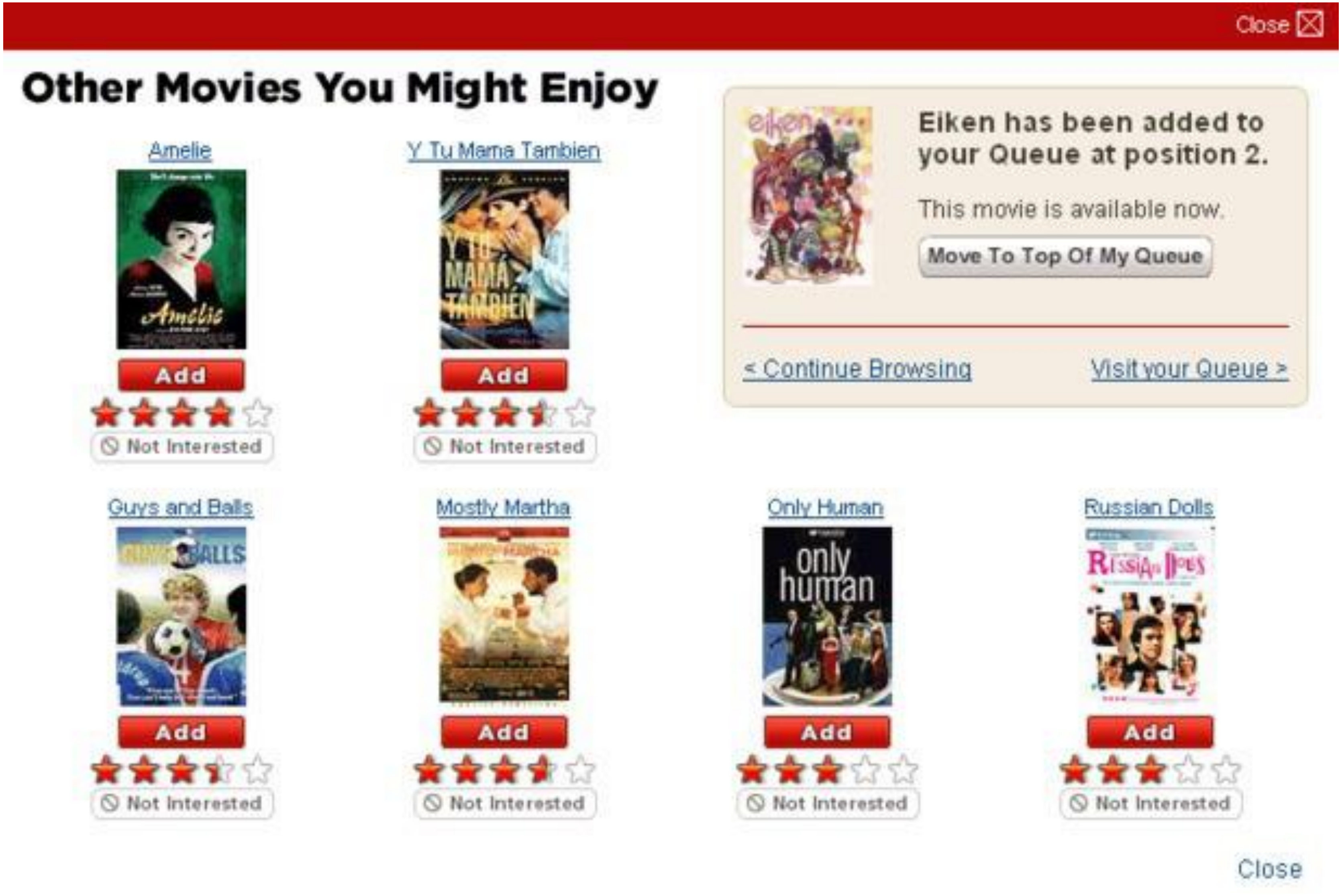
TABLE 4
The Accuracy of Clustering on Leukemia Dataset

Number of types	samples	NMF	NMFL21	RMNMF	RSGNMF	GDNMF	GNMF	orth-NMF	K-means
K = 2	38	92.10%	**97.38%**	92.10%	86.84%	92.10%	92.10%	92.10%	94.70%
K = 3	38	86.84%	89.47%	92.10%	84.21%	92.10%	**94.83%**	84.21%	81.50%

https://www.researchgate.net/publication/313454264

Applications of Matrix Factorizations for Collaborative Filtering Recommendation System

Netflix

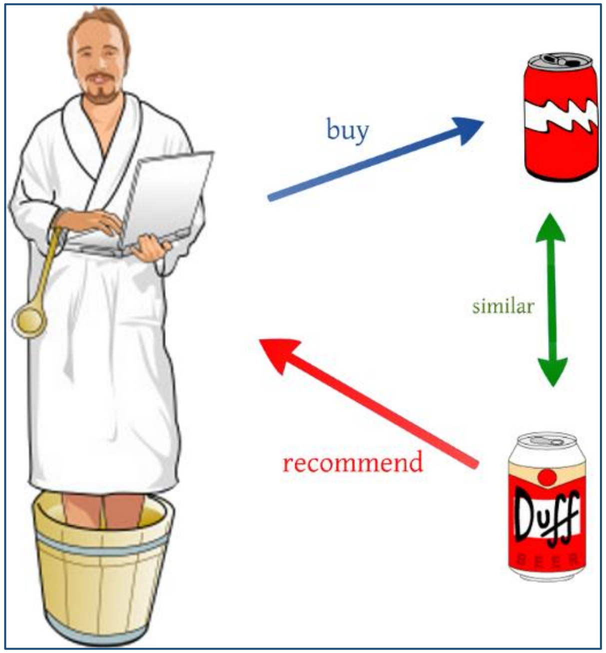


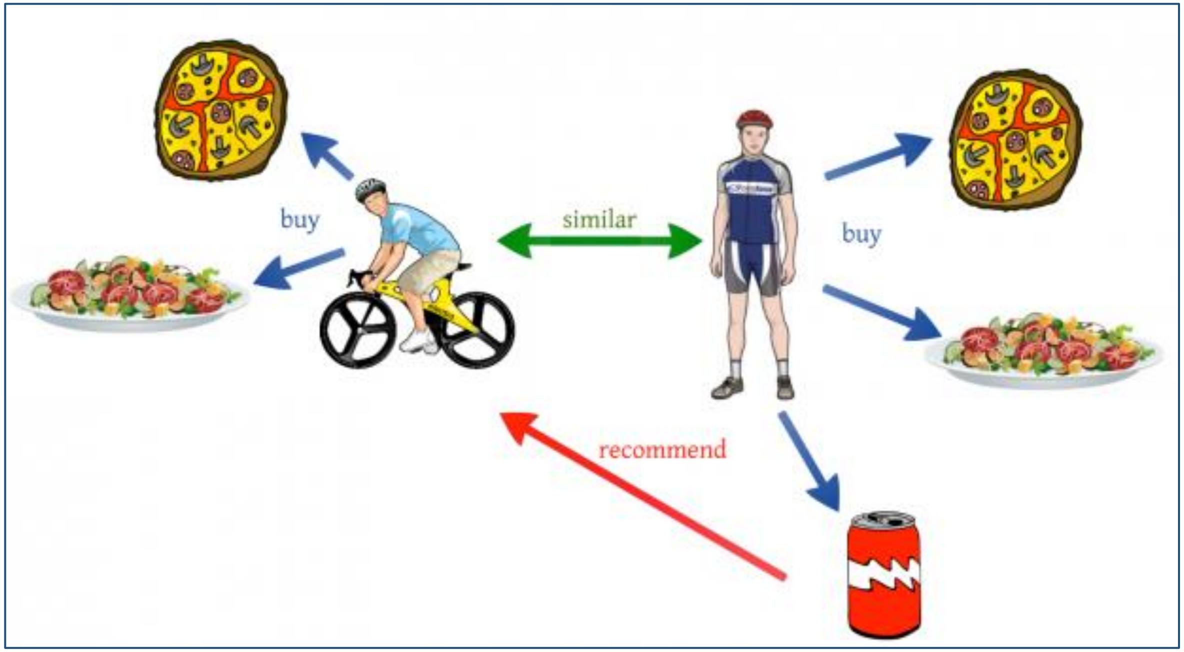
Other Movies You Might Enjoy

Eiken has been added to your Queue at position 2.

This movie is available now.

Recommender System Techniques



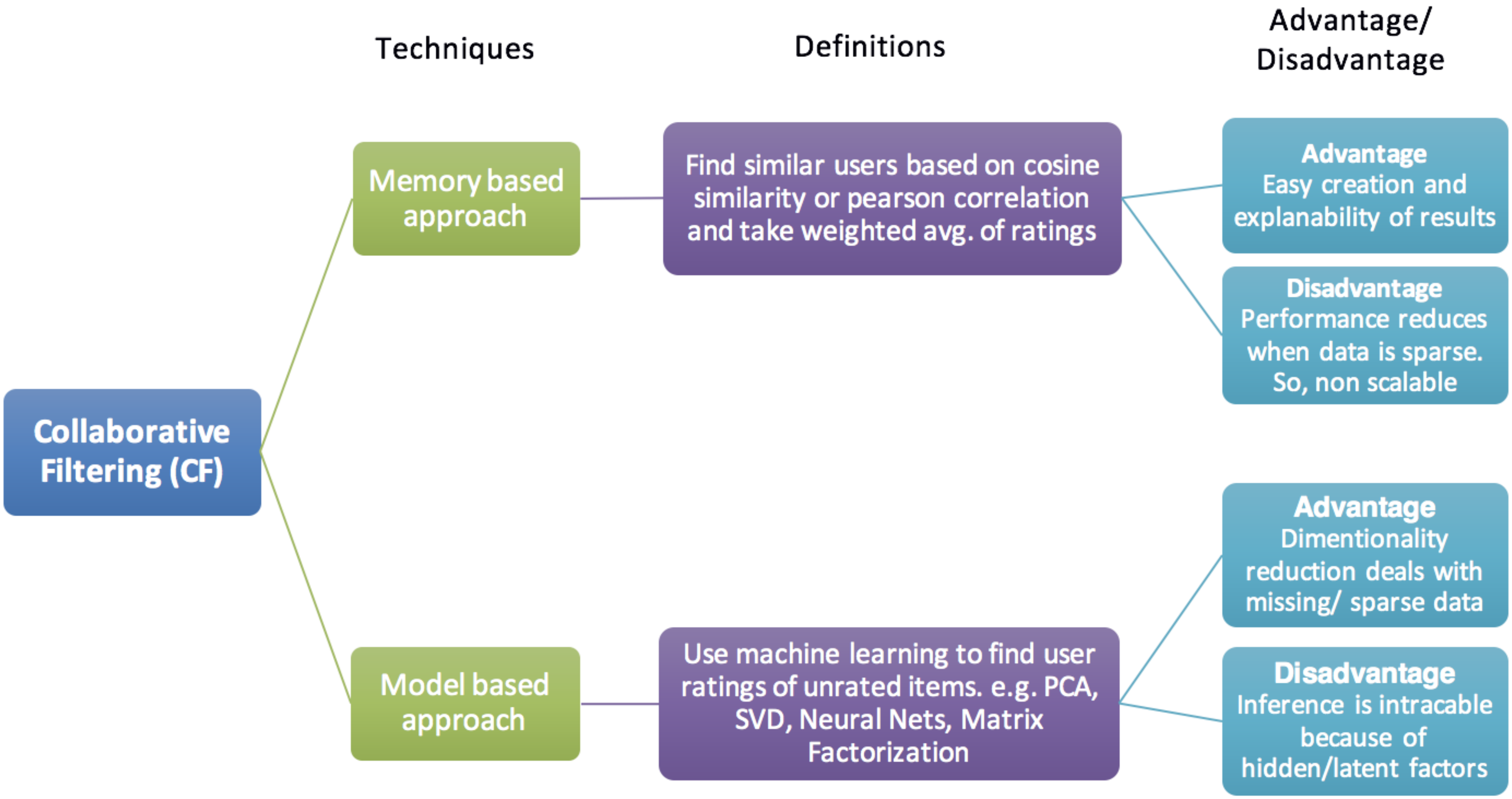


Content-Based Filtering
recommend the item based on the similarity of item that highly rated by user before

Collaborative Filtering
recommend the item based on based on the idea that people who share the same interest in certain kind of items will also share the same interest in some other kind of items

https://d4datascience.wordpress.com/2016/07/22/recommender-systems-101/

Collaborative Filtering Approaches



Techniques

Definitions

Advantage/ Disadvantage

Collaborative Filtering (CF)

Memory based approach

Find similar users based on cosine similarity or pearson correlation and take weighted avg. of ratings

Advantage
Easy creation and explanability of results

Disadvantage
Performance reduces when data is sparse. So, non scalable

Model based approach

Use machine learning to find user ratings of unrated items. e.g. PCA, SVD, Neural Nets, Matrix Factorization

Advantage
Dimentionality reduction deals with missing/ sparse data

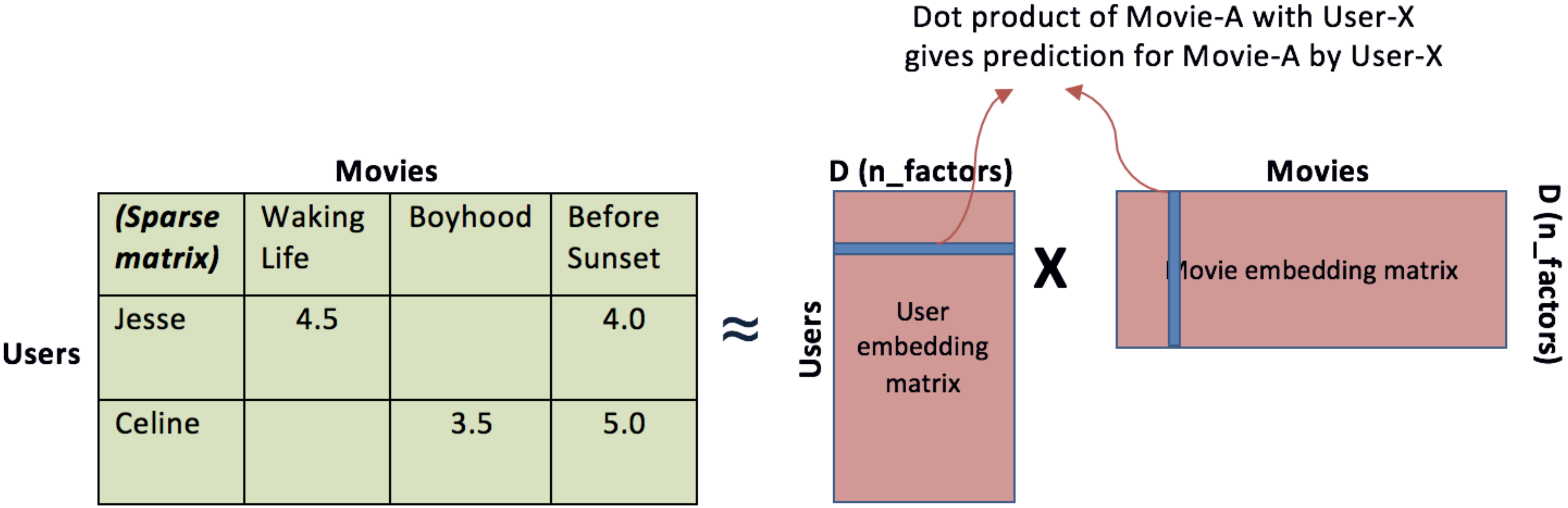
Disadvantage
Inference is intracable because of hidden/latent factors

Types of collaborative filtering approaches

https://en.wikipedia.org/wiki/Collaborative_filtering

Matrix Factorizations for Collaborative Filtering Recommendation System

$$V_{F\times N} \approx W_{F\times K} \times H_{K\times N}$$



Dot product of Movie-A with User-X gives prediction for Movie-A by User-X

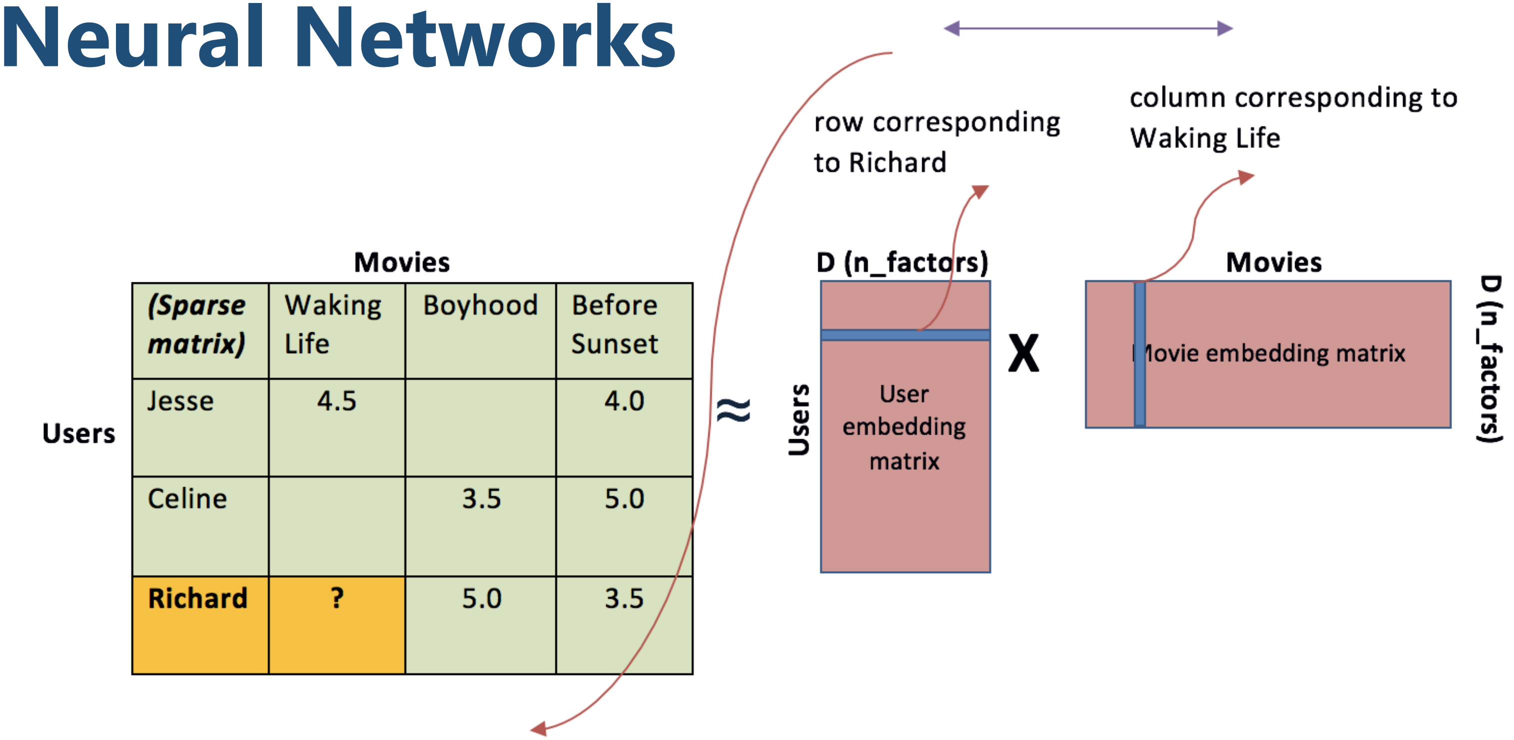
(Sparse matrix)	Waking Life	Boyhood	Before Sunset
Jesse	4.5		4.0
Celine		3.5	5.0

Visualization of Matrix Factorization

Assumes that latent factors exist in user/movies

https://towardsdatascience.com/various-implementations-of-collaborative-filtering-100385c6dfe0

Matrix factorization and Embeddings for Neural Networks



Movies

Users

(Sparse matrix)	Waking Life	Boyhood	Before Sunset
Jesse	4.5		4.0
Celine		3.5	5.0
Richard	**?**	5.0	3.5

Linear and nonlinear layers

Input (X) → Target (Y)

Users	**Movies**
Jesse	Waking Life
Jesse	Boyhood
Jesse	Before Sunset
Celine	Waking Life
Celine	Boyhood
Celine	Before Sunset
Richard	**Waking Life**
Richard	Boyhood
Richard	Before Sunset

User latent features	**Movie latent features**
0.2, 0.4, 2.8, 4.8, 2.4	0.1, 0.5, 5.0, 3.7, 2.8
……	….
……	….
……	….
……	….
……	….
2.2, 1.4, 2, 1.8, 4.4	0.1, 0.25, 4.5, 3.1, 2
……	….
……	….

Ratings
4.5
4.0
3.5
5.0
?
5.0
3.5

Matrix Factorization Methods for Collaborative Filtering Recommendation System

Non-negative Matrix Factorization (NMF)

Probabilistic Matrix Factorization (PMF)

Singular Value Decomposition (SVD)

Principal Component Analysis (PCA)

Deep Matrix Factorization (DMF)

Applications of Matrix Factorizations for Image Processing

Face Images by NMF

Eigenfaces

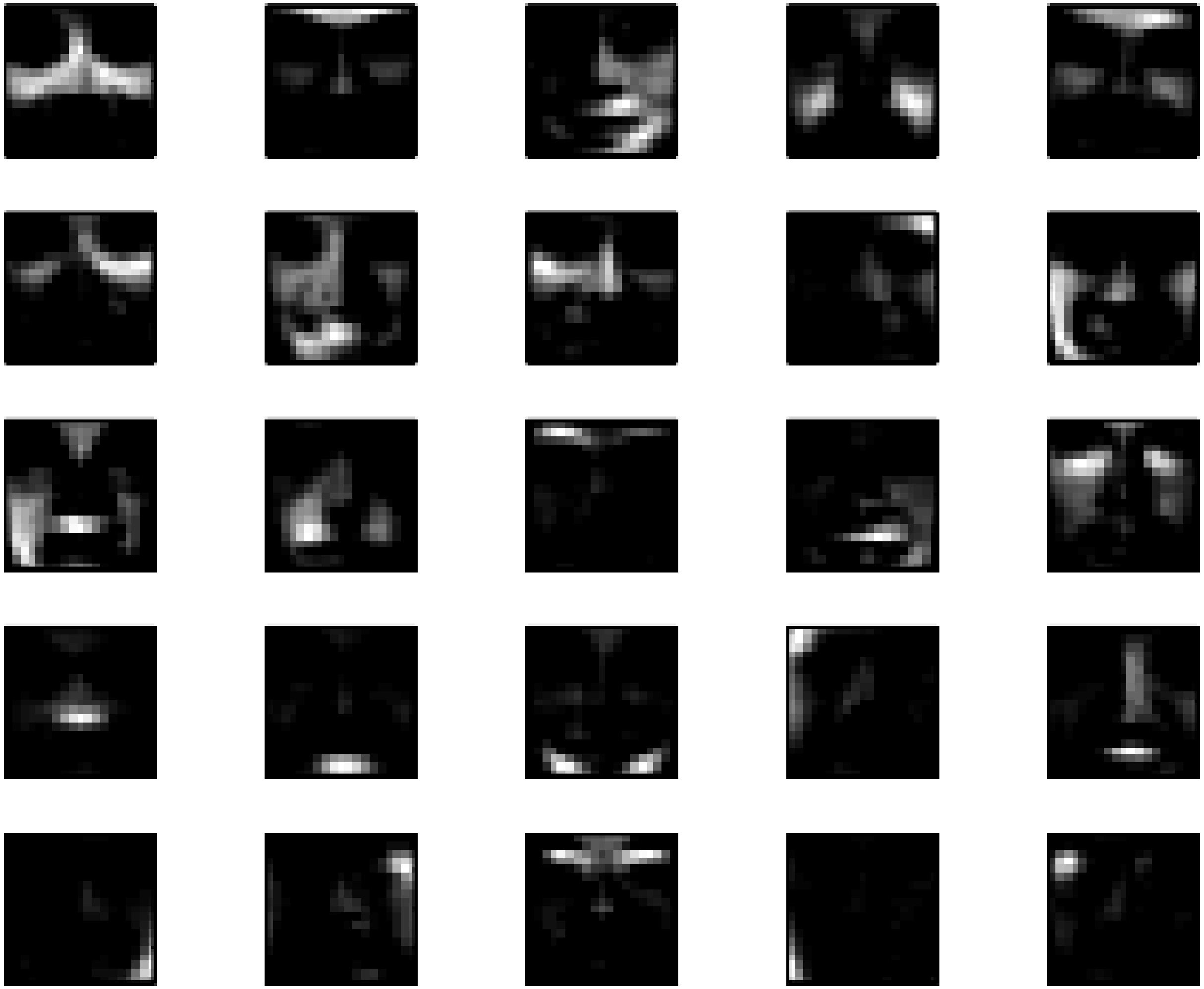


Illustration by C. Fevotte

Image Reconstruction

using Eigenfaces

Original images corrupted by occlusion from ORL dataset

NMF_EUD

GNMF

NMF_KL

SpatialNMF

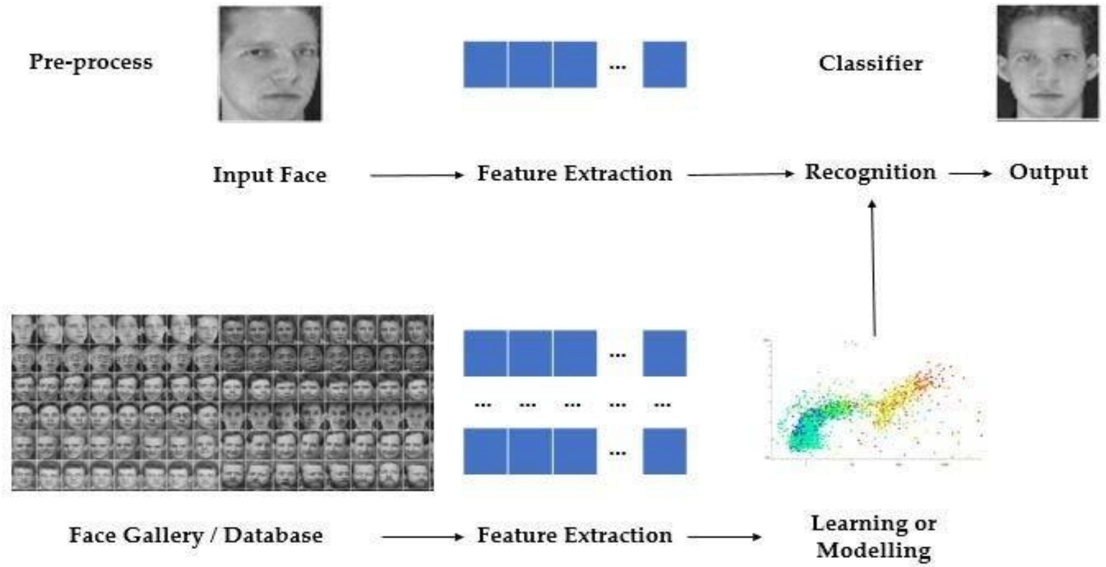
NMF_PGD

CoupledNMF

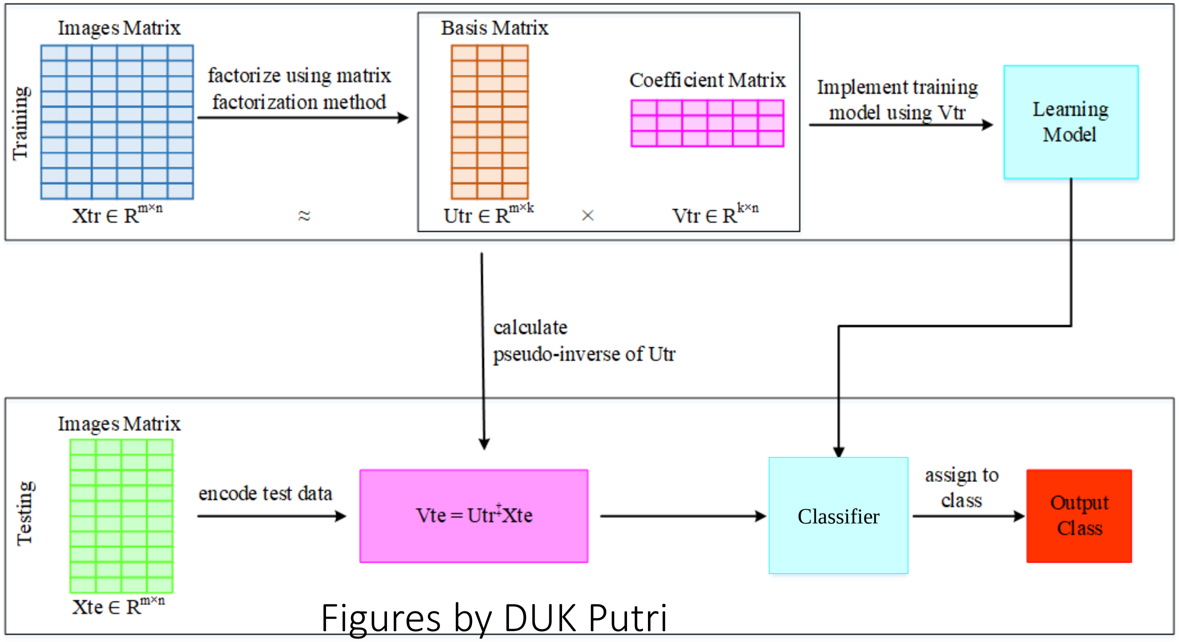
SNMF

Figures by DUK Putri

Face Recognition



Figures by DUK Putri



Figures by DUK Putri

Extension of NMF

Complex Matrix Factorization

CMF will be decomposed the complex-valued matrix into two matrices of bases and coefficients.

- ☐ The real data matrix is transformed into a complex number based on the **Euler representation**
- ☐ *Wirtinger*'s calculus is used to compute derivative of the cost function.
- ☐ The gradient descent method is used to solve complex matrix factorization problems.

Transform Real to Complex Field

- Let the input data matrix $V = (V_1, V_2, \ldots, V_N)$ contains *N* data vectors as columns.
- Using the **Euler's formula**, the elements of real matrix V are normalized and transformed into a complex number field to yield the complex data matrix Z.

The mapping function $f: \mathbb{R}^M \rightarrow \mathbb{C}^M$ defined by:

$$Z_t = f(V_t) = \frac{1}{\sqrt{2}} e^{ia\pi V_t} = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{ia\pi V_t(1)} \\ \vdots \\ e^{ia\pi \boldsymbol{V}_t(M)} \end{bmatrix}$$

where the Euler's formula is

$$e^{ia\pi \boldsymbol{V}_t} = cos(a\pi V_t) + i\, sin(a\pi V_t)$$

- The nonlinear function f is a special feature mapping which transform the real valued features to complex feature space.

Complex Matrix Factorization

Notation I

- ❑Given a matrix $\mathrm{Z} \in \mathrm{C}^{\mathrm{m} \times \mathrm{n}}$, find two matrices $\mathrm{W} \in \mathrm{C}^{\mathrm{m} \times \mathrm{k}}$ and $\mathrm{H} \in \mathrm{C}^{\mathrm{k} \times \mathrm{n}}$ that minimize the objective function using Euclidean distance (Frobenius norm)

$$\mathrm{O}_{\mathrm{CMF}}(\mathrm{W}, \mathrm{H}) = \frac{1}{2} ||\mathrm{Z} - \mathrm{WH}||_{\mathrm{F}}^{2}$$

- ❑These problem can be solved by using **block coordinate descent (BCD)** with two matrix block alternatingly to obtain a local solution

Complex Matrix Factorization

Notation II

- **With W fixed**, the optimization objective function is modified as follows

$$\min_H O(H) = \min_H \left(\frac{1}{2}\|Z - WH\|_F^2\right)$$

- To solve the subproblem, the one variable $O(H)$ is treated as bivariable function $O(H, H^*)$

$$O(H, H^*) = \frac{1}{2} Tr(Z^H Z - 2(H^*)^T W^H Z + (H^*)^T W^H W H)$$

- The **gradient descent method** is utilized to update the current solution $H^{(t)}$ to $H^{(t+1)}$

$$\boldsymbol{H^{(\mathrm{t+1})} = H^{(\mathrm{t})} - \beta^{(\mathrm{t})} \nabla_{H^*} \mathrm{O}\left(H^{(\mathrm{t})}, H^{*(\mathrm{t})}\right)}$$

- where $\beta^{(t)}$ is the step size. Backtracking line search (Armijo rule) is used to estimate step size
- The first order partial derivative with respect to H^* for have the form

$$\nabla_{H^*} O(H, H^*) = -W^H Z + W^H W V$$

- **With H fixed**, the update rule for obtaining W is $W = H^{\ddagger} Z$, where $\ddagger$ denoted the Moore-Penrose pseudoinverse.

Complex Matrix Factorization

Stopping Condition

Algorithm
Complex Gradient Descent Algorithm

```
Input: Z, H, $0 < \mu < 1, 0 < \sigma < 1$
Output: H
1. Initialize any feasible $H^{(0)}$
Set $\beta^{(0)} = 1$
2. repeat
(a) Compute gradient $\nabla_{H^*} O\left(H^{(t)}, H^{*(t)}\right)$
(b) for t = 1, 2, ….
    (b1) Assign $\beta^{(t)} \leftarrow \beta^{(t-1)}$
    (b2) if $\beta^{(t)}$ satisfies
            repeat
              $\beta^{(t)} = \beta^{(t)}/\mu$
            until either $\beta^{(t)}$ does not satisfy
              or $H\left(\beta^{(t)}/\mu\right) = H\left(\beta^{(t)}\right)$
         else
            repeat
              $\beta^{(t)} = \beta^{(t)}\mu$
            until $\beta^{(t)}$ satisfies
    (b3) set $H^{(t+1)} = H^{(t)} - \beta^{(t)} \nabla_{H^*} O\left(H^{(t)}, H^{*(t)}\right)$
until Stopping criterion is satisfied
```

The **complex gradient descent (CGD) method** for optimizing the objective function.

An ideal criterion can be used to stop the iterative process when a **local minimum of the objective function is reached**.

The stopping condition is adapted by applying the follow condition

$$\|\nabla_{H^*} O(H)\|_F \leq \varepsilon$$

where ε is pre-defined threshold.

Extensions of CMF

1. **Sparse Complex Matrix Factorization using Ridge Term (SCMF-L_2)** which used L_2-norm regularization to provide smoothness of the coefficient matrix.
2. **Spatial Complex Matrix Factorization (SpatialCMF)** used spatial locality using pixel dispersion penalty to the basis matrix.
3. **Coupled Complex Matrix Factorization (CoupledCMF)** used combination of pixel images representation and class activity annotation

Results of Face Recognition Rate

un-occluded ORL dataset with different subspace dimensions

k	NMF_ EUD	NMF_ KL	NMF_ PGD	CMF	SNMF	SCMF	SCMF-L_2	GNMF	GCMF	Spatial NMF	Spatial CMF	Coupled NMF	Coupled CMF
10	0,8781 ± 0,0246	0,8720 ± 0,0242	0,8789 ± 0,0277	0,9194 ± 0,0258	0,8532 ± 0,0350	0,9216 ± 0,0250	0,9298 ± 0,0206	0,8913 ± 0,0284	0,9183 ± 0,0227	0,1373 ± 0,0684	0,9084 ± 0,0248	0,9412 ± 0,0879	0,9933 ± 0,0093
20	0,9134 ± 0,0246	0,9081 ± 0,0234	0,9118 ± 0,0260	0,9398 ± 0,0181	0,9132 ± 0,0280	0,9437 ± 0,0177	0,9475 ± 0,0147	0,9212 ± 0,0255	0,9406 ± 0,0223	0,2708 ± 0,2533	0,9308 ± 0,0170	0,9657 ± 0,0417	**0,9892 ± 0,0093**
30	0,9322 ± 0,0216	0,9307 ± 0,0176	0,9343 ± 0,0220	0,9585 ± 0,0188	0,9242 ± 0,0236	0,9545 ± 0,0168	0,9601 ± 0,0148	0,9361 ± 0,0225	0,9552 ± 0,0166	0,9377 ± 0,0164	0,9444 ± 0,0237	0,9666 ± 0,0246	**0,9902 ± 0,0053**
40	0,9334 ± 0,0247	0,9349 ± 0,0223	0,9376 ± 0,0218	0,9548 ± 0,0196	0,9315 ± 0,0203	0,9567 ± 0,0167	0,9619 ± 0,0199	0,9388 ± 0,0169	0,9585 ± 0,0174	0,9392 ± 0,0232	0,9488 ± 0,0167	0,9767 ± 0,0109	**0,9875 ± 0,0075**
50	0,9390 ± 0,0140	0,9362 ± 0,0190	0,9400 ± 0,0195	0,9510 ± 0,0169	0,9288 ± 0,0188	0,9546 ± 0,0178	0,9600 ± 0,0132	0,9453 ± 0,0198	0,9507 ± 0,0209	0,9459 ± 0,0194	0,9430 ± 0,0215	0,9785 ± 0,0101	**0,9860 ± 0,0039**
60	0,9371 ± 0,0204	0,9276 ± 0,0227	0,9362 ± 0,0185	0,9535 ± 0,0191	0,9288 ± 0,0227	0,9529 ± 0,0173	0,9609 ± 0,0168	0,9456 ± 0,0190	0,9567 ± 0,0182	0,9470 ± 0,0149	0,9431 ± 0,0213	0,9773 ± 0,0078	**0,9860 ± 0,0051**
70	0,9330 ± 0,0238	0,9358 ± 0,0152	0,9285 ± 0,0210	0,9517 ± 0,0206	0,9270 ± 0,0224	0,9501 ± 0,0222	0,9607 ± 0,0114	0,9397 ± 0,0203	0,9498 ± 0,0184	0,9470 ± 0,0131	0,9502 ± 0,0187	0,9776 ± 0,0109	**0,9855 ± 0,0059**
80	0,9322 ± 0,0207	0,9354 ± 0,0197	0,9225 ± 0,0263	0,9522 ± 0,0172	0,9306 ± 0,0197	0,9483 ± 0,0178	0,9614 ± 0,0119	0,9507 ± 0,0204	0,9517 ± 0,0171	0,9535 ± 0,0134	0,9385 ± 0,0221	0,9780 ± 0,0084	**0,9850 ± 0,0062**
90	0,9317 ± 0,0247	0,9232 ± 0,0192	0,9386 ± 0,0184	0,9533 ± 0,0140	0,9218 ± 0,0223	0,9505 ± 0,0154	0,9616 ± 0,0118	0,9525 ± 0,0150	0,9512 ± 0,0199	0,9464 ± 0,0166	0,9443 ± 0,0151	0,9775 ± 0,0093	**0,9835 ± 0,0075**
100	0,9332 ± 0,0247	0,9194 ± 0,0245	0,9229 ± 0,0203	0,9509 ± 0,0168	0,9270 ± 0,0244	0,9477 ± 0,0152	0,9613 ± 0,0162	0,9525 ± 0,0148	0,9496 ± 0,0200	0,9473 ± 0,0175	0,9424 ± 0,0238	0,9774 ± 0,0107	**0,9829 ± 0,0075**
AVG	0,9263	0,9223	0,9251	0,9485	0,9186	0,9481	0,9565	0,9374	0,9482	0,7972	0,9394	0,9717	**0,9869**
MAX	0,9390	0,9362	0,9400	0,9585	0,9315	0,9567	0,9619	0,9525	0,9585	0,9535	0,9502	0,9785	**0,9933**

Results of research by DUK Putri

Results of Face Recognition Rate

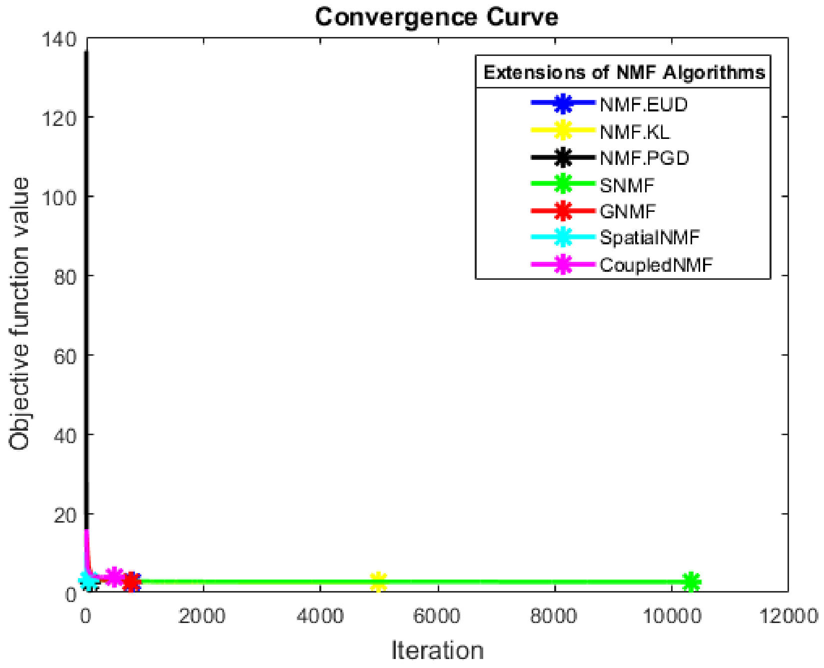
Occluded ORL dataset with different Patch Sizes

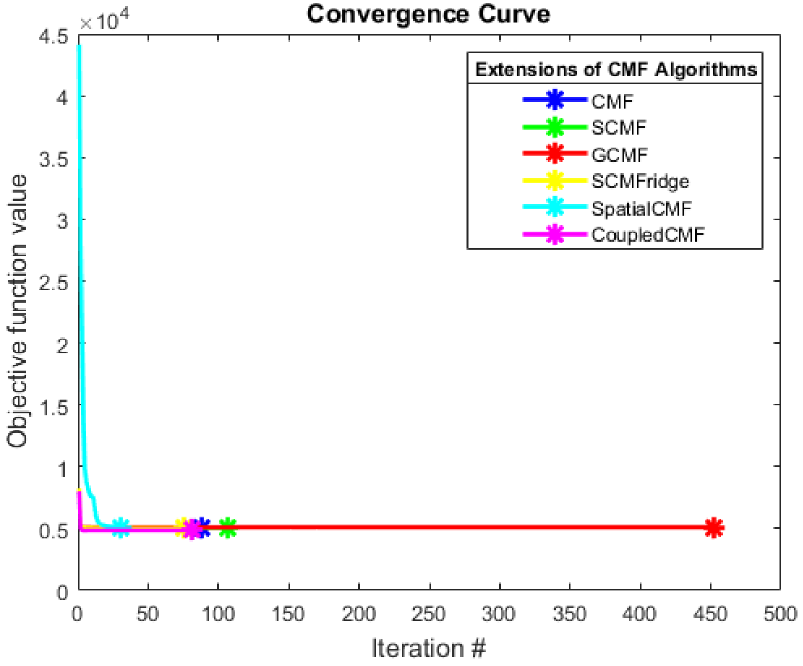
Patch	NMF_ EUD	NMF_ KL	NMF_ PGD	CMF	SNMF	SCMF	SCMF-L$_2$	GNMF	GCMF	Spatial NMF	Spatial CMF	Coupled NMF	Coupled CMF
15x15	0,8296 ± 0,0270	0,8366 ± 0,0303	0,8314 ± 0,0325	0,9314 ± 0,0181	0,8157 ± 0,0263	0,9319 ± 0,0275	0,9444 ± 0,0208	0,8610 ± 0,0239	0,9320 ± 0,0181	0,8456 ± 0,0370	0,9176 ± 0,0267	0,9823 ± 0,0091	**0,9872 ± 0,0050**
20x20	0,6340 ± 0,0368	0,6510 ± 0,0462	0,6461 ± 0,0339	0,8736 ± 0,0234	0,6371 ± 0,0398	0,8650 ± 0,0210	0,8794 ± 0,0296	0,7086 ± 0,0296	0,8605 ± 0,0267	0,6641 ± 0,0386	0,8513 ± 0,0390	0,9836 ± 0,0089	**0,9882 ± 0,0074**
25x25	0,4090 ± 0,0335	0,4380 ± 0,0307	0,4173 ± 0,0347	0,7009 ± 0,0311	0,4225 ± 0,0567	0,6946 ± 0,0426	0,7202 ± 0,0389	0,4922 ± 0,0555	0,6889 ± 0,0392	0,4492 ± 0,0525	0,6858 ± 0,0528	0,9880 ± 0,0115	**0,9925 ± 0,0052**
30x30	0,2972 ± 0,0379	0,3291 ± 0,0359	0,3118 ± 0,0457	0,5501 ± 0,0489	0,3044 ± 0,0361	0,5397 ± 0,0355	0,5525 ± 0,0373	0,3513 ± 0,0378	0,5346 ± 0,0404	0,3312 ± 0,0457	0,5063 ± 0,0553	0,9878 ± 0,0105	**0,9915 ± 0,0043**
AVG	0,5425	0,5637	0,5517	0,7640	0,5449	0,7578	0,7741	0,6033	0,7540	0,5725	0,7403	0,9854	0,9899

Results of research by DUK Putri

Convergence Curves

of Extensions of NMF and CMF Algorithms





Figures by DUK Putri

Various Research for NMF Applications

- **Text mining:**
  - (Xu et al., 2003; Berry and Browne, 2006; Kim and Park, 2008)
- **Images:**
  - unsupervised object discovery (Sivic et al., 2005)
  - object and face recognition (Soukup and Bajla, 2008)
  - tagging (Kalayeh et al., 2014)
  - denoising and inpainting (Mairal et al., 2010)
  - texture classification (Sandler and Lindenbaum, 2011)
  - spectral data (Berry et al.)
  - hashing (Monga and Mihcak, 2007)
  - watermarking (Lu et al., 2009)
- **Electroencephalography (EEG) data:**
  - feature extraction (Cichocki and Rutkowski, 2006; Lee et al., 2009)
  - artifact rejection (Damon et al., 2013a,b)

https://www.cs.rochester.edu/u/jliu/CSC-576/NMF-tutorial.pdf

Various Research for NMF Applications

- **Audio and music processing**
  - Source separation (speech) (Virtanen, 2007; Virtanen and Cemgil, 2009; Mohammadiha et al., 2013)
  - Source separation (music) (Durrieu et al., 2009; Ozerov and Fevotte, 2010; Hennequin et al., 2011; Ozerov et al., 2013; Rafii et al., 2013)
  - Signal enhancement/denoising (Wilson et al., 2008; Schmidt et al., 2007; Sun and Mazumder, 2013)
  - Audio inpainting (Roux et al., 2011; Yilmaz et al., 2011)
  - Compression (Ozerov et al., 2011b; Nikunen et al., 2011)
  - Music transcription (Smaragdis and Brown, 2003; Abdallah and Plumbley, 2004; Vincent et al., 2007; E. Vincent et al., 2008; Févotte et al., 2009; Bertin et al., 2010; Vincent et al., 2010)

https://www.cs.rochester.edu/u/jliu/CSC-576/NMF-tutorial.pdf

Various Research for NMF Applications

- ❑Video processing
    - ❑Video summarization (Cooper and Foote, 2002)
    - ❑Dynamic video content representation and scene change detection (Bucak and Gunsel, 2007)
    - ❑Onscreen person spotting and shot-type classification (Essid and Fevotte, 2012, 2013)
    - ❑Fingerprinting (Cirakman et al., 2010)
    - ❑Action recognition (Krausz and Bauckhage, 2010; Masurelle et al., 2014)
    - ❑Compression (Türkan and Guillemot, 2011)

https://www.cs.rochester.edu/u/jliu/CSC-576/NMF-tutorial.pdf

Various Research for NMF Applications

- **Bioinformatics:**
    - gene expression analysis (Brunet et al., 2004; Gao and Church, 2005)
    - protein interaction clustering (Greene et al., 2008)
- **Other:**
    - collaborative filtering (Melville and Sindhwai, 2010)
    - community discovery (Wang et al., 2010)
    - portfolio diversification (Drakakis et al., 2007)
    - food consumption analysis (Zetlaoui et al., 2010)
    - industrial source apportionment (Limem et al., 2013)

https://www.cs.rochester.edu/u/jliu/CSC-576/NMF-tutorial.pdf

THANK YOU