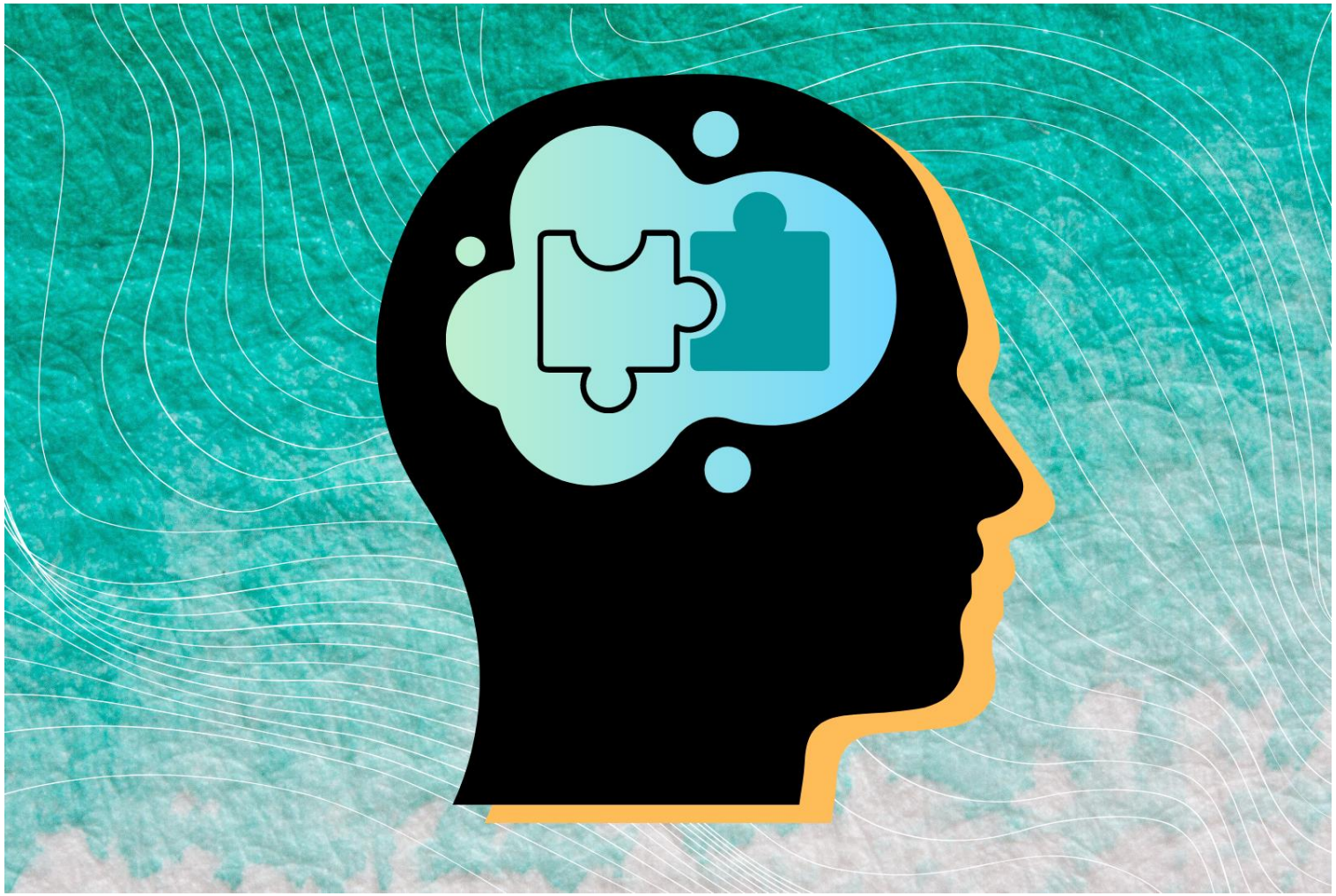
NON NEGATIVE MATRIX FACTORISATION

Made By

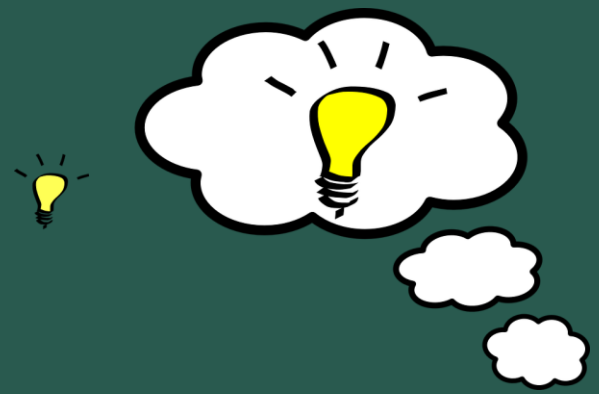
Siddhant Chandel(22M0400)

Prateek Kumar(22M0396)

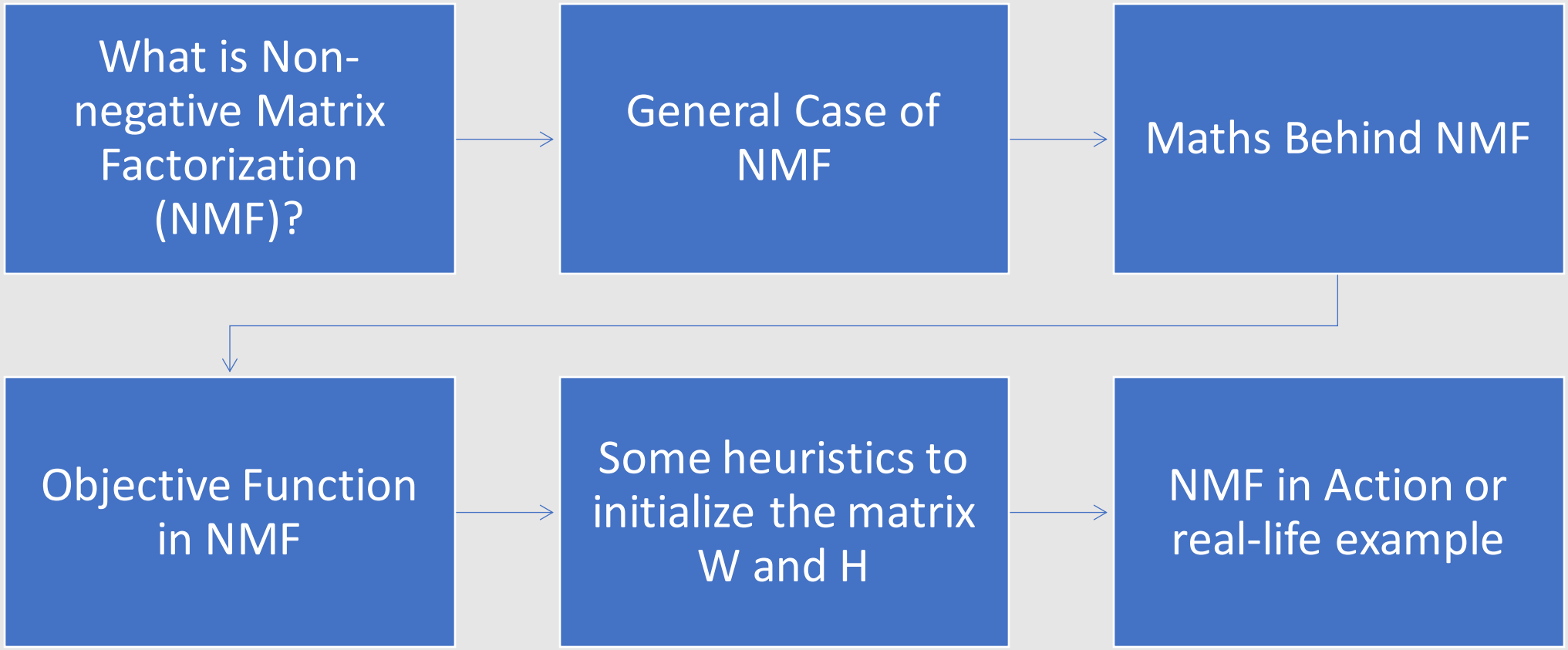
Vivek Kumar(22M0430)



Is perception of whole based on perception of its parts?



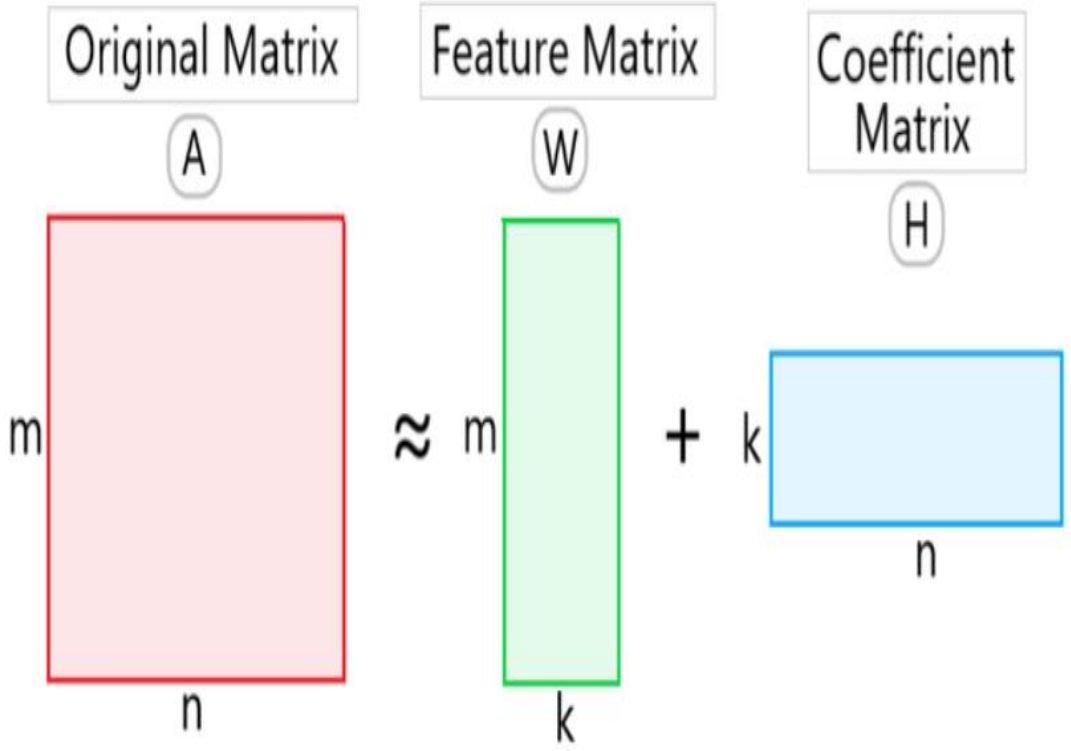
Content


What is Non-negative Matrix Factorization (NMF)?
General Case of NMF
Maths Behind NMF
Objective Function in NMF
Some heuristics to initialize the matrix W and H
NMF in Action or real-life example

What is NMF?

The objective of NMF is dimensionality reduction and feature extraction.

the goal of NMF is to find two matrices $W \in R(m \times k)$ and $H \in R(n \times k)$ having only nonnegative elements.



NMF Algorithms

$$f(\mathbf{W},\mathbf{H}) = \|\mathbf{V} - \mathbf{WH}\|_F^2$$
$$s.t., \mathbf{W} \geq 0, \mathbf{H} \geq 0$$

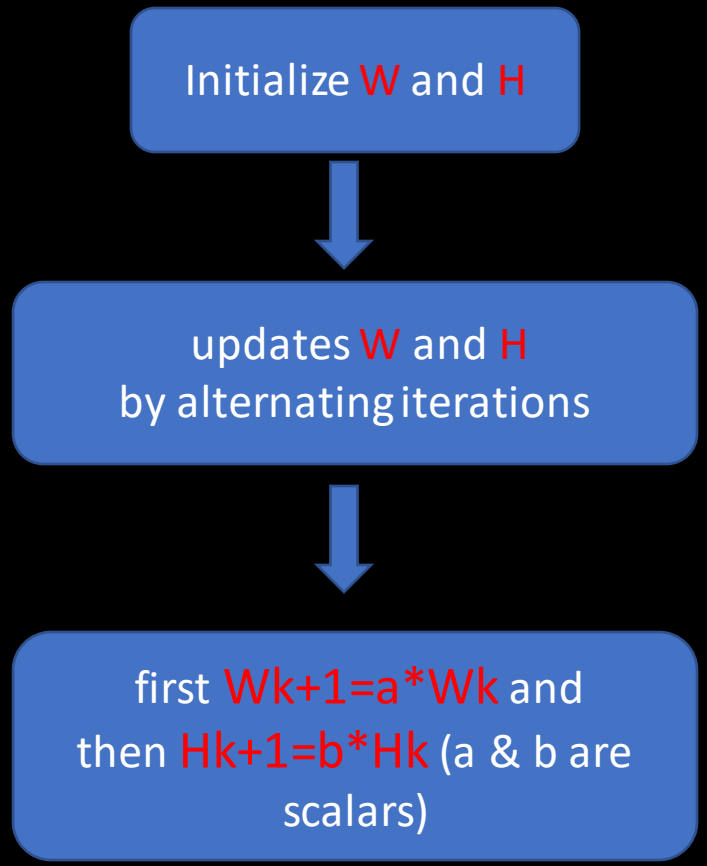
- Multiplicative update(MU)
- Alternating non-negative least squares method(ANLS)
- Newton like method

  & many more....

NMF Algorithms

$$f(\mathbf{W},\mathbf{H}) = \|\mathbf{V} - \mathbf{WH}\|_F^2$$
$$s.t., \mathbf{W} \geq 0, \mathbf{H} \geq 0$$

- Multiplicative update(MU)



NMF Algorithms

$$f(\mathbf{W},\mathbf{H}) = \|\mathbf{V} - \mathbf{WH}\|_F^2$$
$$s.t., \mathbf{W} \geq 0, \mathbf{H} \geq 0$$

- **Alternating non-negative least squares method(ANLS)**

Decomposes NMF problem into two suboptimization problems with non-negative constraints

↓

Solve only for H keeping W constant

↓

Update H and solve for W keeping H (updated) as constant

NMF Algorithms

$$f(\mathbf{W},\mathbf{H}) = \|\mathbf{V} - \mathbf{WH}\|_F^2$$
$$s.t., \mathbf{W} \geq 0, \mathbf{H} \geq 0$$

- **Newton like method**

first approximates the objective function around the current iterate Hk by using the quadratic model

Minimise the quadratic model and obtain Hbar

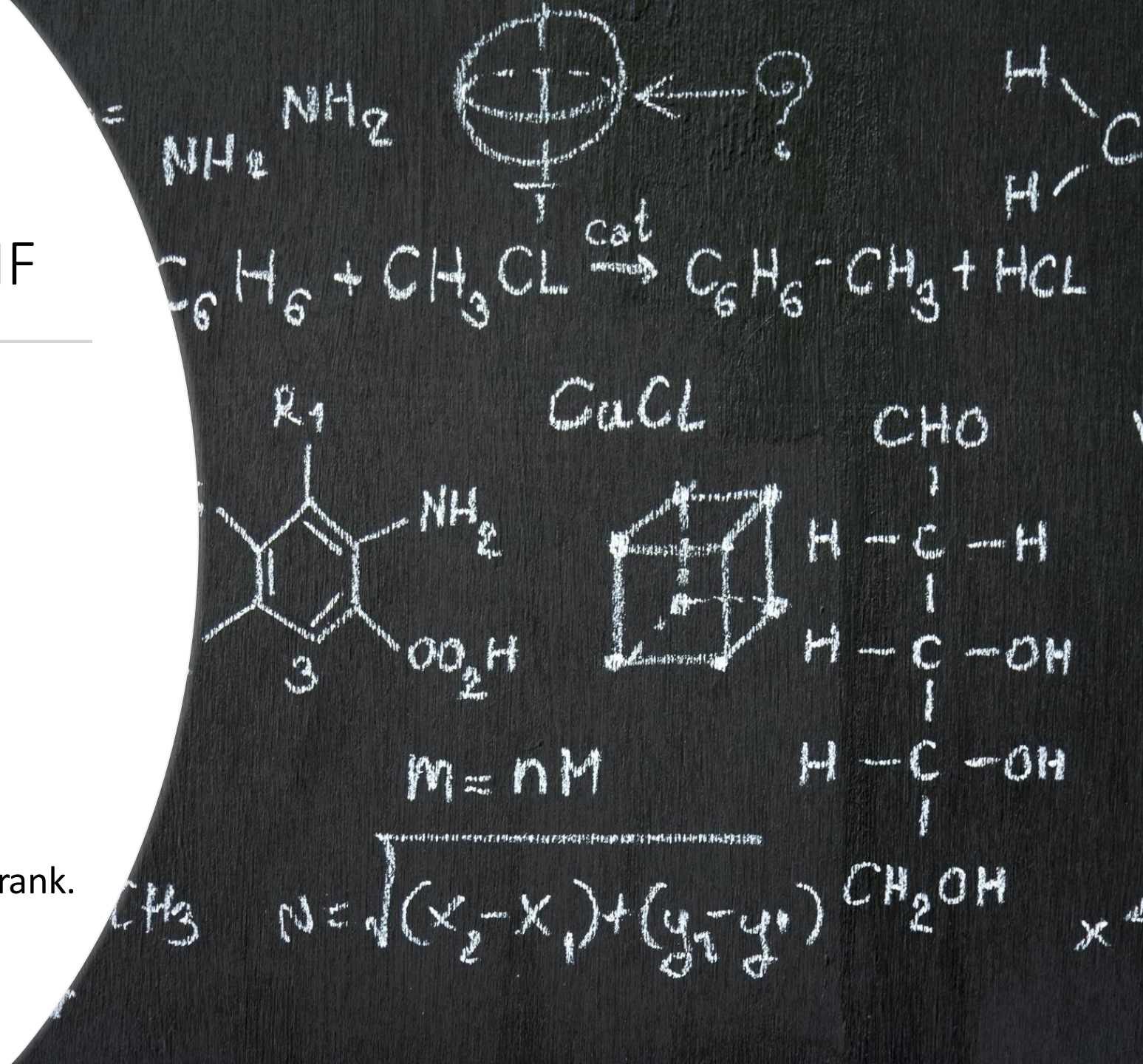
Update Hk+1 = Hk + β(Hbar k – Hk)

Objective Function in NMF

- We show here an algorithm to minimize least square error by developing an auxiliary function .
- Non-negative constraint for matrix factorization
- V =WH

  i.e V is nxm matrix where n is n dimensional data vector set and m is number of example in the set.

- W= nxr and H is rxm, r in the variable rank.



Cost Function

- Euclidean Distance between A and B ,

J = $$\|A - B\|^2 = \sum_{ij} (A_{ij} - B_{ij})^2$$

- Minimization of J wrt W and H , subject to constraint W,H > 0.

$$W_{ia} \leftarrow W_{ia} \frac{(V H^T)_{ia}}{(W H H^T)_{ia}}$$

$$H_{a\mu} \leftarrow H_{a\mu} \frac{(W^T V)_{a\mu}}{(W^T W H)_{a\mu}}$$

Algorithm for NMF based on iterative updates of W and H because they are easy to code and use.

Ex- Astronomical images

At each iteration new W and H is multiplied by a current value with some factor that depend on quality of approximation.

Quality of approximation improves with each iteration

Heuristics to initialize

- Using additive update rule:

$$H_{a\mu} \leftarrow H_{a\mu} + \eta_{a\mu}\left[(W^T V)_{a\mu} - (W^T W H)_{a\mu}\right].$$

where

$$\eta_{a\mu} = \frac{H_{a\mu}}{(W^T W H)_{a\mu}},$$

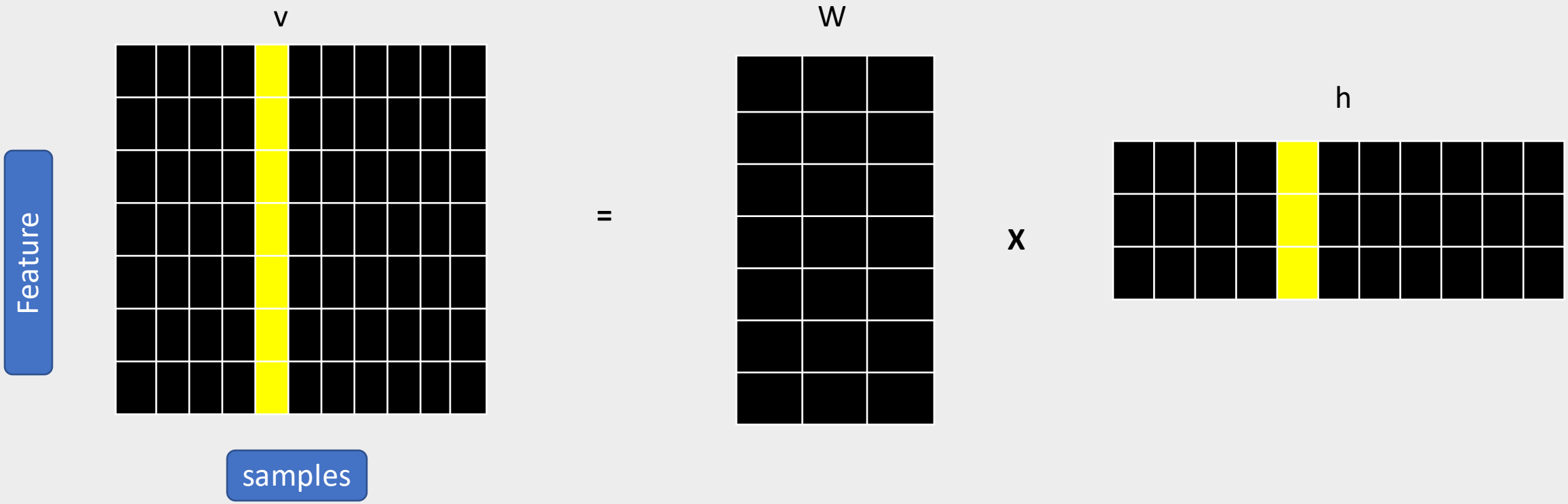
G(h,h′) is an auxiliary function for F (h) if the conditions

G(h,h′) ≥ F(h) and G(h,h) = F(h) ae satisfied.

Significance of approximation.

Column by column approximation as v=W x h where i.e each vector v (column of V) can be a linear combination of W (basis vector) with weighted by h.

.



Problem statement

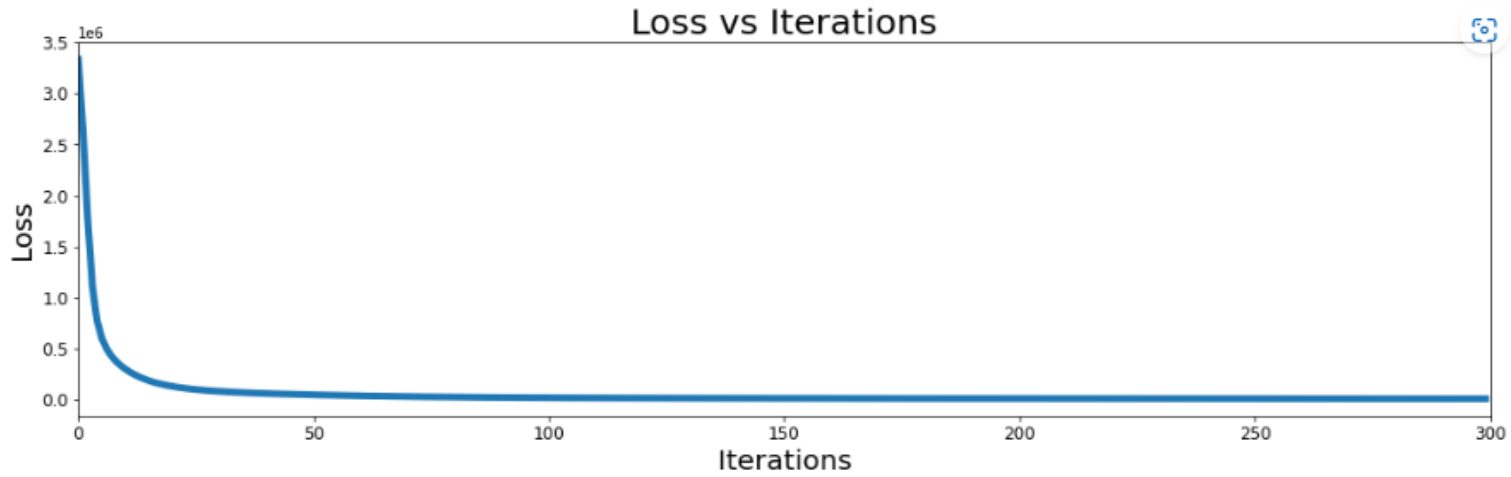
So here we implemented NMF on the image processing for the feature extraction.

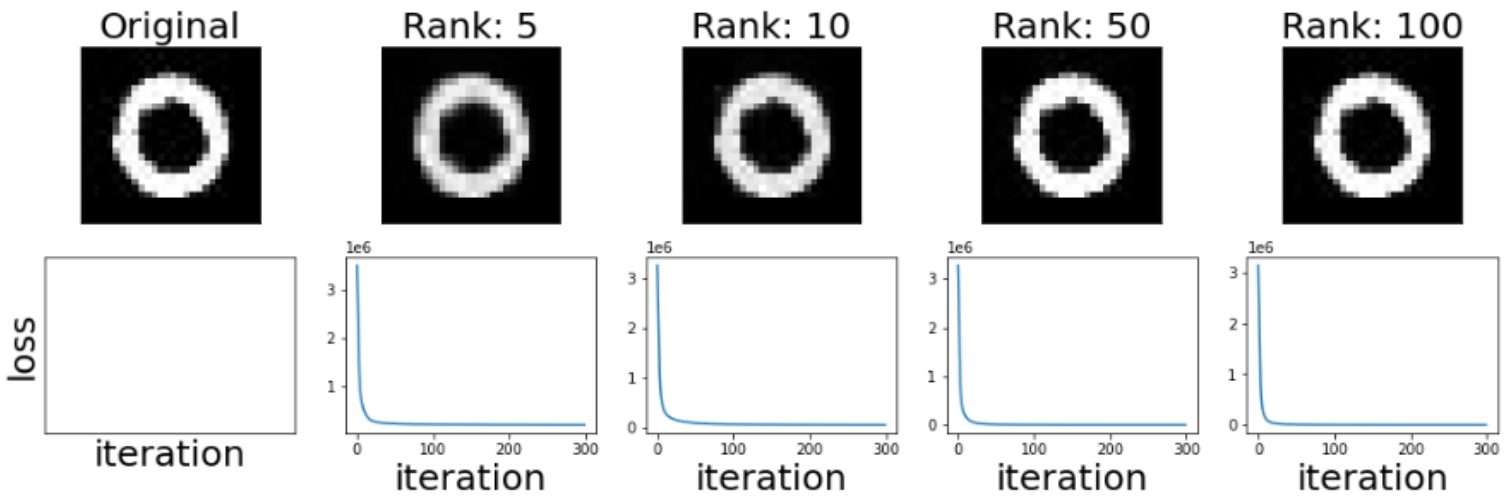
And in this it could be extract meaningful features from a high-dimensional dataset (in the form of Matrix), where the features are non-negative and linearly dependent.

And In the code section we are input an image which have dimension V=28*28.

And after dimensionality reduction image size is W= 28*20 and H =20*28.

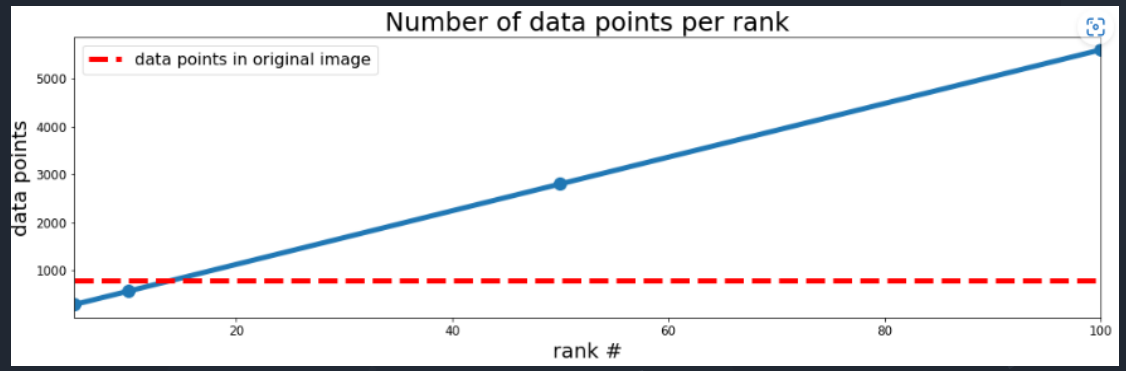
Results

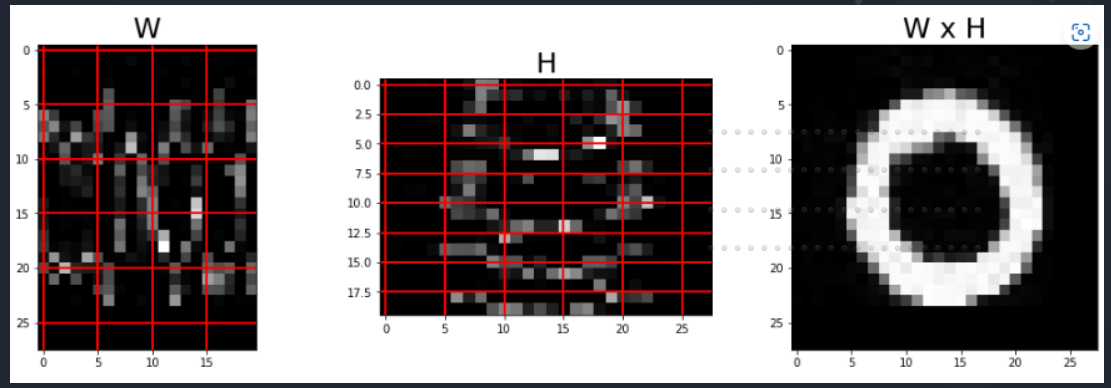

Loss vs Iterations
1e6
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Loss
0
50
100
150
200
250
300
Iterations


Original
Rank: 5
Rank: 10
Rank: 50
Rank: 100
loss
iteration
iteration
iteration
iteration
iteration

Here we are showing the data points of our original image and datapoints of dimensionalized image.

Checking by multiplying W and H we are getting approximately the same image as original one.





Some real life application

- Astronomy, computer vision, document clustering, missing data imputation, chemometrics, audio signal processing, recommender systems, and bioinformatics.

Comaparision

Vector Quantisation:

- Compressing dataset by mapping high-dimensional data points to a set of discrete codes.

Principle Component Analysis:

- It reduce the dimensionality of a data set by finding the principal components, or directions of greatest variance, and projecting the data onto those components. The resulting reduced-dimensionality data set captures most of the variance in the original data set.

Non Negative Matrix Factorisation:

- It is similar to PCA in that it is used to find a low-dimensional representation of a high-dimensional data set.
- Both requires input data and the resulting factors to be non-negative.

Comment

- NMF is often the preferred technique when working with non-negative data, particularly in image and audio processing applications, due to its superior accuracy and interpretability compared to VQ and PCA.

Thankyou