



WWW.DAUPHINE.PSL.EU

09/10/2024

Recommendation systems

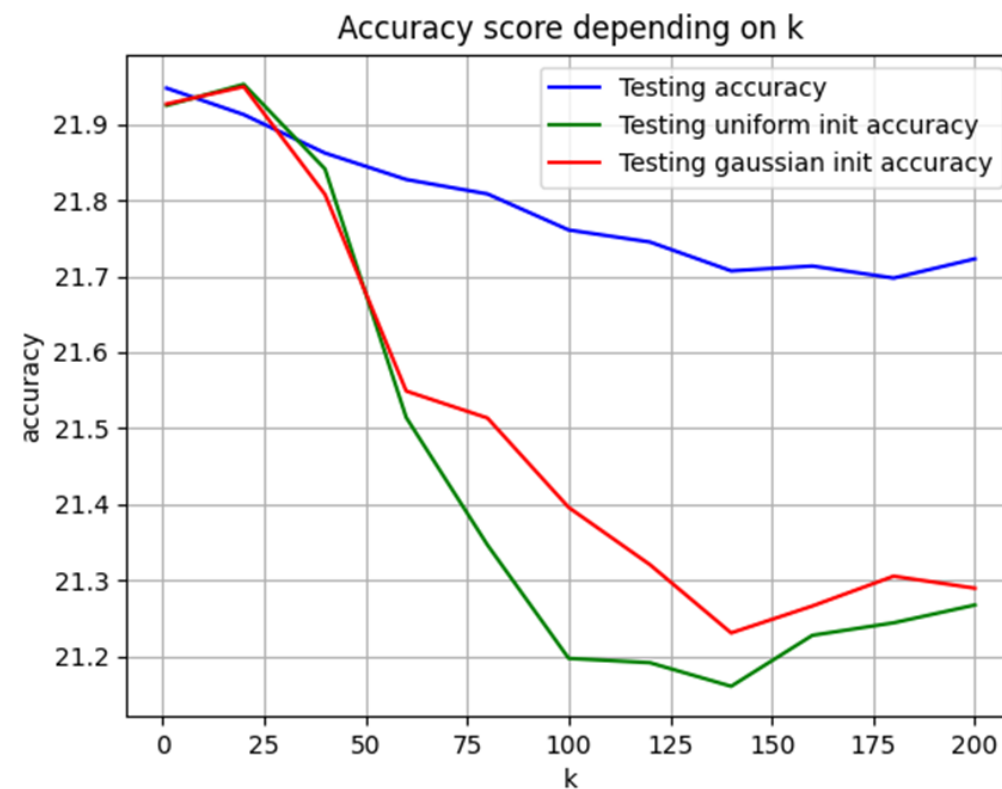
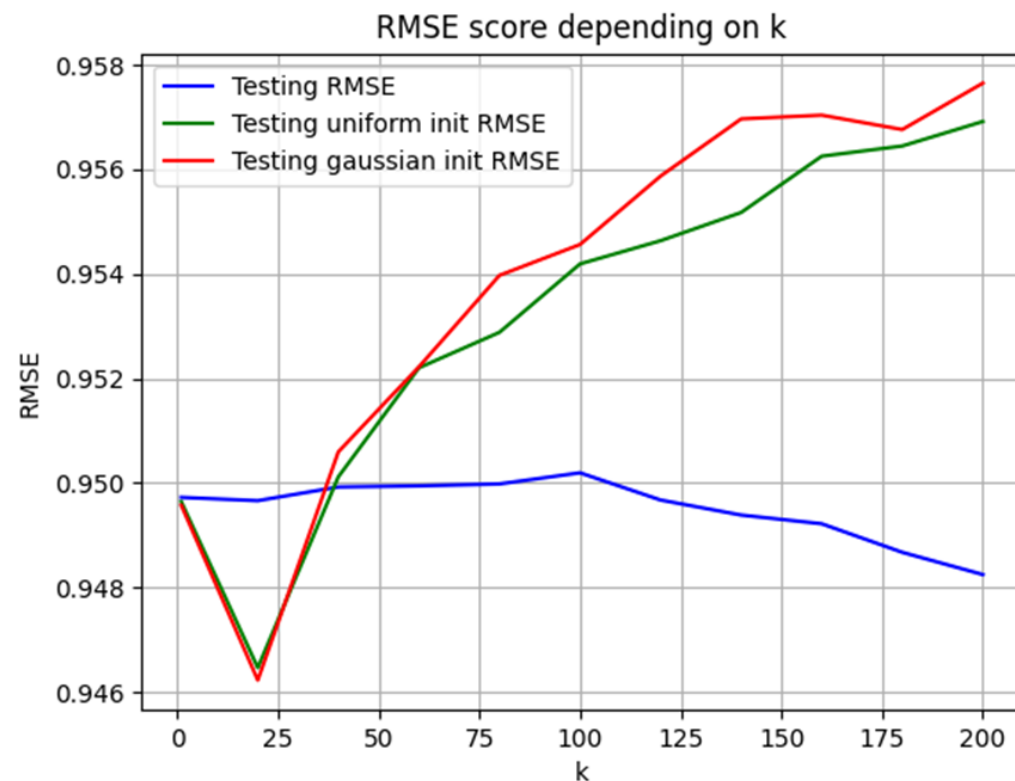
Group : RiMaYa
Rithy SOCHET
Marc KASPAR
Yannis FONTAINE

Matrix factorisation

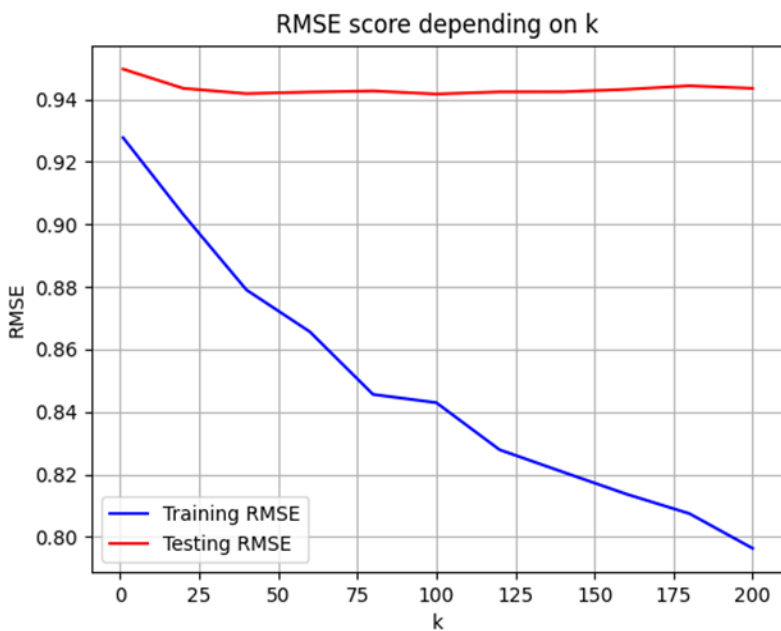
- **Cost function:** $\| R - IU^T \|_F^2 + \lambda \| I \|_F^2 + \mu \| U \|_F^2$
- **Optimiser:** Alternated Least-Square (ALS)

$$I_{t+1} = RU_t(U_t^T U_t + \lambda I)^{-1}$$
$$U_{t+1} = R^T I_t(I_t^T I_t + \mu I)^{-1}$$

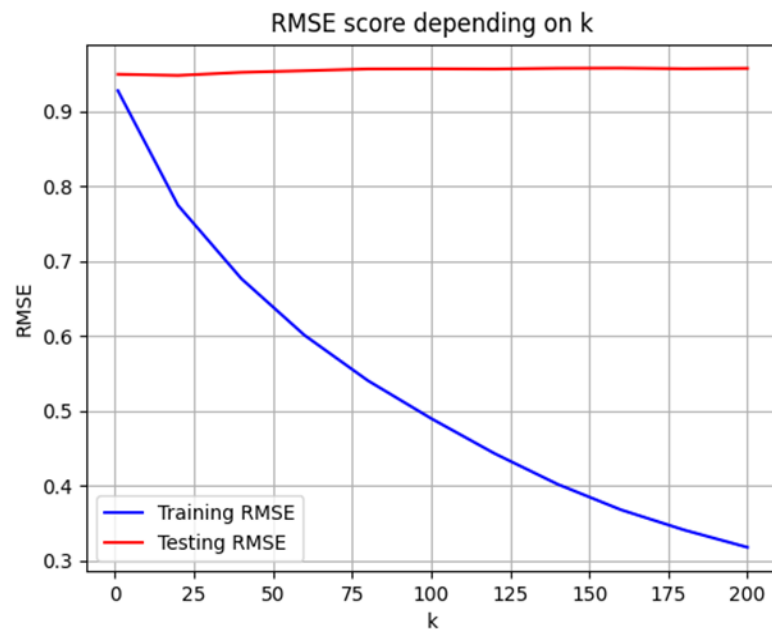
Some results: different initialisation



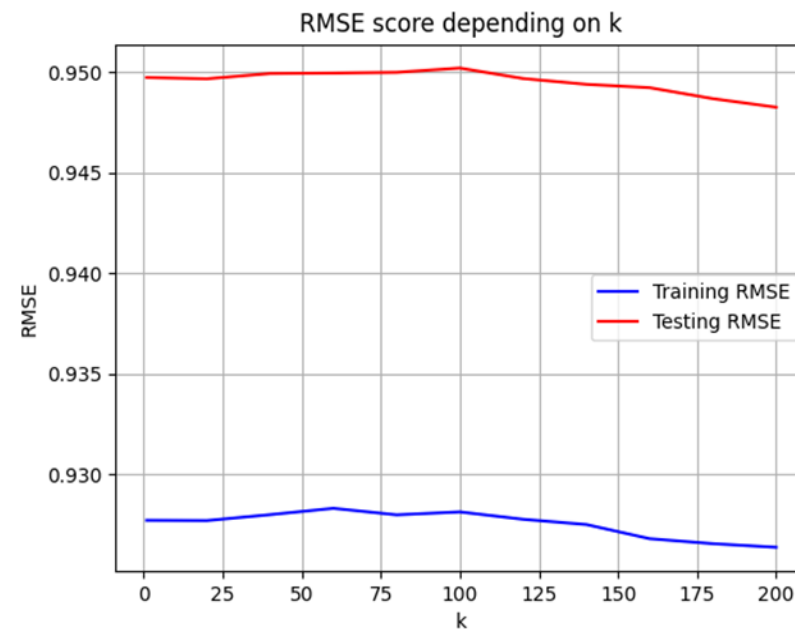
Some results: regularisation



$$\lambda = 0.1, \mu = 1$$



$$\lambda = \mu = 0.1$$

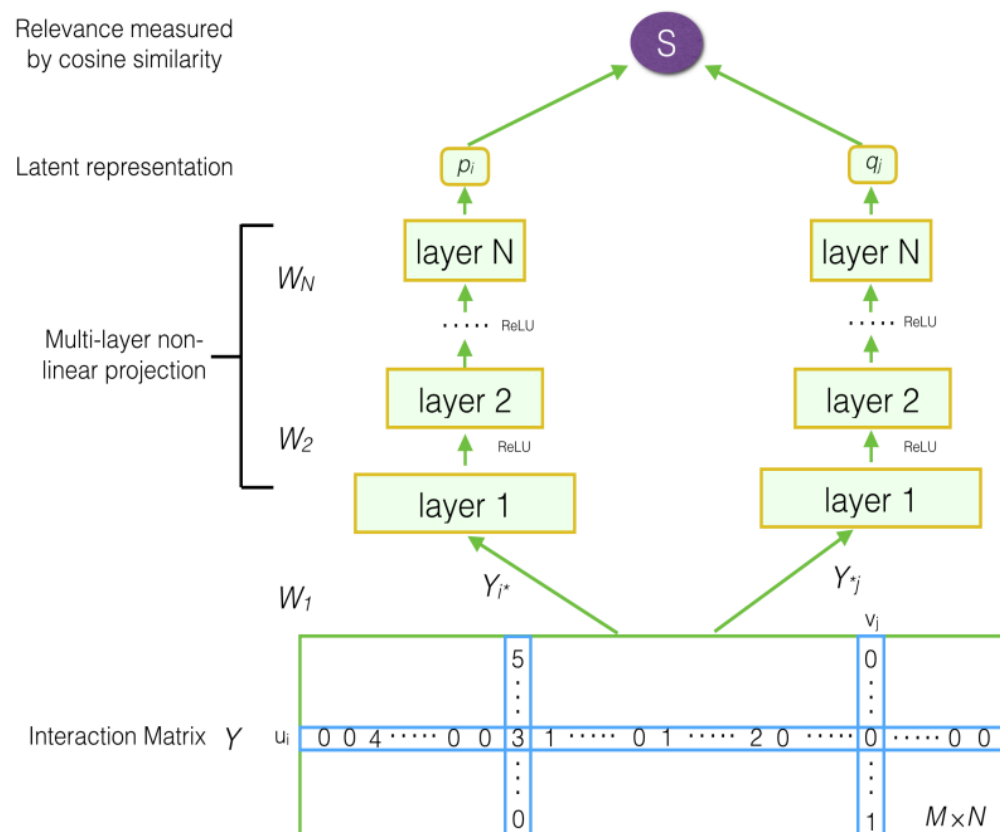


$$\lambda = 1, \mu = 0.1$$

Deep matrix factorisation

- Advantages:
 - DMF can model complex non-linear interactions between users and items.
 - Fit very well on large datasets
- Disadvantages:
 - Costly in terms of time and resources
 - Risk of overfitting
 - Large data volumes required

Architecture



$$\hat{Y}_{ij} = F^{DMF}(u_i, v_j | \Theta) = \text{cosine}(p_i, q_j) * 5 = \frac{p_i^T q_j}{\|p_i\| \|q_j\|} * 5$$

Model enhancement

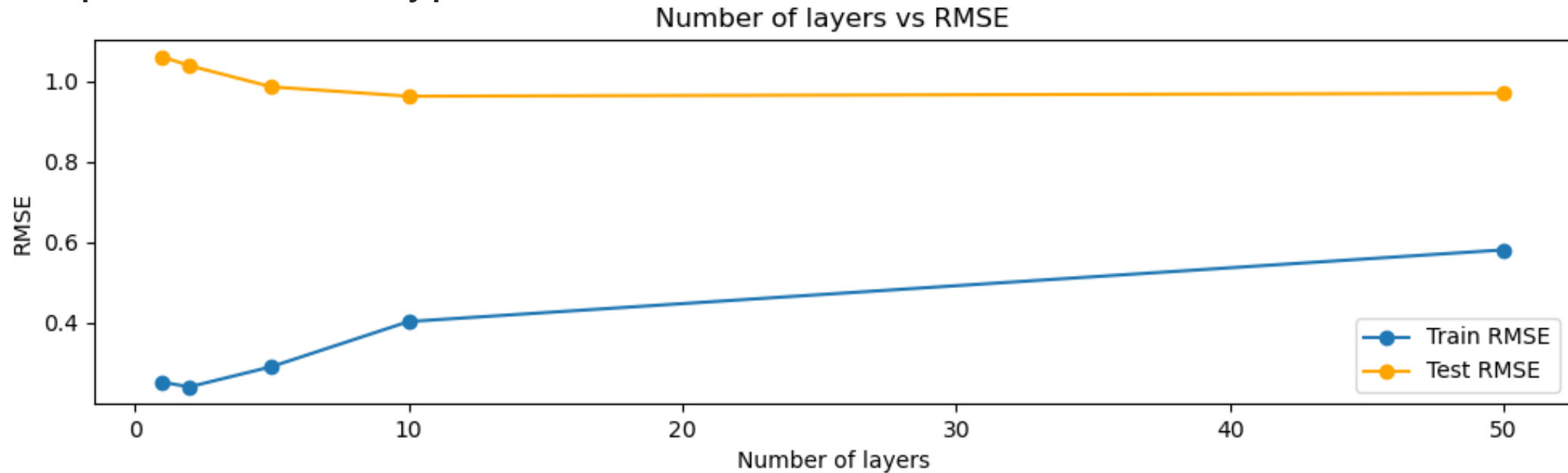
We added the film genre in our model

- Preprocessing of genre:

We tokenize our genders datas so that each gender is represented by a number.

Hyperparameters optimisation

After optimisation of hypermaters:



We obtain similar results :

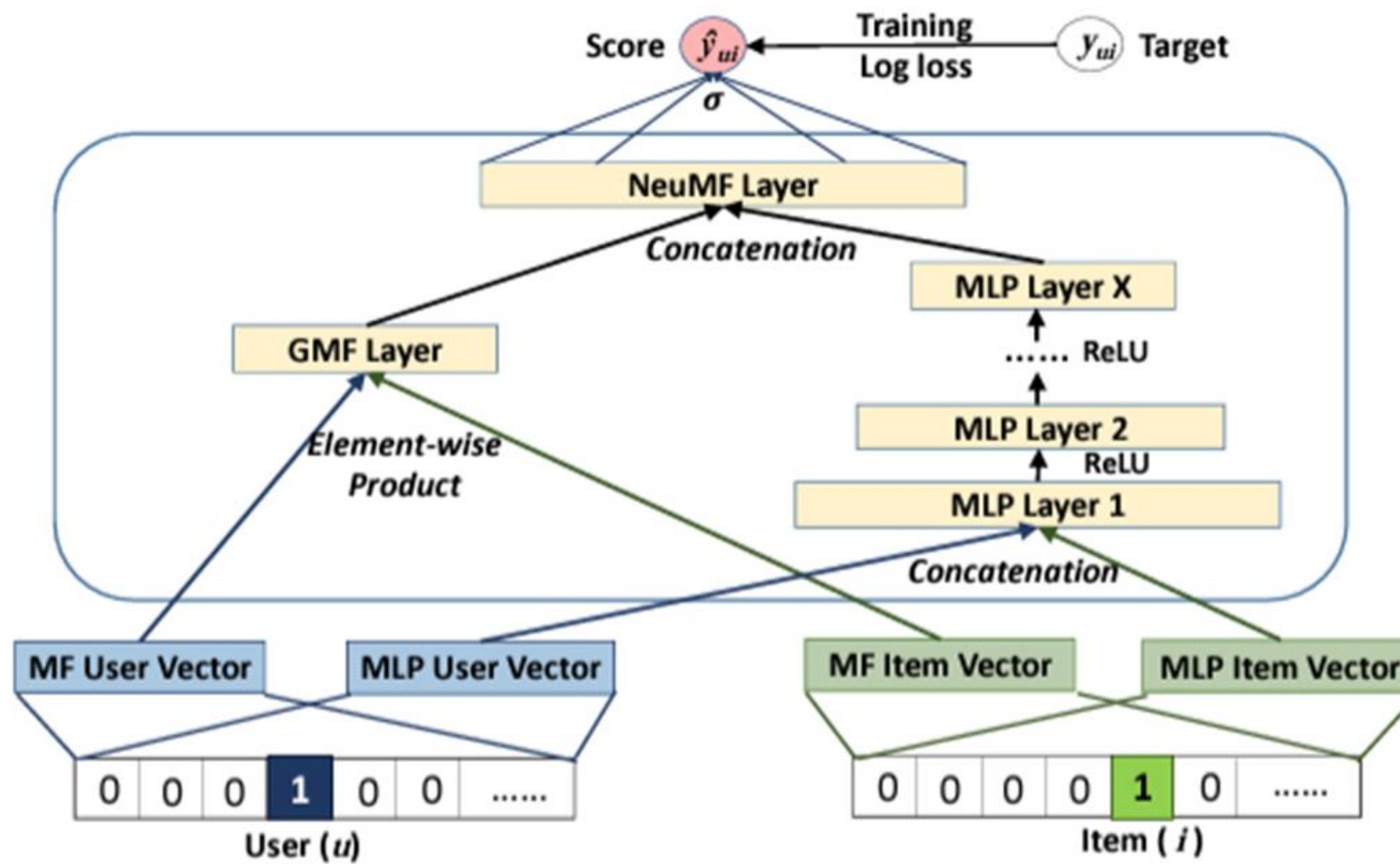
RMSE-test (DMF) = 1.053

RMSE-test (DMF with rouding and genre datas) = 1.045

Neural Collaborative Filtering

- Advantages:
 - It captures more complex non-linear relationships between users and items
 - Generalization of Matrix Factorization
- Disadvantages:
 - Risk of overfitting

Architecture



Some results

Parameters:

Optimizer: Stochastic Gradient Descent
with learning rate = 0.001

Batch size: 128

Number of epochs: 30

MLP structure: [64,32, 16, 8]

Test results:

RMSE = 1.0379

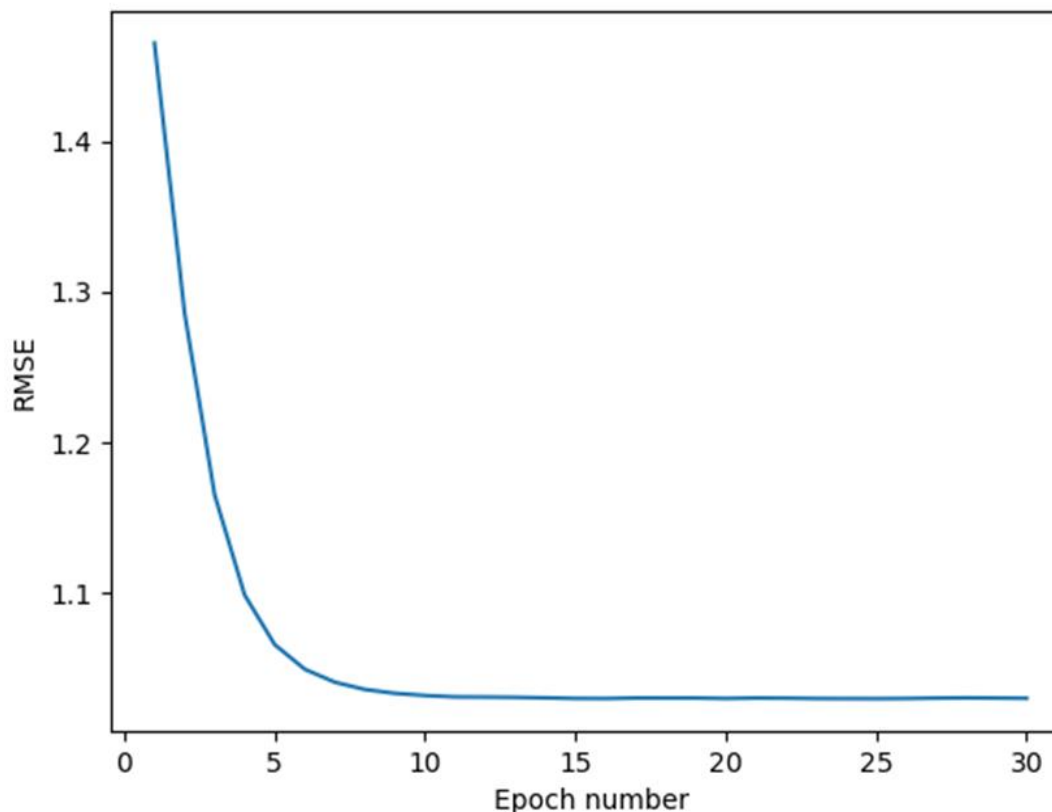
Accuracy = 0

Time = 17.22s

With Rounding:

RMSE= 1.1473

Time = 17.38s



Thank you for listening



UNIVERSITÉ PARIS DAUPHINE - PSL

Place du Maréchal de Lattre de Tassigny – 75775 Paris cedex 16