

Machine Learning : Projet - Réseau de neurones : DIY

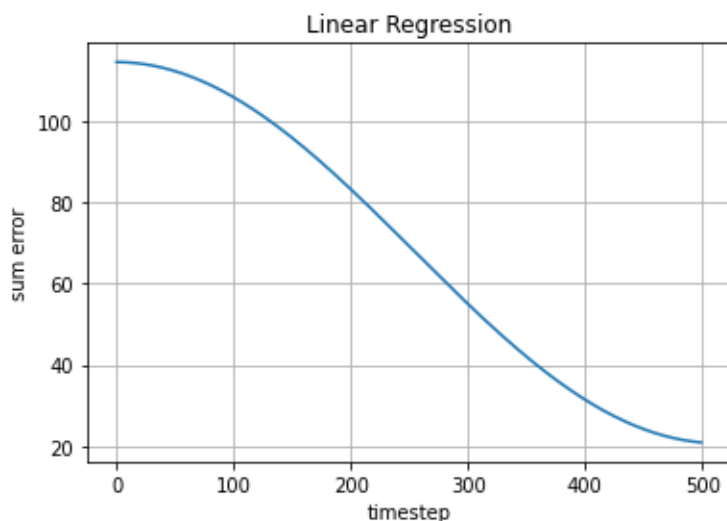
Semestre 2 - Données, Apprentissages, Connaissances

Vincent Fu - 3681685

Yuhao Liu - 3703494

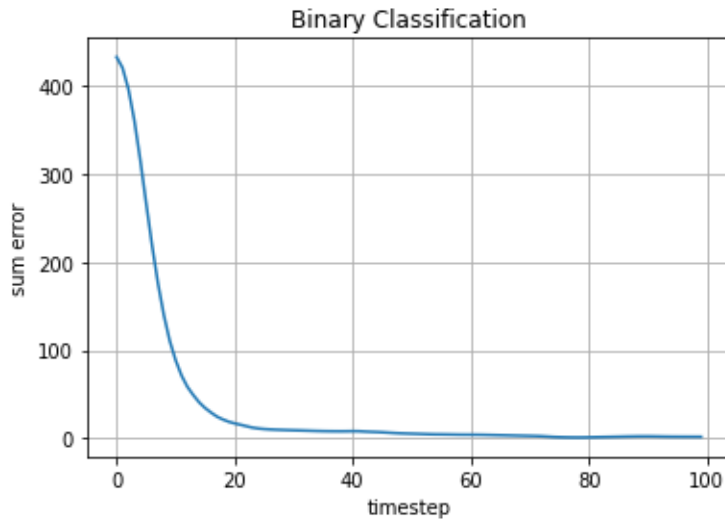
1. Mon premier est ... linéaire !

Voici la courbe de l'erreur en fonction du pas de temps d'apprentissage en prenant comme pas de gradient $\alpha = 1.5 \times 10^{-7}$ pour la régression linéaire :

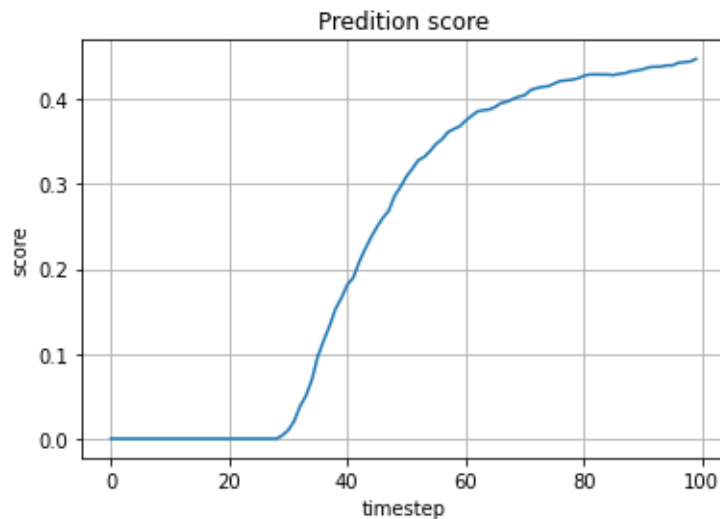


2. Mon deuxième est ... non linéaire !

Voici la courbe de l'erreur en fonction du pas de temps d'apprentissage en prenant comme pas de gradient $\alpha = 1.0 \times 10^{-4}$ pour la classification binaire :



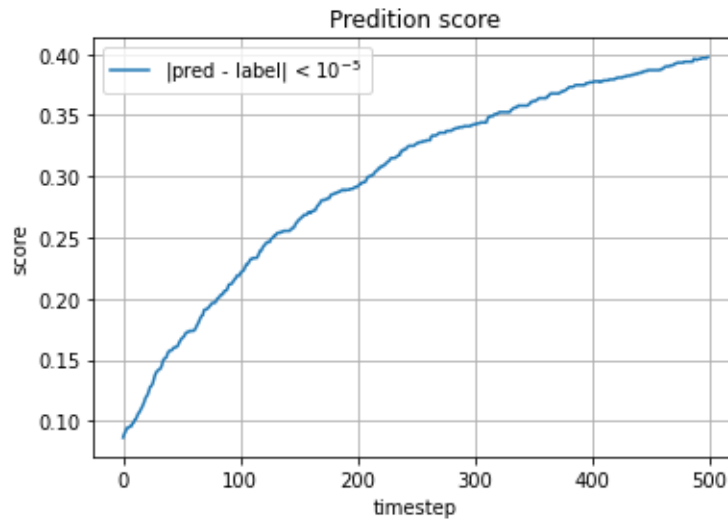
Voici la courbe de score de prédiction :



Malgré que l'erreur converge vers 0, on remarque que le score de prédiction n'est seulement qu'à 45%. Cela est dû au fait que certaines prédictions ne renvoient pas exactement les valeurs des classes comme 0 ou 1 à cause des erreurs d'arrondis. Ils renvoient plutôt des valeurs de l'ordre de 10^{-15} ou moins.

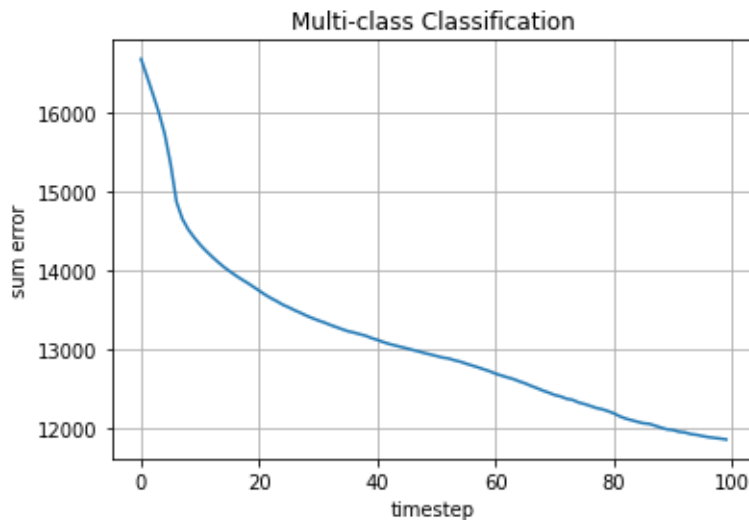
3. Mon troisième est un encapsulage

Voici la courbe de score de prédiction pour le réseau de neurones de la section précédente :

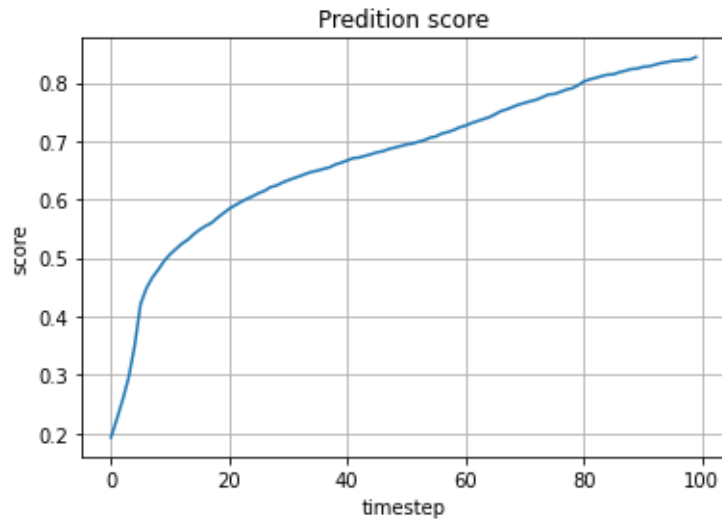


4. Mon quatrième est multi-classe

Voici la courbe de l'erreur en fonction du pas de temps d'apprentissage en prenant comme pas de gradient $\alpha = 1.0 \times 10^{-3}$ pour la classification multi-classe :

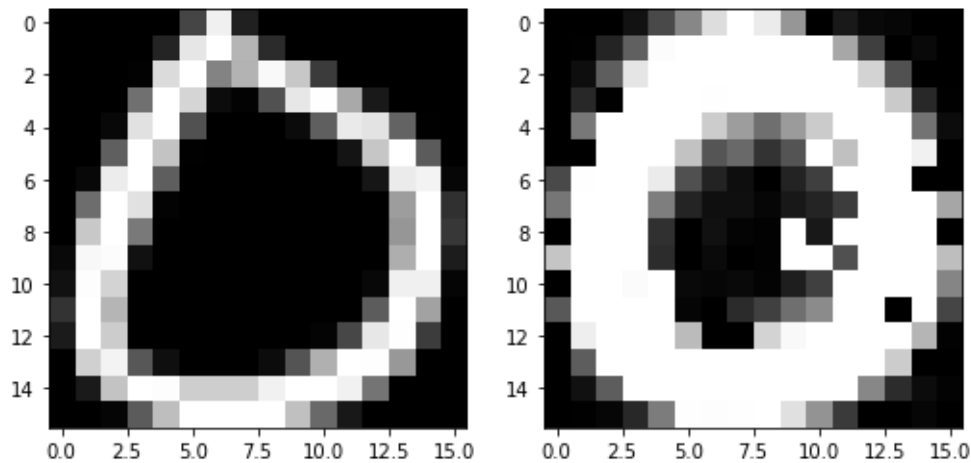


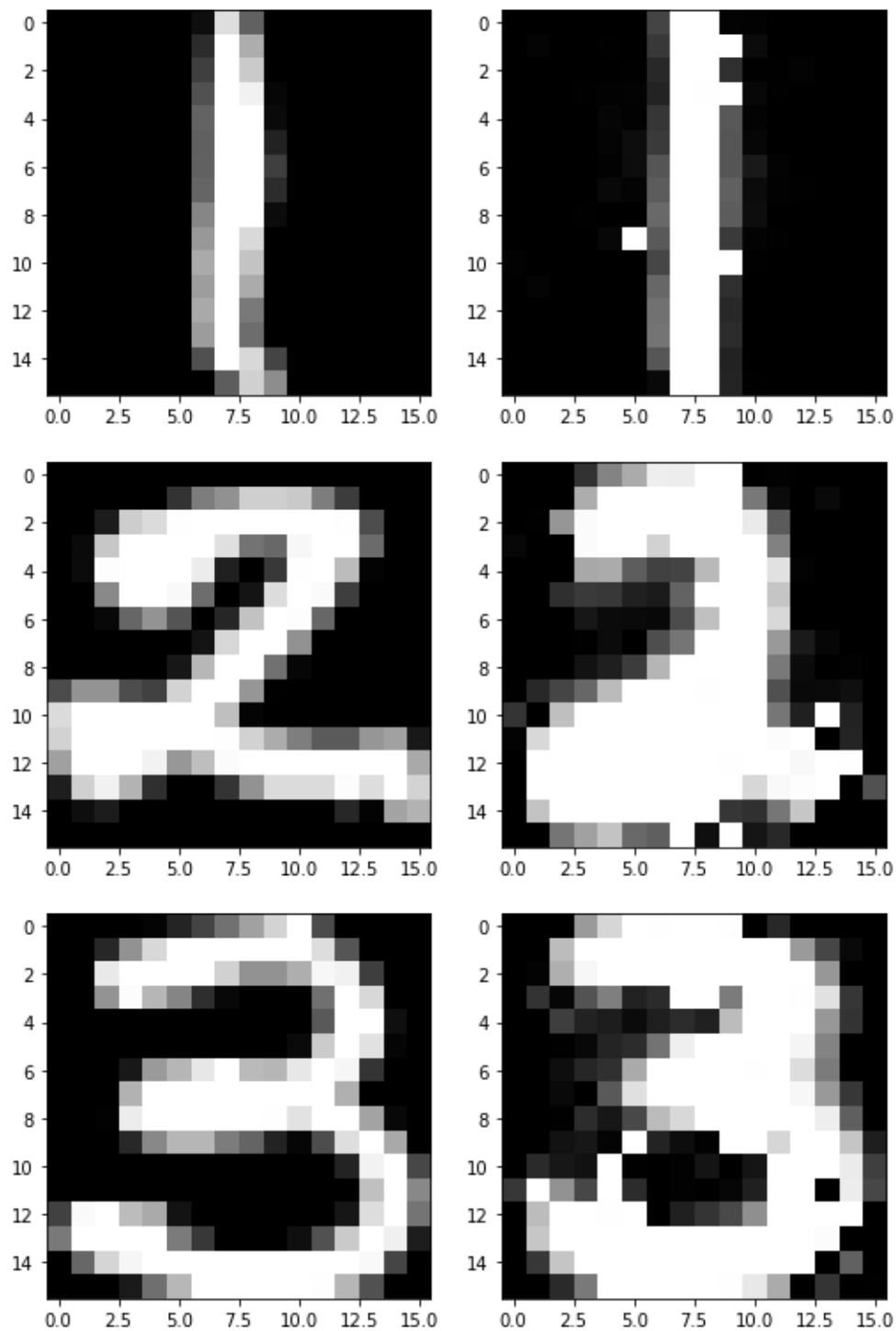
Voici la courbe de score de prédiction :

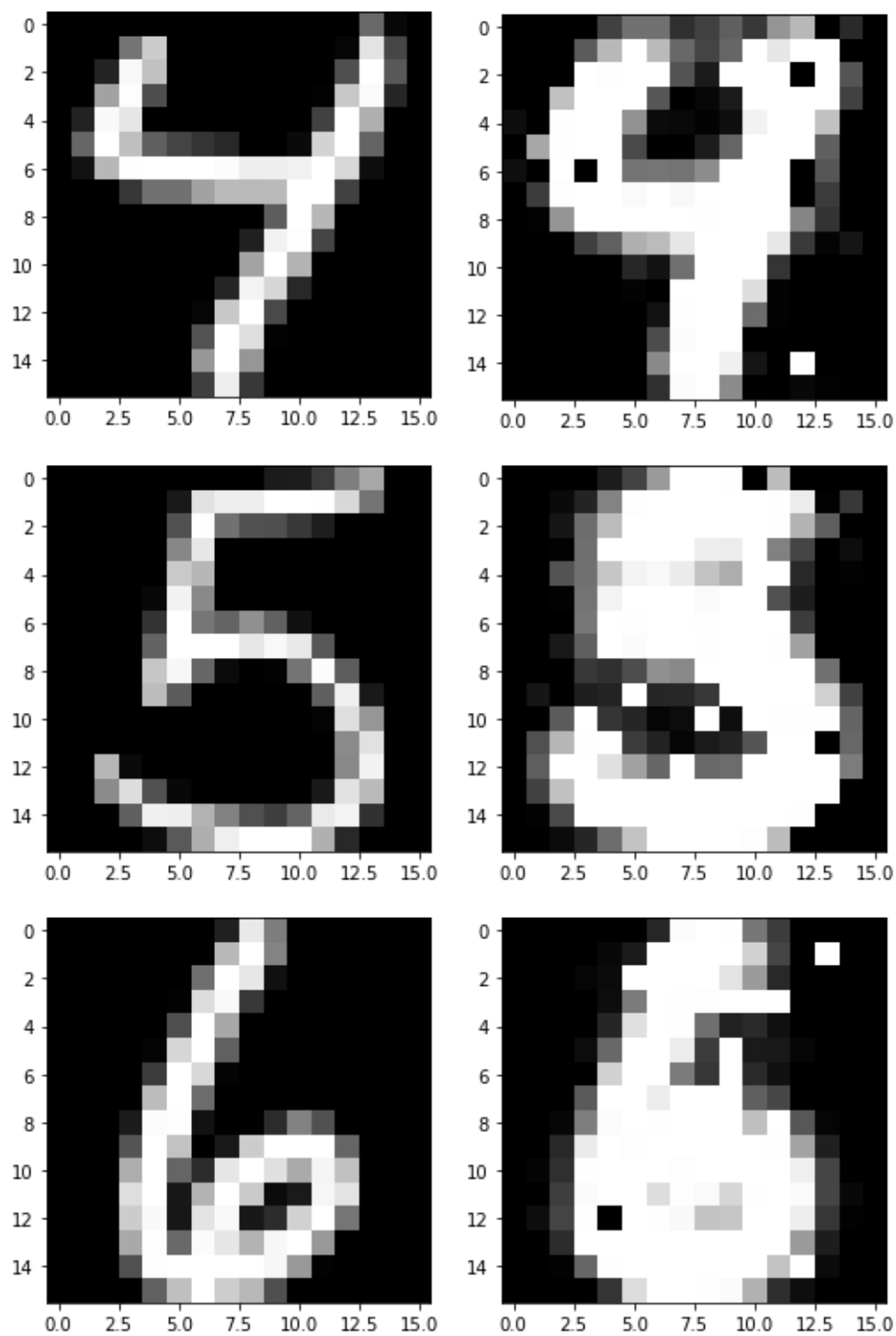


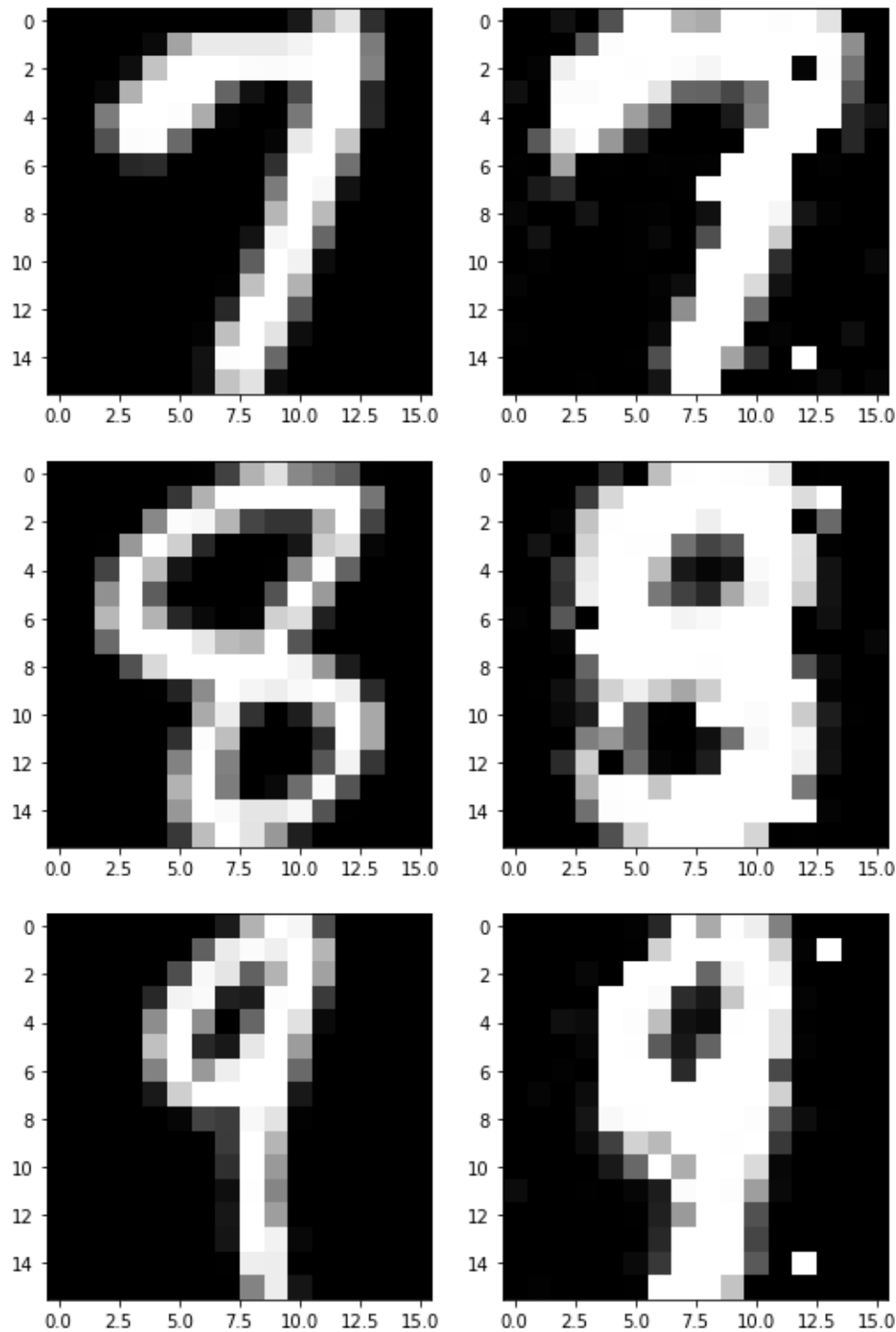
5. Mon cinquième se comprsse (et s'expérimente beaucoup)

Voici les résultats des reconstructions de l'encodeur-décodeur (Linear(256, 100), Tanh(), Linear(100, 10), Tanh(), Linear(10, 100), Tanh(), Linear(100, 256), Sigmoide()) avec un MSELoss, $\alpha = 1e-4$ et 1000 pas d'apprentissages (les originales à gauche, et reconstruites à droite) :





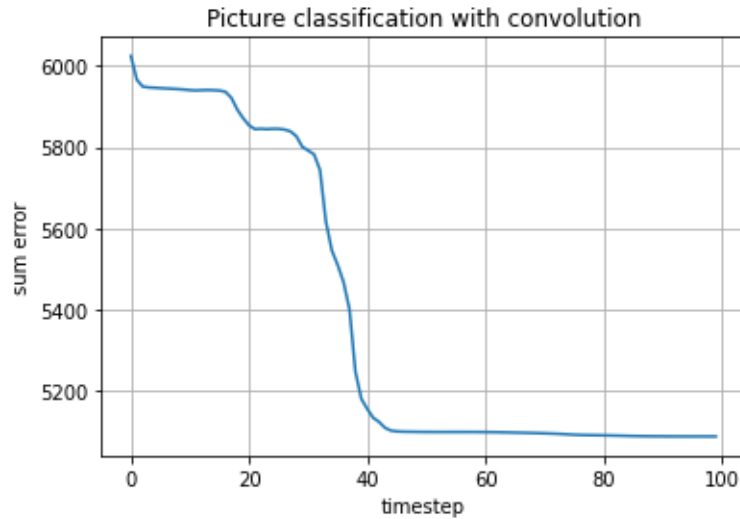




En observant les images, on remarque certaines images reconstruites ne correspondent pas aux données initiales, comme par exemple le 4 et 5 qu'on peut confondre respectivement avec un 9 et un 3. Cela est dû un manque de données de ces classes là par rapport aux classes.

6. Mon sixième se convole

Voici la courbe de l'erreur en fonction du pas de temps d'apprentissage en prenant comme pas de gradient $\alpha = 1.0 \times 10^{-2}$ pour la classification avec convolution avec un Softmax Cross Entropy Loss sur 100 pas d'apprentissage :



Voici la courbe de score de prédiction :

