

**INF8225**

**Hiver 2020**

**TP No. 1**

**1878557 – Bernard Meunier**

**Soumis à :** Christopher Pal

**Question 1**

**A)** P(H=1)

=

Avec la loi des sommes :

= P(H = 1, P = 1, A = 1) + P(H = 1, P = 1, A = 0) + P(H = 1, P = 0, A = 1) + P(H = 1, P = 0, A = 0)

= P(H = 1|P = 1, A = 1)\*P(P = 1) \* P(A = 1) + P(H = 1|P = 1, A = 0) \* P(P = 1)\*P(A = 0) + P(H = 1|P = 0, A = 1) \*

P(P = 0)\*P(A = 1) + P(H = 1|P = 0, A = 0)\*P(P = 0)\*P(A = 0)

= 0.272

**B)** P(H=1|W=1)

=

Avec la loi des sommes et la règle de la chaîne des probabilités, on obtient :

= 0.5956

**C)** P(H=1|W=0)

=

Avec la règle de la chaîne des probabilités et que P(W=0) = 1-P(W=1), on obtient :

=0.0576 / 0.64

=0.09

**D)** P(H=1|P=0,W=1)

= (0.072 \* 0.2 + 1 \* 0) / (0.2\*0.8)

=0.374

**E)** P(W=1H=1)

=

= 0.2144 / 0.272

=0.7882

**F)** P(W=1|H=1,A=1)

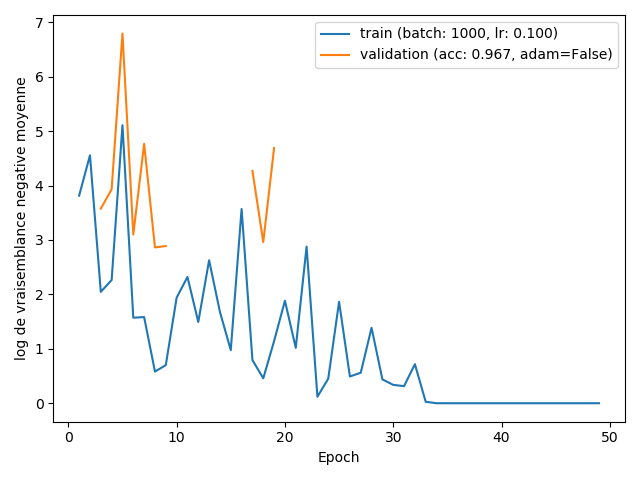
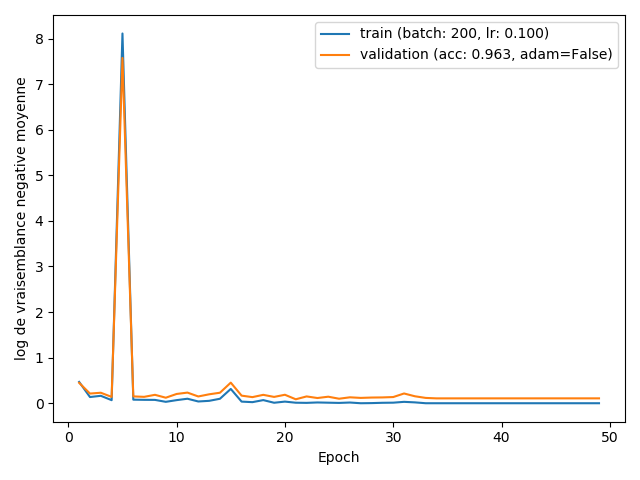
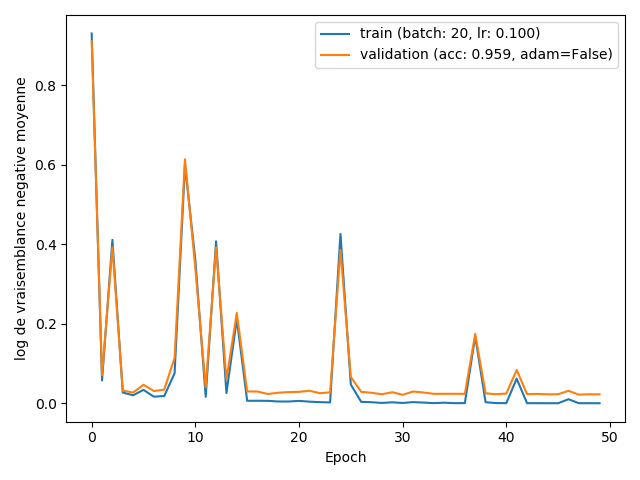
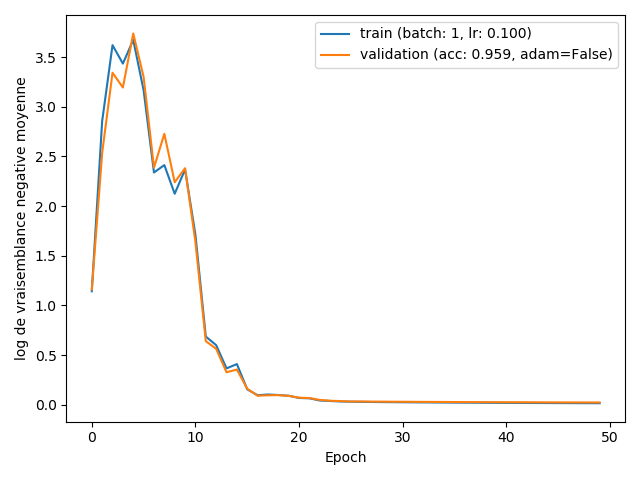
= (0.072\*0.2+1\*0.02) / (0.02+0.072)

=0.374

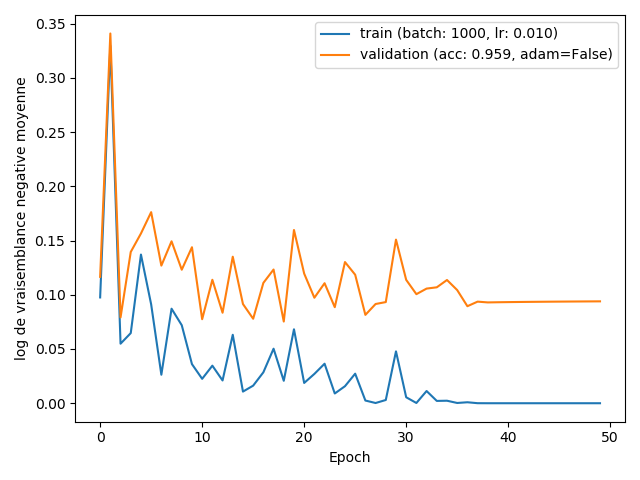
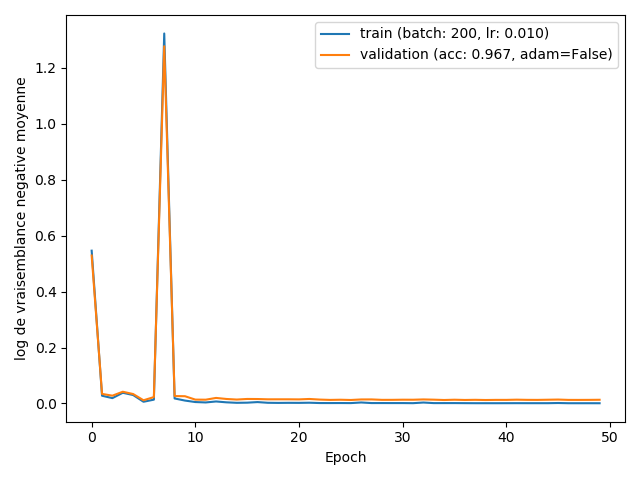
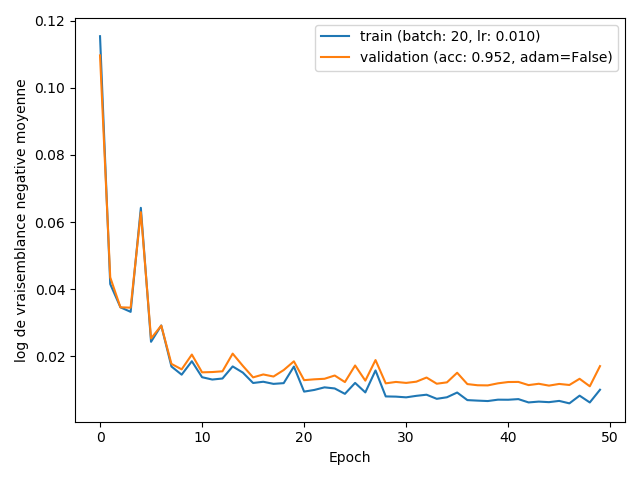
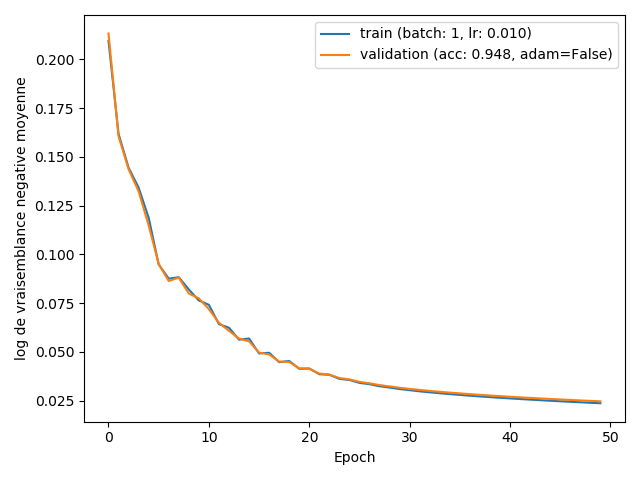
**Question 2**

a)

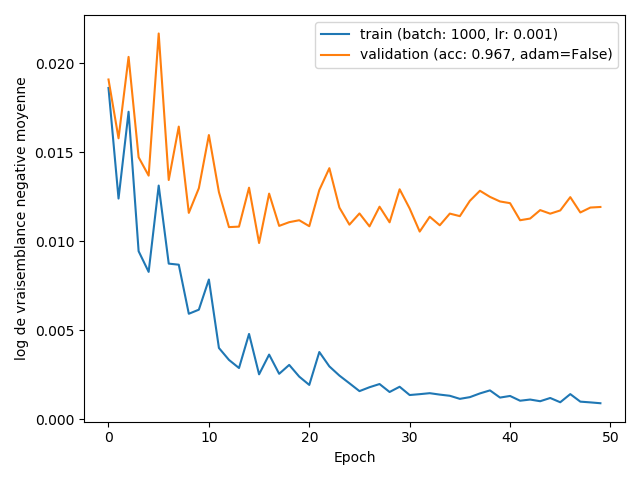
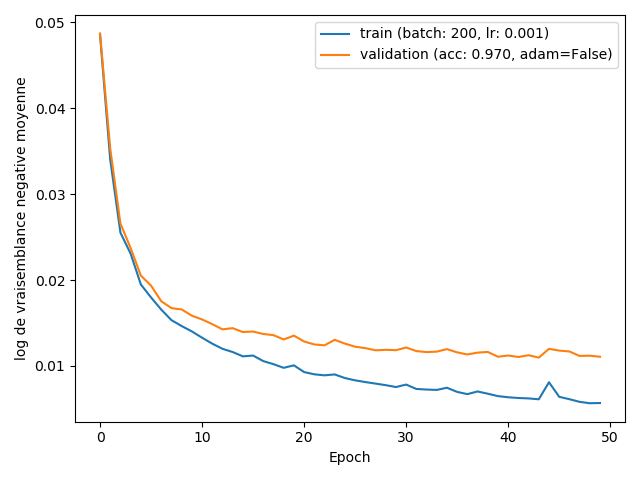
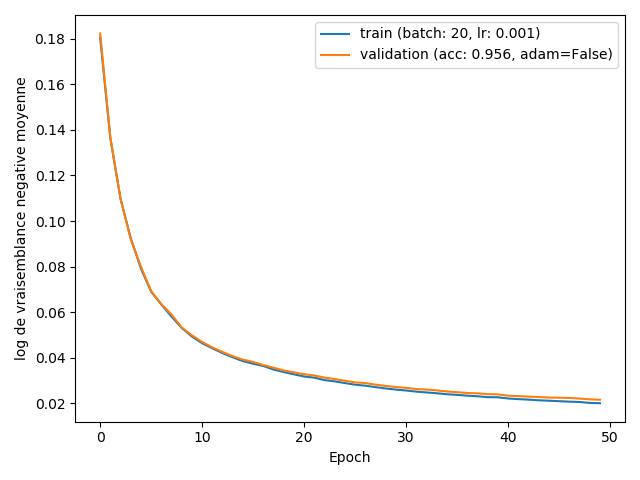
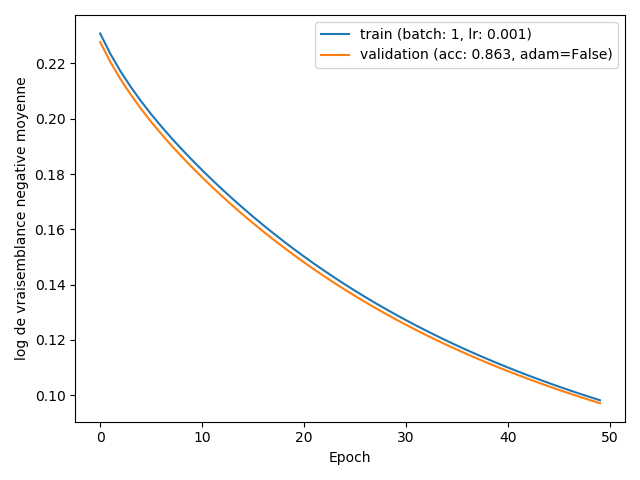
Learning rate = 0.1



Learning rate = 0.01



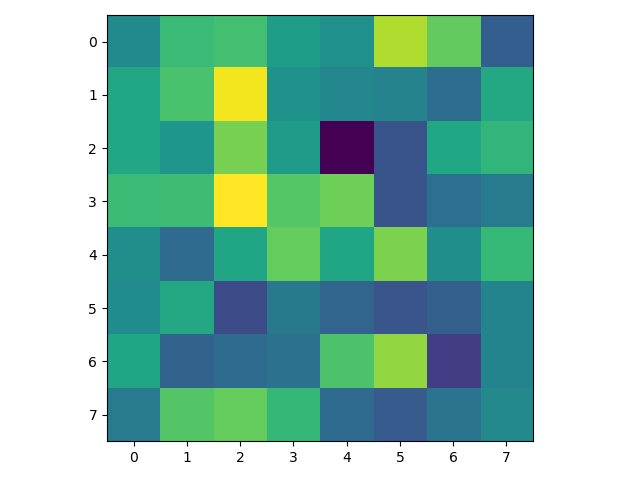
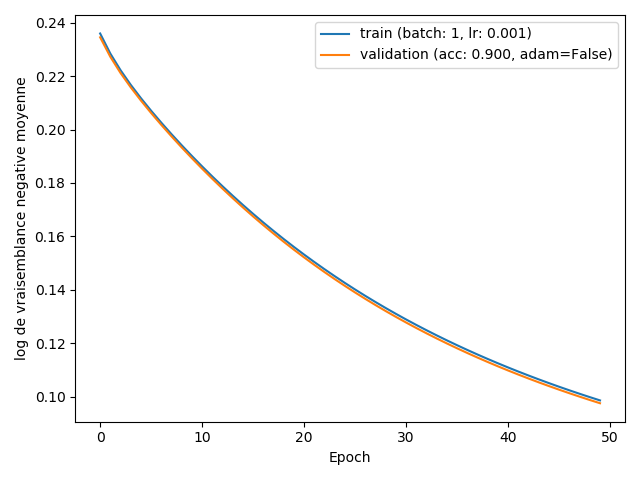
Learning rate = 0.001

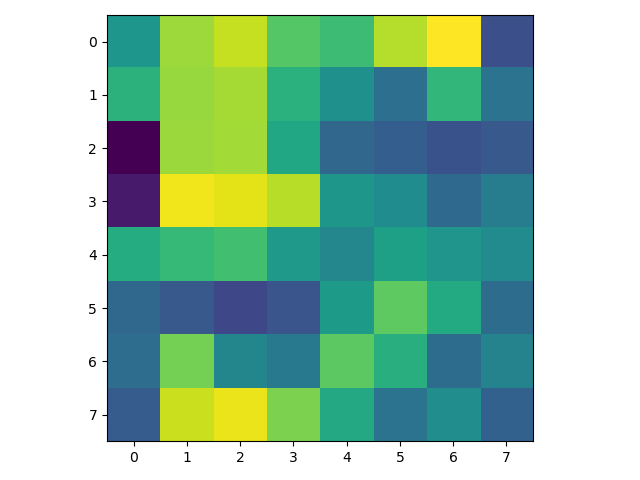
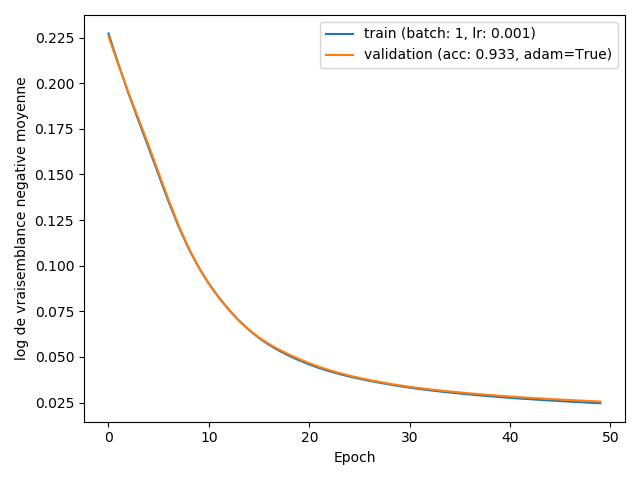


B) ADAM

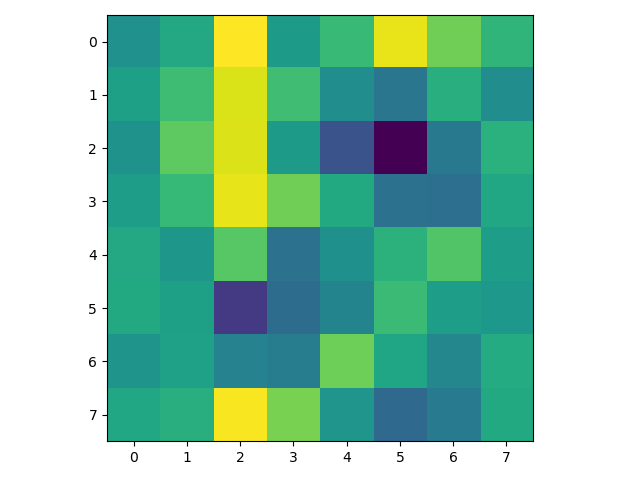
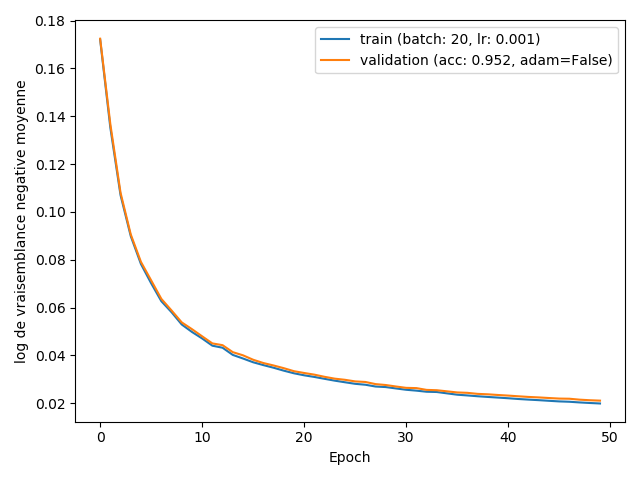
J’ai utilisé un Learning Rate de 0.001 pour mes tests puisque c’est celui-ci qui avait obtenue le meilleur résultat dans les autres tests. Les poids du chiffre 5 sont utilisés dans les images de poids.

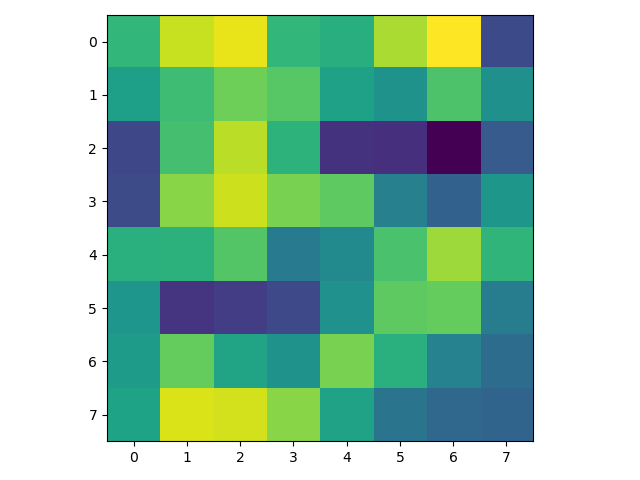
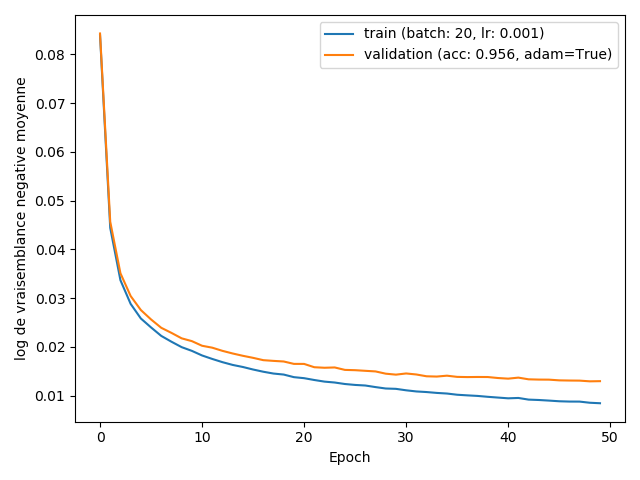
Batch size = 1



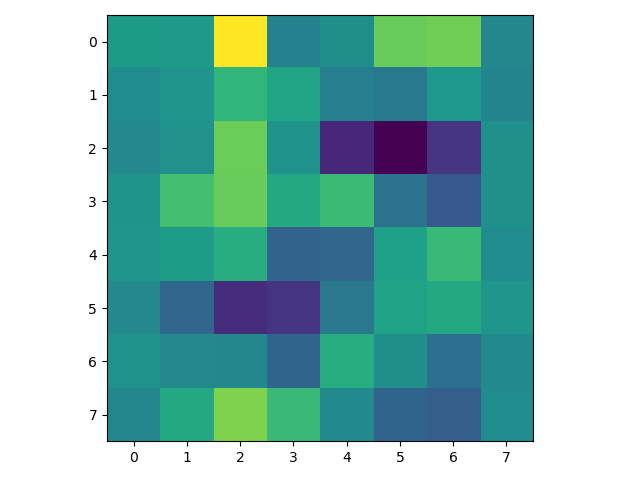
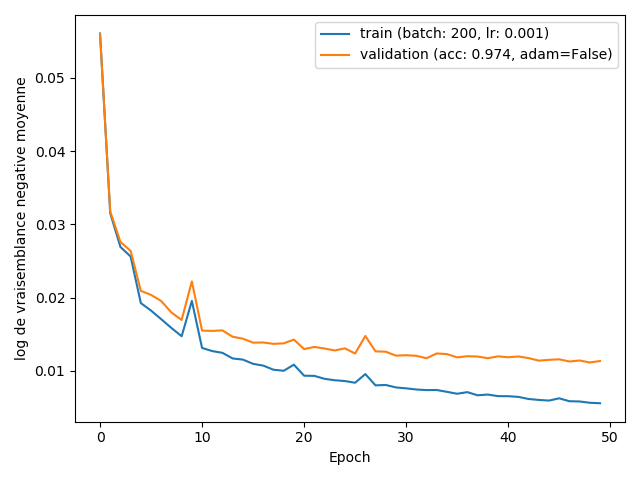


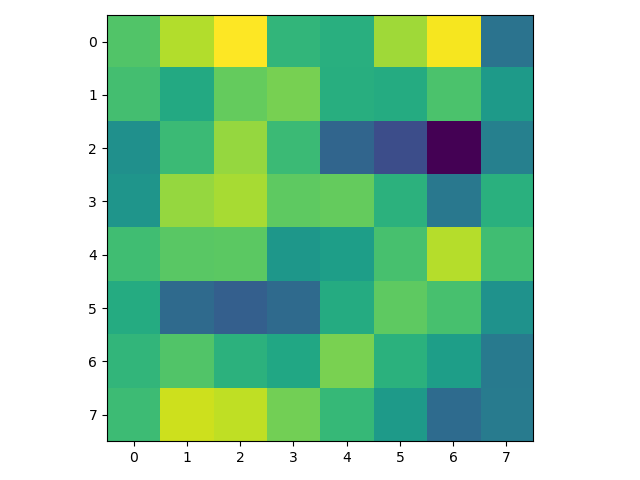
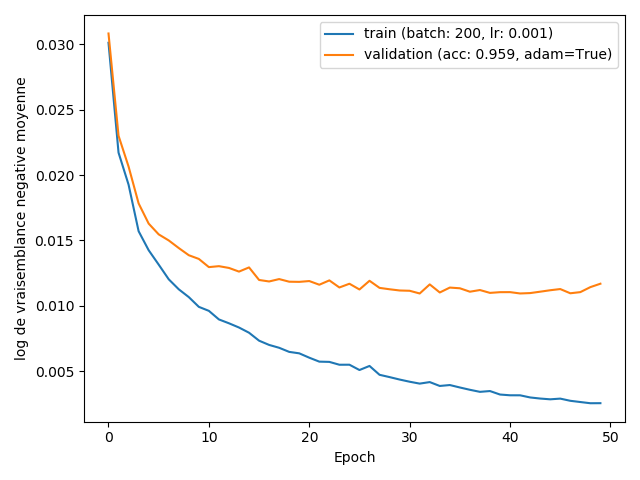
Batch size = 20



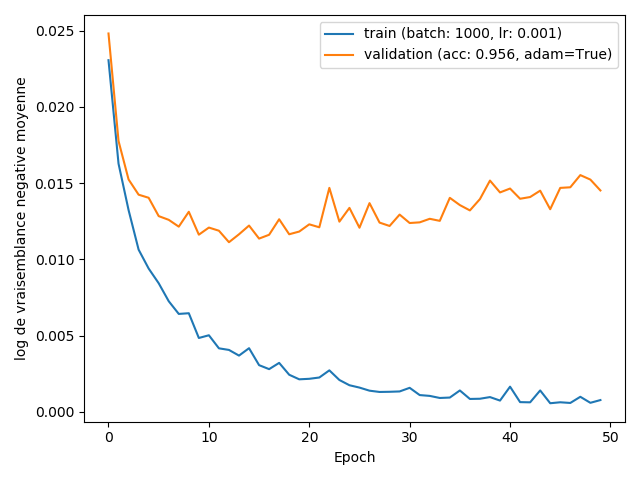
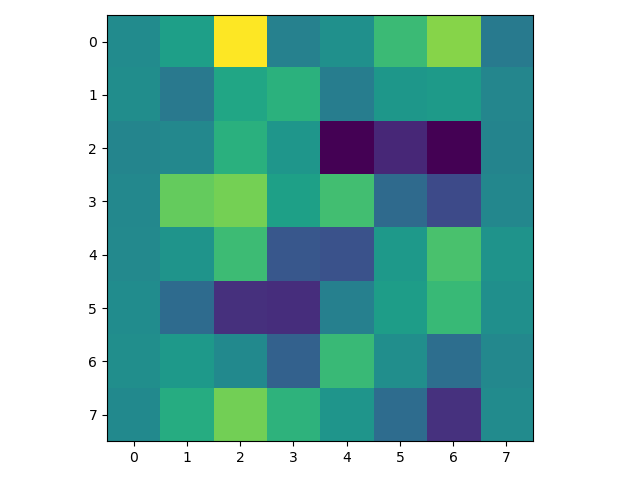
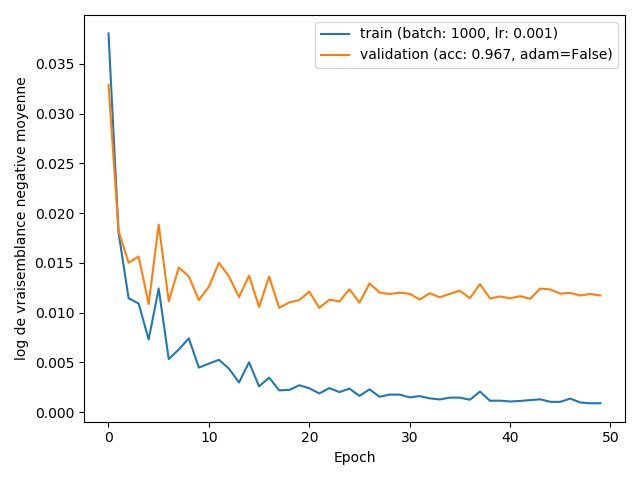
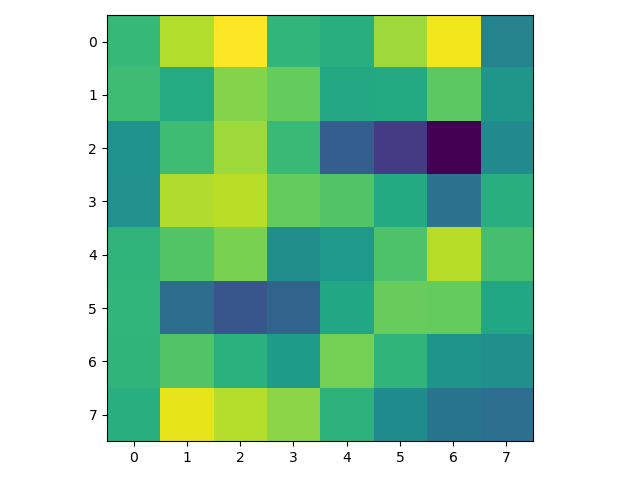


Batch size = 200





Batch size = 1000 :

ADAM ne semble pas faire une grande différence à la précision des prédictions des données sur l’ensemble de test. Dans le cas de la grosseur de batch de 1 et 20, ADAM remporte par une faible marge, alors que pour les batch de 200 et 1000, ADAM semble perdre par une faible marge. Dans tous les cas, la différence de résultat est si faible, qu’elle pourrait simplement venir des éléments aléatoires de l’algorithme.