

Feed Forward Neural Network with Back Propagation

Application on MNIST Dataset

Eden BELOUADAH

6 décembre 2017

1 Introduction

Le domaine d'apprentissage profond est l'un des domaines les plus récents qui s'inspirent du domaine biomédical. Il trouve des applications dans plein de problématiques réelles parmi lesquelles la classification d'image.

Le but de ce travail pratique est de mettre en oeuvre un réseau de neurones Feed Forward avec la rétropropagation de l'erreur, et l'appliquer sur la base de données MNIST.

La base de données MNIST est contruite de 50000 exemple d'images binaires numérisées. Chaque image contient un chiffre manuscrit et le but du réseau est d'associer l'image au bon chiffre correspondant parmi 10.

2 Meilleure architecture du réseau de neurones

La meilleure architecture obtenue pour ce réseau de neurone est la suivante :

- Nombre de couches cachées : 1,
- Nombre de neurones dans chaque couche cachée : 20,
- Fonction d'activation : ReLU,
- Erreur : Quadratique.

3 Analyse des résultats

Nous remarquons que lorsque les dérivées sont très petites, les deux fonctions *sigmoid* et *tanh* ne deviennent plus efficaces et l'effet du gradient disparaît. Donc la fonction ReLU reste le meilleur choix pour ce problème.

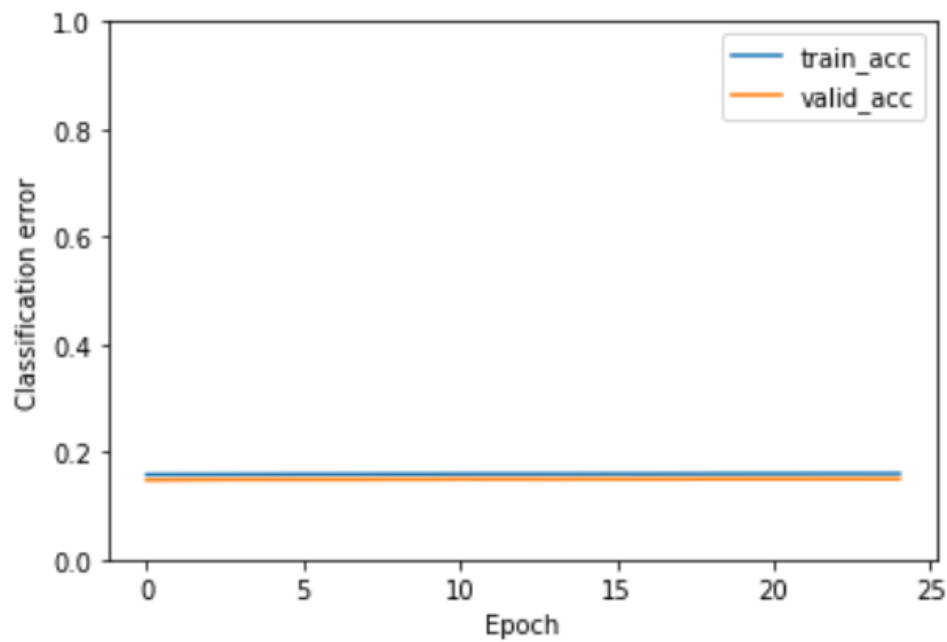


Figure 1. Sample image

Architecture	1	2	3	4	5
Taux de précision	100%	97.5%	92.06%	89.66%	89.19%
Taux d'erreur	100%	97.5%	92.06%	89.66%	89.19%

Tableau 1. Précision et Erreur en fonction d'architecture

- Architecture1 ;
- Architecture2 ;
- Architecture3 ;
- Architecture4 ;

4 Conclusion

Ce travail pratique était une occasion pour voir de plus proche le fonctionnement des réseaux de neurones. Ces derniers peuvent changer d'architecture en variant le type du problème, sa difficulté, ses particularités...

La base de données MNIST représente une tâche de classification simple qui peut être résolue avec un réseau de neurone simple, même sans couches cachées, comme nous l'avons vu dans le travail pratique précédent. En effet, il était possible d'atteindre une précision de 87% facilement et en peu de temps.

Les architectures complexes sont plus adaptées aux problèmes complexes qui nécessitent vraiment un réseau profond.