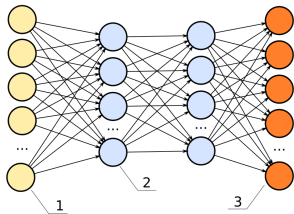


Deep Learning par la Pratique

Optimisation des hyperparamètres

Optimisation des hyperparamètres



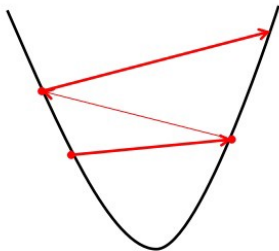
Les hyperparamètres :

- Le learning rate
- La taille de batch
- Le nombre de couches
- La taille des couches

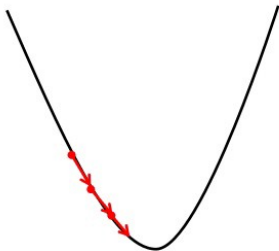
Optimisation du learning rate

Learning Rate

trop grand



trop petit



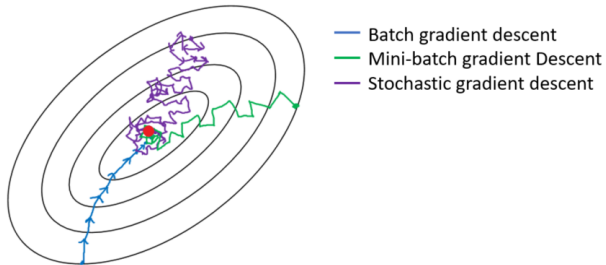
Optimisation du learning rate

Utilisation d'une échelle logarithmique (dans un premier temps) :

```
1 lr_grid = [ 0.1 0.001 0.0001 0.00001]
2 for lr in lr_grid:
3     train = tf.train.GradientDescentOptimizer(lr).minimize(loss)
4     ...
```

Optimisation de la taille du batch

Usuellement, des puissances de 2, pour optimiser les ressources GPU



Optimisation du nombre de couches

```
1 while (perf > perf_precedente):  
2     perf_precedente = perf  
3     ajouter_une_couche(model)  
4     apprendre(model, data_train)  
5     perf = calcul_perf(model, data_validation)
```

Optimisation de la taille des couches

Usuellement, des puissances de 2, pour optimiser les ressources GPU.
Même intuition que pour le nombre de couches.