**Sample Code & 作業內容**

請參閱作業範例 :Day76-Opimizer\_example

作業1:以同一模型, 分別驗證 SGD, Adam, Rmsprop 的 accurancy

作業2:以, Adam, 為例, 調整 batch\_size, epoch , 觀察aaurancy, loss 的變化

作業請提交Day76-Optimizer\_HW

檔案 Day76-Optimizer\_進階 額外提供給學員作為參考

[檢視範例](https://ai100-2.cupoy.com/samplecodelist/D76)

**參考資料**

**延伸閱讀**

An overview of gradient descent optimization algorithms

[**https://arxiv.org/pdf/1609.04747.pdf**](https://arxiv.org/pdf/1609.04747.pdf)

在很多機器學習和深度學習的應用中，我們發現用的最多的優化器是Adam，為什麼呢？

下面是TensorFlow中的優化器，**[https://www.tensorflow.org/api\_guides/python/train](https://www.tensorflow.org/api_guides/python/train" \t "_blank)**

在keras中也有SGD，RMSprop，Adagrad，Adadelta，Adam等：**[https://keras.io/optimizers/](https://keras.io/optimizers/" \t "_blank)**

我們可以發現除了常見的梯度下降，還有Adadelta，Adagrad，RMSProp 等幾種優化器，都是什麼呢，又該怎麼選擇呢？**[https://blog.csdn.net/qq\_35860352/article/details/80772142](https://blog.csdn.net/qq_35860352/article/details/80772142" \t "_blank)**

Sebastian Ruder的這篇論文中給出了常用優化器的比較 [**https://arxiv.org/pdf/1609.04747.pdf**](https://link.jianshu.com/?t=https://arxiv.org/pdf/1609.04747.pdf)

**延伸閱讀:優化器是編譯Keras模型所需的兩個參數之一**

**from** keras **import** optimizers

model = Sequential() model.add(Dense(64, kernel\_initializer='uniform', input\_shape=(10,)))

model.add(Activation('softmax’))

sgd = optimizers.SGD(lr=0.01, decay=1e-6, momentum=0.9, nesterov=**True**)

model.compile(loss='mean\_squared\_error', optimizer=sgd)

您可以在傳遞優化器之前將其實例化model.compile()，如上例所示，或者您可以通過其名稱來調用它。

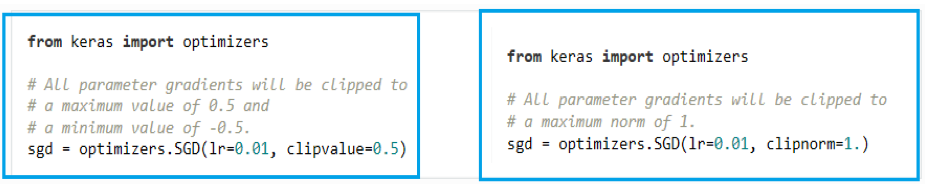
在後一種情況下，將使用優化程序的默認參數。

*# pass optimizer by name: default parameters will be used*

model.compile(loss='mean\_squared\_error', optimizer='sgd')

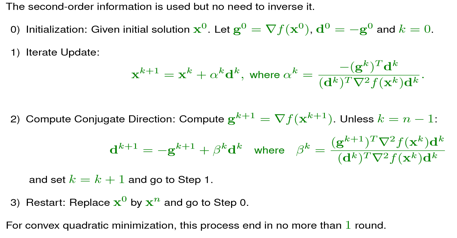
**所有Keras優化器通用的參數**

的參數clipnorm和clipvalue可以與所有優化可以用來控制限幅梯度

**延伸閱讀:二階優化算法**

[**https://web.stanford.edu/class/msande311/lecture13.pdf**](https://web.stanford.edu/class/msande311/lecture13.pdf)

二階優化算法使用了二階導數(也叫做**Hessian方法** )來最小化或最大化損失函數。由於二階導數的計算成本很高，所以這種方法並沒有廣泛使用。



**延伸閱讀:自適應的算法**

如果需要更快的收斂，或者是訓練更深更複雜的神經網絡，需要用一種自適應的算法。

[**http://www.redcedartech.com/pdfs/Select\_Optimization\_Method.pdf**](http://www.redcedartech.com/pdfs/Select_Optimization_Method.pdf)

