# Algoritmo hyperband

Tiempo inicio 11:19 del 21-08-2023

Parada a 11:32

Ejecuto devuelta 11:52

Finalizo a las 18:39

Por un error mio no se activó la gpu de colab antes de comenzar a entrenar

Datos del google cloud

!df -h

Filesystem Size Used Avail Use% Mounted on

overlay 108G 27G 82G 25% /

tmpfs 64M 0 64M 0% /dev

shm 5.8G 0 5.8G 0% /dev/shm

/dev/root 2.0G 1.1G 887M 55% /usr/sbin/docker-init

tmpfs 6.4G 272K 6.4G 1% /var/colab

/dev/sda1 44G 28G 16G 64% /etc/hosts

tmpfs 6.4G 0 6.4G 0% /proc/acpi

tmpfs 6.4G 0 6.4G 0% /proc/scsi

tmpfs 6.4G 0 6.4G 0% /sys/firmware

drive 15G 7.9G 7.2G 53% /content/drive

!cat /proc/cpuinfo

processor : 0

vendor\_id : GenuineIntel

cpu family : 6

model : 79

model name : Intel(R) Xeon(R) CPU @ 2.20GHz

stepping : 0

microcode : 0xffffffff

cpu MHz : 2199.998

cache size : 56320 KB

physical id : 0

siblings : 2

core id : 0

cpu cores : 1

apicid : 0

initial apicid : 0

fpu : yes

fpu\_exception : yes

cpuid level : 13

wp : yes

flags : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse36 clflush mmx fxsr sse sse2 ss ht syscall nx pdpe1gb rdtscp lm constant\_tsc rep\_good nopl xtopology nonstop\_tsc cpuid tsc\_known\_freq pni pclmulqdq ssse3 fma cx16 pcid sse4\_1 sse4\_2 x2apic movbe popcnt aes xsave avx f16c rdrand hypervisor lahf\_lm abm 3dnowprefetch invpcid\_single ssbd ibrs ibpb stibp fsgsbase tsc\_adjust bmi1 hle avx2 smep bmi2 erms invpcid rtm rdseed adx smap xsaveopt arat md\_clear arch\_capabilities

bugs : cpu\_meltdown spectre\_v1 spectre\_v2 spec\_store\_bypass l1tf mds swapgs taa mmio\_stale\_data retbleed

bogomips : 4399.99

clflush size : 64

cache\_alignment : 64

address sizes : 46 bits physical, 48 bits virtual

power management:

processor : 1

vendor\_id : GenuineIntel

cpu family : 6

model : 79

model name : Intel(R) Xeon(R) CPU @ 2.20GHz

stepping : 0

microcode : 0xffffffff

cpu MHz : 2199.998

cache size : 56320 KB

physical id : 0

siblings : 2

core id : 0

cpu cores : 1

apicid : 1

initial apicid : 1

fpu : yes

fpu\_exception : yes

cpuid level : 13

wp : yes

flags : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse36 clflush mmx fxsr sse sse2 ss ht syscall nx pdpe1gb rdtscp lm constant\_tsc rep\_good nopl xtopology nonstop\_tsc cpuid tsc\_known\_freq pni pclmulqdq ssse3 fma cx16 pcid sse4\_1 sse4\_2 x2apic movbe popcnt aes xsave avx f16c rdrand hypervisor lahf\_lm abm 3dnowprefetch invpcid\_single ssbd ibrs ibpb stibp fsgsbase tsc\_adjust bmi1 hle avx2 smep bmi2 erms invpcid rtm rdseed adx smap xsaveopt arat md\_clear arch\_capabilities

bugs : cpu\_meltdown spectre\_v1 spectre\_v2 spec\_store\_bypass l1tf mds swapgs taa mmio\_stale\_data retbleed

bogomips : 4399.99

clflush size : 64

cache\_alignment : 64

address sizes : 46 bits physical, 48 bits virtual

power management:

!cat /proc/meminfo

MemTotal: 13294244 kB

MemFree: 8939212 kB

MemAvailable: 11941300 kB

Buffers: 68264 kB

Cached: 3125788 kB

SwapCached: 0 kB

Active: 923096 kB

Inactive: 3198304 kB

Active(anon): 1196 kB

Inactive(anon): 927600 kB

Active(file): 921900 kB

Inactive(file): 2270704 kB

Unevictable: 12 kB

Mlocked: 12 kB

SwapTotal: 0 kB

SwapFree: 0 kB

Dirty: 5428 kB

Writeback: 0 kB

AnonPages: 925124 kB

Mapped: 562680 kB

Shmem: 1444 kB

KReclaimable: 101168 kB

Slab: 138096 kB

SReclaimable: 101168 kB

SUnreclaim: 36928 kB

KernelStack: 5680 kB

PageTables: 15892 kB

NFS\_Unstable: 0 kB

Bounce: 0 kB

WritebackTmp: 0 kB

CommitLimit: 6647120 kB

Committed\_AS: 2945048 kB

VmallocTotal: 34359738367 kB

VmallocUsed: 10628 kB

VmallocChunk: 0 kB

Percpu: 1312 kB

HardwareCorrupted: 0 kB

AnonHugePages: 22528 kB

ShmemHugePages: 0 kB

ShmemPmdMapped: 0 kB

FileHugePages: 0 kB

FilePmdMapped: 0 kB

CmaTotal: 0 kB

CmaFree: 0 kB

HugePages\_Total: 0

HugePages\_Free: 0

HugePages\_Rsvd: 0

HugePages\_Surp: 0

Hugepagesize: 2048 kB

Hugetlb: 0 kB

DirectMap4k: 80696 kB

DirectMap2M: 4110336 kB

DirectMap1G: 11534336 kB

Resultados

n\_days 7 {'units': 128, 'lr': 0.004203071113584753, 'dropout': 0.30000000000000004, 'recurrent\_dropout': 0.0, 'batch\_size': 64, 'tuner/epochs': 50, 'tuner/initial\_epoch': 17, 'tuner/bracket': 2, 'tuner/round': 2, 'tuner/trial\_id': '0071'}

otro 7 <class '\_\_main\_\_.MyTuner'>

Results summary

Results in /content/drive/MyDrive/a Tesis de Grado/Tesis Versiones/tesis2023/modelos /salidasModelosHypertuner/hyperband/SMAPE-max\_trials50- epochs100-aleatorio 7

Showing 1 best trials

Objective(name="val\_loss", direction="min")

Trial 0073 summary

Hyperparameters:

units: 128

lr: 0.004203071113584753

dropout: 0.30000000000000004

recurrent\_dropout: 0.0

batch\_size: 64

tuner/epochs: 50

tuner/initial\_epoch: 17

tuner/bracket: 2

tuner/round: 2

tuner/trial\_id: 0071

Score: 0.08344406634569168

results\_summary <class 'NoneType'>

n\_days 17 {'units': 384, 'lr': 0.0002571412776290704, 'dropout': 0.0, 'recurrent\_dropout': 0.0, 'batch\_size': 32, 'tuner/epochs': 50, 'tuner/initial\_epoch': 0, 'tuner/bracket': 0, 'tuner/round': 0}

otro 17 <class '\_\_main\_\_.MyTuner'>

WARNING:tensorflow:Detecting that an object or model or tf.train.Checkpoint is being deleted with unrestored values. See the following logs for the specific values in question. To silence these warnings, use `status.expect\_partial()`. See <https://www.tensorflow.org/api_docs/python/tf/train/Checkpoint#restorefor> details about the status object returned by the restore function.

WARNING:tensorflow:Value in checkpoint could not be found in the restored object: (root).optimizer.\_variables.16

WARNING:tensorflow:Detecting that an object or model or tf.train.Checkpoint is being deleted with unrestored values. See the following logs for the specific values in question. To silence these warnings, use `status.expect\_partial()`. See <https://www.tensorflow.org/api_docs/python/tf/train/Checkpoint#restorefor> details about the status object returned by the restore function.

Results summary

Results in /content/drive/MyDrive/a Tesis de Grado/Tesis Versiones/tesis2023/modelos /salidasModelosHypertuner/hyperband/SMAPE-max\_trials50- epochs100-aleatorio 17

Showing 1 best trials

Objective(name="val\_loss", direction="min")

Trial 0087 summary

Hyperparameters:

units: 384

lr: 0.0002571412776290704

dropout: 0.0

recurrent\_dropout: 0.0

batch\_size: 32

tuner/epochs: 50

tuner/initial\_epoch: 0

tuner/bracket: 0

tuner/round: 0

Score: 0.08234339207410812

results\_summary <class 'NoneType'>

n\_days 27 {'units': 288, 'lr': 0.0013600983479762927, 'dropout': 0.1, 'recurrent\_dropout': 0.1, 'batch\_size': 32, 'tuner/epochs': 50, 'tuner/initial\_epoch': 0, 'tuner/bracket': 0, 'tuner/round': 0}

otro 27 <class '\_\_main\_\_.MyTuner'>

Results summary

Results in /content/drive/MyDrive/a Tesis de Grado/Tesis Versiones/tesis2023/modelos /salidasModelosHypertuner/hyperband/SMAPE-max\_trials50- epochs100-aleatorio 27

Showing 1 best trials

Objective(name="val\_loss", direction="min")

Trial 0089 summary

Hyperparameters:

units: 288

lr: 0.0013600983479762927

dropout: 0.1

recurrent\_dropout: 0.1

batch\_size: 32

tuner/epochs: 50

tuner/initial\_epoch: 0

tuner/bracket: 0

tuner/round: 0

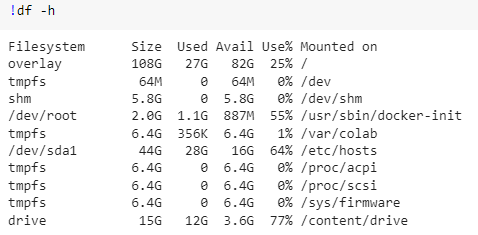
Score: 0.07837852090597153

results\_summary <class 'NoneType'>

# ***Algoritmo RandomSearch***

Tiempo inicio 9:24 22/08/2023

Paro 16:32



!cat /proc/cpuinfo

processor : 0

vendor\_id : GenuineIntel

cpu family : 6

model : 79

model name : Intel(R) Xeon(R) CPU @ 2.20GHz

stepping : 0

microcode : 0xffffffff

cpu MHz : 2200.202

cache size : 56320 KB

physical id : 0

siblings : 2

core id : 0

cpu cores : 1

apicid : 0

initial apicid : 0

fpu : yes

fpu\_exception : yes

cpuid level : 13

wp : yes

flags : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse36 clflush mmx fxsr sse sse2 ss ht syscall nx pdpe1gb rdtscp lm constant\_tsc rep\_good nopl xtopology nonstop\_tsc cpuid tsc\_known\_freq pni pclmulqdq ssse3 fma cx16 pcid sse4\_1 sse4\_2 x2apic movbe popcnt aes xsave avx f16c rdrand hypervisor lahf\_lm abm 3dnowprefetch invpcid\_single ssbd ibrs ibpb stibp fsgsbase tsc\_adjust bmi1 hle avx2 smep bmi2 erms invpcid rtm rdseed adx smap xsaveopt arat md\_clear arch\_capabilities

bugs : cpu\_meltdown spectre\_v1 spectre\_v2 spec\_store\_bypass l1tf mds swapgs taa mmio\_stale\_data retbleed

bogomips : 4400.40

clflush size : 64

cache\_alignment : 64

address sizes : 46 bits physical, 48 bits virtual

power management:

processor : 1

vendor\_id : GenuineIntel

cpu family : 6

model : 79

model name : Intel(R) Xeon(R) CPU @ 2.20GHz

stepping : 0

microcode : 0xffffffff

cpu MHz : 2200.202

cache size : 56320 KB

physical id : 0

siblings : 2

core id : 0

cpu cores : 1

apicid : 1

initial apicid : 1

fpu : yes

fpu\_exception : yes

cpuid level : 13

wp : yes

flags : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse36 clflush mmx fxsr sse sse2 ss ht syscall nx pdpe1gb rdtscp lm constant\_tsc rep\_good nopl xtopology nonstop\_tsc cpuid tsc\_known\_freq pni pclmulqdq ssse3 fma cx16 pcid sse4\_1 sse4\_2 x2apic movbe popcnt aes xsave avx f16c rdrand hypervisor lahf\_lm abm 3dnowprefetch invpcid\_single ssbd ibrs ibpb stibp fsgsbase tsc\_adjust bmi1 hle avx2 smep bmi2 erms invpcid rtm rdseed adx smap xsaveopt arat md\_clear arch\_capabilities

bugs : cpu\_meltdown spectre\_v1 spectre\_v2 spec\_store\_bypass l1tf mds swapgs taa mmio\_stale\_data retbleed

bogomips : 4400.40

clflush size : 64

cache\_alignment : 64

address sizes : 46 bits physical, 48 bits virtual

power management:

!cat /proc/meminfo

MemTotal: 13294252 kB

MemFree: 8551400 kB

MemAvailable: 11659584 kB

Buffers: 68224 kB

Cached: 3230236 kB

SwapCached: 0 kB

Active: 920492 kB

Inactive: 3573112 kB

Active(anon): 1148 kB

Inactive(anon): 1195540 kB

Active(file): 919344 kB

Inactive(file): 2377572 kB

Unevictable: 12 kB

Mlocked: 12 kB

SwapTotal: 0 kB

SwapFree: 0 kB

Dirty: 436 kB

Writeback: 0 kB

AnonPages: 1193212 kB

Mapped: 562396 kB

Shmem: 1540 kB

KReclaimable: 104444 kB

Slab: 142192 kB

SReclaimable: 104444 kB

SUnreclaim: 37748 kB

KernelStack: 5584 kB

PageTables: 23200 kB

NFS\_Unstable: 0 kB

Bounce: 0 kB

WritebackTmp: 0 kB

CommitLimit: 6647124 kB

Committed\_AS: 2971064 kB

VmallocTotal: 34359738367 kB

VmallocUsed: 10532 kB

VmallocChunk: 0 kB

Percpu: 1296 kB

HardwareCorrupted: 0 kB

AnonHugePages: 22528 kB

ShmemHugePages: 0 kB

ShmemPmdMapped: 0 kB

FileHugePages: 0 kB

FilePmdMapped: 0 kB

CmaTotal: 0 kB

CmaFree: 0 kB

HugePages\_Total: 0

HugePages\_Free: 0

HugePages\_Rsvd: 0

HugePages\_Surp: 0

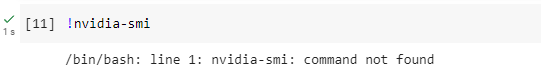
Hugepagesize: 2048 kB

Hugetlb: 0 kB

DirectMap4k: 72504 kB

DirectMap2M: 5167104 kB

DirectMap1G: 10485760 kB



El comando de arriba debería de darme la gpu usada pero tampoco se activó en esta ejecución

Habia sido que habia que ir a una parte de la configuración y activar manualmente antes de ejecutar

Paro 16:32

Ya había hecho el 7 y 17 n days atras

Continúo a las 16:43 con gpu

Completo 18:25

!df -h

Filesystem Size Used Avail Use% Mounted on

overlay 79G 27G 52G 34% /

tmpfs 64M 0 64M 0% /dev

shm 5.7G 0 5.7G 0% /dev/shm

/dev/root 2.0G 1.1G 887M 55% /usr/sbin/docker-init

tmpfs 6.4G 80K 6.4G 1% /var/colab

/dev/sda1 50G 29G 22G 58% /opt/bin/.nvidia

tmpfs 6.4G 0 6.4G 0% /proc/acpi

tmpfs 6.4G 0 6.4G 0% /proc/scsi

tmpfs 6.4G 0 6.4G 0% /sys/firmware

drive 15G 9.0G 6.1G 60% /content/drive

!cat /proc/cpuinfo

processor : 0

vendor\_id : GenuineIntel

cpu family : 6

model : 63

model name : Intel(R) Xeon(R) CPU @ 2.30GHz

stepping : 0

microcode : 0xffffffff

cpu MHz : 2299.998

cache size : 46080 KB

physical id : 0

siblings : 2

core id : 0

cpu cores : 1

apicid : 0

initial apicid : 0

fpu : yes

fpu\_exception : yes

cpuid level : 13

wp : yes

flags : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse36 clflush mmx fxsr sse sse2 ss ht syscall nx pdpe1gb rdtscp lm constant\_tsc rep\_good nopl xtopology nonstop\_tsc cpuid tsc\_known\_freq pni pclmulqdq ssse3 fma cx16 pcid sse4\_1 sse4\_2 x2apic movbe popcnt aes xsave avx f16c rdrand hypervisor lahf\_lm abm invpcid\_single ssbd ibrs ibpb stibp fsgsbase tsc\_adjust bmi1 avx2 smep bmi2 erms invpcid xsaveopt arat md\_clear arch\_capabilities

bugs : cpu\_meltdown spectre\_v1 spectre\_v2 spec\_store\_bypass l1tf mds swapgs mmio\_stale\_data retbleed

bogomips : 4599.99

clflush size : 64

cache\_alignment : 64

address sizes : 46 bits physical, 48 bits virtual

power management:

processor : 1

vendor\_id : GenuineIntel

cpu family : 6

model : 63

model name : Intel(R) Xeon(R) CPU @ 2.30GHz

stepping : 0

microcode : 0xffffffff

cpu MHz : 2299.998

cache size : 46080 KB

physical id : 0

siblings : 2

core id : 0

cpu cores : 1

apicid : 1

initial apicid : 1

fpu : yes

fpu\_exception : yes

cpuid level : 13

wp : yes

flags : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse36 clflush mmx fxsr sse sse2 ss ht syscall nx pdpe1gb rdtscp lm constant\_tsc rep\_good nopl xtopology nonstop\_tsc cpuid tsc\_known\_freq pni pclmulqdq ssse3 fma cx16 pcid sse4\_1 sse4\_2 x2apic movbe popcnt aes xsave avx f16c rdrand hypervisor lahf\_lm abm invpcid\_single ssbd ibrs ibpb stibp fsgsbase tsc\_adjust bmi1 avx2 smep bmi2 erms invpcid xsaveopt arat md\_clear arch\_capabilities

bugs : cpu\_meltdown spectre\_v1 spectre\_v2 spec\_store\_bypass l1tf mds swapgs mmio\_stale\_data retbleed

bogomips : 4599.99

clflush size : 64

cache\_alignment : 64

address sizes : 46 bits physical, 48 bits virtual

power management:

!cat /proc/meminfo

MemTotal: 13294244 kB

MemFree: 9287496 kB

MemAvailable: 11649068 kB

Buffers: 70604 kB

Cached: 2491856 kB

SwapCached: 0 kB

Active: 745416 kB

Inactive: 2943608 kB

Active(anon): 1060 kB

Inactive(anon): 1126796 kB

Active(file): 744356 kB

Inactive(file): 1816812 kB

Unevictable: 12 kB

Mlocked: 12 kB

SwapTotal: 0 kB

SwapFree: 0 kB

Dirty: 872 kB

Writeback: 0 kB

AnonPages: 1125324 kB

Mapped: 559096 kB

Shmem: 1288 kB

KReclaimable: 83004 kB

Slab: 121864 kB

SReclaimable: 83004 kB

SUnreclaim: 38860 kB

KernelStack: 5824 kB

PageTables: 21364 kB

NFS\_Unstable: 0 kB

Bounce: 0 kB

WritebackTmp: 0 kB

CommitLimit: 6647120 kB

Committed\_AS: 2891136 kB

VmallocTotal: 34359738367 kB

VmallocUsed: 74616 kB

VmallocChunk: 0 kB

Percpu: 1320 kB

HardwareCorrupted: 0 kB

AnonHugePages: 24576 kB

ShmemHugePages: 0 kB

ShmemPmdMapped: 0 kB

FileHugePages: 0 kB

FilePmdMapped: 0 kB

CmaTotal: 0 kB

CmaFree: 0 kB

HugePages\_Total: 0

HugePages\_Free: 0

HugePages\_Rsvd: 0

HugePages\_Surp: 0

Hugepagesize: 2048 kB

Hugetlb: 0 kB

DirectMap4k: 164664 kB

DirectMap2M: 2977792 kB

DirectMap1G: 12582912 kB



# Resultados Randomsearch

WARNING:tensorflow:Layer gru\_2 will not use cuDNN kernels since it doesn't meet the criteria. It will use a generic GPU kernel as fallback when running on GPU.

WARNING:tensorflow:Layer gru\_3 will not use cuDNN kernels since it doesn't meet the criteria. It will use a generic GPU kernel as fallback when running on GPU.

n\_days 7 {'units': 288, 'lr': 0.00011291796479765002, 'dropout': 0.2, 'recurrent\_dropout': 0.30000000000000004, 'batch\_size': 32}

otro 7 <class 'keras\_tuner.tuners.randomsearch.RandomSearch'>

WARNING:tensorflow:Layer gru will not use cuDNN kernels since it doesn't meet the criteria. It will use a generic GPU kernel as fallback when running on GPU.

WARNING:tensorflow:Layer gru\_1 will not use cuDNN kernels since it doesn't meet the criteria. It will use a generic GPU kernel as fallback when running on GPU.

WARNING:tensorflow:Layer gru\_2 will not use cuDNN kernels since it doesn't meet the criteria. It will use a generic GPU kernel as fallback when running on GPU.

Results summary

Results in /content/drive/MyDrive/a Tesis de Grado/Tesis Versiones/tesis2023/modelos /salidasModelosHypertuner/randomSearch/SMAPE-max\_trials50- epochs100-aleatorio 7

Showing 1 best trials

Objective(name="val\_loss", direction="min")

Trial 48 summary

Hyperparameters:

units: 288

lr: 0.00011291796479765002

dropout: 0.2

recurrent\_dropout: 0.30000000000000004

batch\_size: 32

Score: 0.08344393968582153

results\_summary <class 'NoneType'>

WARNING:tensorflow:Layer gru\_3 will not use cuDNN kernels since it doesn't meet the criteria. It will use a generic GPU kernel as fallback when running on GPU.

n\_days 17 {'units': 192, 'lr': 0.0001948525099608208, 'dropout': 0.1, 'recurrent\_dropout': 0.2, 'batch\_size': 32}

otro 17 <class 'keras\_tuner.tuners.randomsearch.RandomSearch'>

WARNING:tensorflow:Layer gru will not use cuDNN kernels since it doesn't meet the criteria. It will use a generic GPU kernel as fallback when running on GPU.

WARNING:tensorflow:Layer gru\_1 will not use cuDNN kernels since it doesn't meet the criteria. It will use a generic GPU kernel as fallback when running on GPU.

WARNING:tensorflow:Detecting that an object or model or tf.train.Checkpoint is being deleted with unrestored values. See the following logs for the specific values in question. To silence these warnings, use `status.expect\_partial()`. See <https://www.tensorflow.org/api_docs/python/tf/train/Checkpoint#restorefor> details about the status object returned by the restore function.

Results summary

Results in /content/drive/MyDrive/a Tesis de Grado/Tesis Versiones/tesis2023/modelos /salidasModelosHypertuner/randomSearch/SMAPE-max\_trials50- epochs100-aleatorio 17

Showing 1 best trials

Objective(name="val\_loss", direction="min")

Trial 37 summary

Hyperparameters:

units: 192

lr: 0.0001948525099608208

dropout: 0.1

recurrent\_dropout: 0.2

batch\_size: 32

Score: 0.0822763592004776

results\_summary <class 'NoneType'>

n\_days 27 {'units': 128, 'lr': 0.0046862973390983135, 'dropout': 0.2, 'recurrent\_dropout': 0.0, 'batch\_size': 32}

otro 27 <class 'keras\_tuner.tuners.randomsearch.RandomSearch'>

WARNING:tensorflow:Detecting that an object or model or tf.train.Checkpoint is being deleted with unrestored values. See the following logs for the specific values in question. To silence these warnings, use `status.expect\_partial()`. See <https://www.tensorflow.org/api_docs/python/tf/train/Checkpoint#restorefor> details about the status object returned by the restore function.

Results summary

Results in /content/drive/MyDrive/a Tesis de Grado/Tesis Versiones/tesis2023/modelos /salidasModelosHypertuner/randomSearch/SMAPE-max\_trials50- epochs100-aleatorio 27

Showing 1 best trials

Objective(name="val\_loss", direction="min")

Trial 18 summary

Hyperparameters:

units: 128

lr: 0.0046862973390983135

dropout: 0.2

recurrent\_dropout: 0.0

batch\_size: 32

Score: 0.0783756896853447

results\_summary <class 'NoneType'>

# Algoritmo Bayesian optimitation

Está en el Word tiempos y resolución 2024

# Marco teórico de los algoritmos de tunning de hyperparametros

Básicamente esta sección es para justificar algunas elecciones de los hyperparametros .

Definiciones teóricas de los hyperparemetros

## Batch Size:

If we take out a subsample of the dataset, it should represent the properties of the whole dataset. This batch can be used to calculate the gradient and update the weights. And we iterate over all the subsamples until we cover the whole dataset. The idea is to save memory space. But you need to choose the optimal value of the size of the subset, because a lesser batch size would cause more fluctuations while reaching the minima, and a greater value can cause memory errors.

Activation Function:

Activation functions are used to introduce a nonlinearity on each node. Few things we need to make sure while deciding activation functions are, they are to be used on thousands and millions of nodes, and back propagation uses their derivatives, so both the function and its derivative should be less computationally complex. Some of the widely used activations are ReLU, Sigmoid, and Leaky ReLU.

## Loss Function:

Loss function is chosen on the basis of output, whether it’s a binary classification, multiclass classification, regression, and so forth. There are also other factors. For example, using sigmoid activation on the last layer and quadratic loss function can result in learning slow down. So things like these are need to be taken care of. And there are internal hyperparameters for loss function as well which can be tuned.

## Optimizer:

In the “Understanding Hyperparameters” section we discussed an optimization method, gradient descent. There are other, more advanced optimization methods, like Adagrad, Adam Optimizer, and so on, and these optimizers also contain various hyperparameters that affect the overall optimization.

# Random Search

It randomly picks hyperparameters, makes a set, and trains the model on it. This method may not find the best set, but there are higher chances of finding a near best set saving a huge amount of time. Unlike grid search, instead of spending a large amount of time on unpromising candidates, random search jumps to random hyperparameters, and even though it does not learn from its past results, it usually delivers satisfactory results. In random search, we define the number of trials, which is the number of sets of hyperparameters to be tried.

## Hyperband Search

Hyperband is a bandit-based approach for solving the problem of hyperparameter optimization. The bandit-based approach addresses our problem perfectly here: we have a limited amount of resources but we need to allocate them to all our trials efficiently. Again we spend more time on better-performing models instead of wasting our resources and time on poor configuration of hyperparameters. Hyperband is an extended version of successive halving (mas info en el libro sobre cómo funciona).

## Bayesian optimitation

Suppose an algorithm could keep a log of all the previous observations and learn from them. For example, suppose it could observe that our model is being optimized near certain values of hyperparameters and could exploit this valuable information and proceed to the hyperparameters nearest to those good-performing ones, hence learning from its history. By doing so, the algorithm would not waste time on bad-performing hyperparameters while reaching the best-performing hyperparameters

Mas definiciones de <https://towardsdatascience.com/epoch-vs-iterations-vs-batch-size-4dfb9c7ce9c9>

## Epochs

One Epoch is when an ENTIRE dataset is passed forward and backward through the neural network only ONCE

## Batch Size

Total number of training examples present in a single batch.

## Iterations

Iterations is the number of batches needed to complete one epoch.

## Cantidad de trials

Cada trial tiene 100 epochs yo probé 50 trials

Según <https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf?ref=blog.floydhub.com>

Random Search for Hyper-Parameter Optimization

32 trials es suficiente para el dataset de MNIST(el de imágenes de numeros)

Según este otro <https://arxiv.org/abs/1912.06059>

[Submitted on 12 Dec 2019]

**Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS**

Cnn CIFAR- 10 un dataset de imágenes

This dataset consists of 50,000 training images and 10,000 test images

Each architecture search is run for 50 epochs on Nvidia Tesla K80 GPU.

The whole training procedure of 5 runs

The best model shows about 86% accuracy. As you can see, this algorithm is faster than Grid Search, but if we take more runs it will be much longer.

También había otro que hacia 300 trials o más.

A seguir buscando cantidad de epochs y trials para los hypertuner que justifique los que usé

También la cantidad de días que mira hacia atrás es de 7 luego 10 en 10 hasta 27

Una mejor justificación seria de 7 en 7 hasta 28 donde se supone agarraríamos de 1 semana en semana hasta casi un mes.

# Fuentes

Libro Hyperparameter Optimization in Machine Learning. Apress, Berkeley, CA. <https://doi.org/10.1007/978-1-4842-6579-6>