OGÓLNIE

1. Data Augmentation

CNN1 – mała sieć, którą sami napisaliśmy

* Różne podejścia – z dropoutem lub bez
* Dwa modele: jeden większy drugi mniejszy

Eksperymenty:

1. Batch size: na małej sieci – najlepszy: 4
2. Learning Rate: na małej sieci – najlepszy: 10e-3
3. Epochs: na małej sieci – najlepszy: 15-20 – przy większych trochę overfitting
4. Dropout: na małej sieci – najlepszy 0.25
5. **DONE**: Weight decay <- ADAMW
6. Na dużej sieci – dla najlepszych parametrów tylko wytrenowana o 75% Accuracy
7. Można porobić eksperymenty na dużej sieci jeszcze

CNN2 – pretrained i Resnet18 ze zmieniona ostatnią warstwą na dwie warstwy fully connected

Eksperymenty:

1. TODO: Batch size
2. TOOD: learning rate
3. TODO: Liczba Epoch
4. TODO: Weight decay

CNN3 – zamrozić początkowe warstwy, ale trenować ostatnią warstwę + klasyfikator

1. Batch size
2. Adam optimizer + Weight decay
3. Spadający learning rate !

Do rozważenia:

1. Dodać Softmax klasyfikator na sam koniec – poczytać co to zmienia czy warto itd. <- chyba nie jeśli CrossEntropyLoss, bo to już zawiera Softmax
2. Tensorboard – wizualizacja danych

PLAN

SIECI:

1. Small Net
2. Bigger Net
3. Resnet18 with FC trainable
4. Resnet18 with Last Layer + FC trainable

EXPERIMENTS:

1. Batch Size
2. Learning Rate
3. Nr of Epoch
4. Dropout Rate
5. Weight Decay
6. Data Augmentation
7. Optimizer

NOTEBOOKS TODO:

1. SMALL NET -> 40 min
   1. Jakies drobne experymenty
2. BIGGER NET -> 40 min
   1. Batch Size
   2. Learning Rate – with best batch size
   3. Dropout - with best BS and LR
   4. Weight Decay – with best BS and LR with/without dropout
3. Resnet\_18 FC
4. Resnet layer 4
   1. Batch Size
   2. Learning Rate - decreasing
   3. Weight Decay

PYTANIA:

1. Czy eksperymenty mogą być przeprowadzone tylko na małych sieciach – bo krócej się liczą?
2. Czy to możliwe że te sieci liczą się tak długo? – około 2-3h: 30 Epoch
3. Augmentacja – czy wystarczy tylko dodać na koniec?