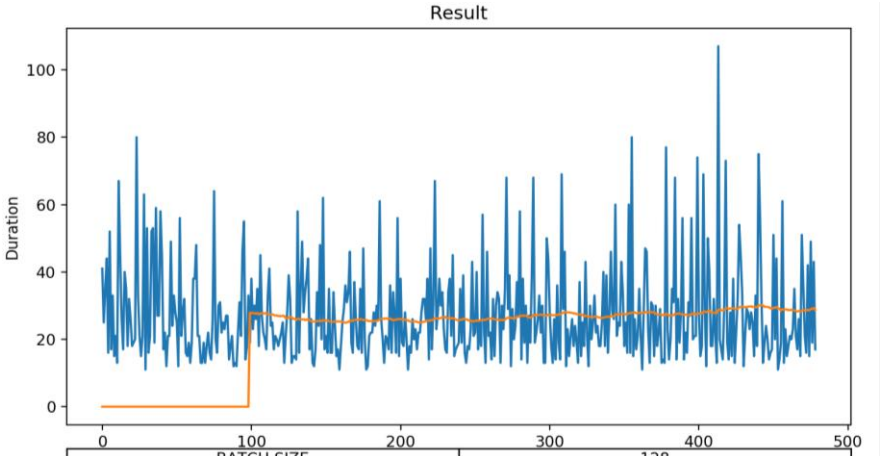
Various network architectures and parameters for DRL simulation

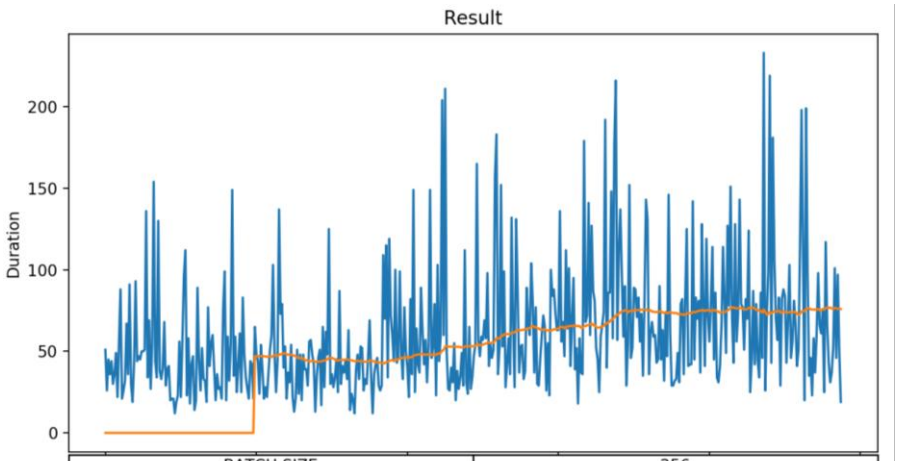
Group K: Fioritto Alessandro, Pilutti Filippo, Simonitti Samuele

The base architecture with no changes in the parameters, as shown in the lecture



BATCH SIZE	128
GAMMA	0.999
TARGET UPDATE	10
EPSILON START	0.9
EPSILON END	0.05
EPSILON DECAY	200
Training time (s)	6066
Max peak	107
Maximal Peak Episode	430

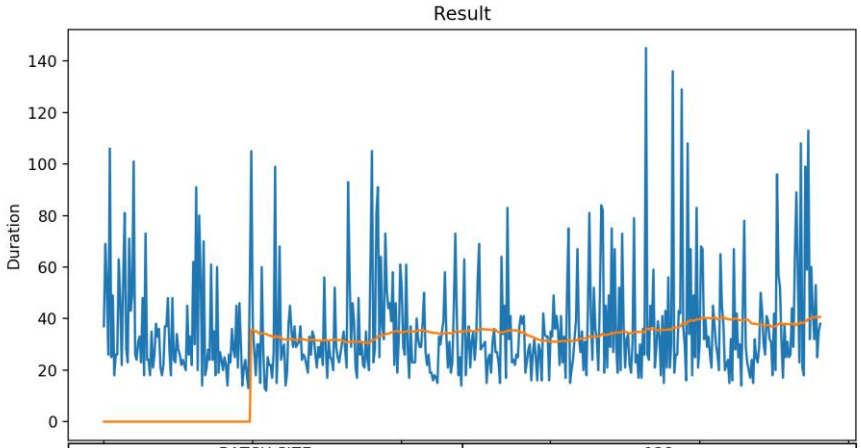
Four Hidden layers architecture



BATCH SIZE	256
GAMMA	0.999
TARGET UPDATE	10
EPSILON START	0.9
EPSILON END	0.01
EPSILON DECAY	2500
LR	0.0001
Optimizer	adamw
Training time (s)	10794
Max peak	233

We modified the architecture adding a hidden layer, the batch size, the optimizer, the EPSILON END and the EPSILON DECAY and, even if we had a far larger training time, we had a far better mean of duration. Even if the peak is lower, this seems like a better result than the previous one.

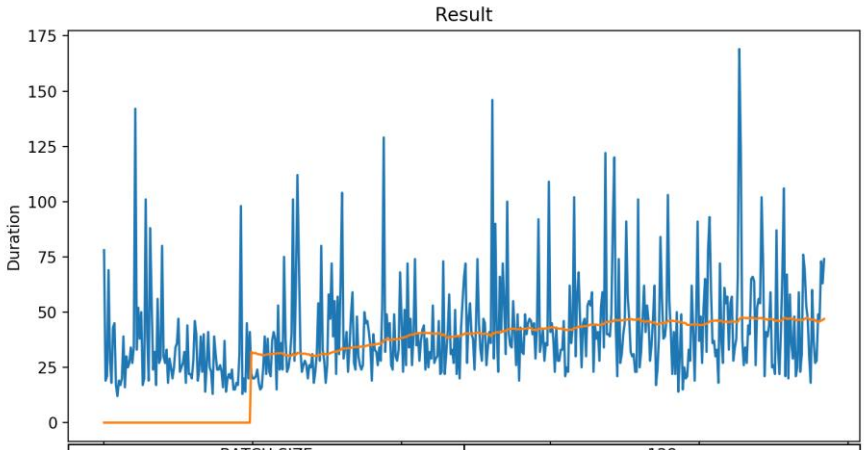
Wide but Shallow architecture with slightly lesser GAMMA value and 10 times the EPSILON DECAY



BATCH SIZE	128
GAMMA	0.95
TARGET UPDATE	10
EPSILON START	0.9
EPSILON END	0.05
EPSILON DECAY	2000
LR	0.001
Optimizer	rmsprop
Training time (s)	5515
Max peak	145

We tried an architecture composed of two fully connected layers with 512 nodes. With a slightly faster training speed the results are not that far from the base models seen during the lecture.

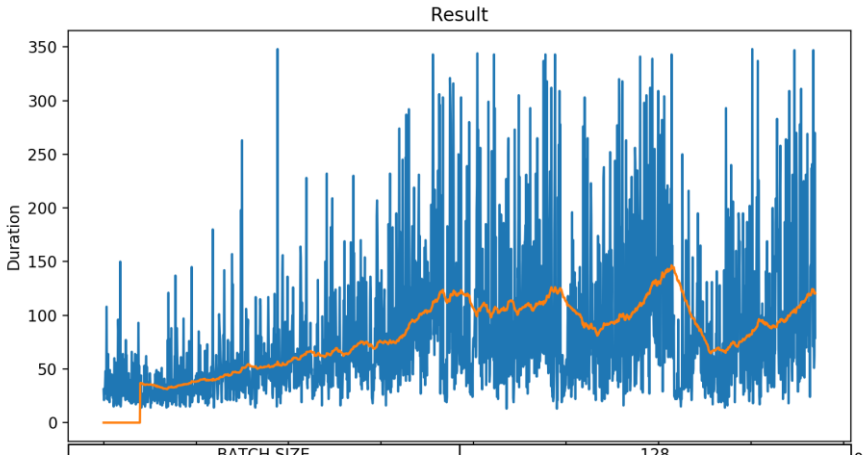
Residual architecture



BATCH SIZE	128
GAMMA	0.999
TARGET UPDATE	10
EPSILON START	0.9
EPSILON END	0.05
EPSILON DECAY	200
LR	0.001
Optimizer	rmsprop
Training time (s)	6954
Max peak	169

We tried to use residual blocks with skip connections to help the model preserve the input information.

We also tried to train for 2000 episodes obtaining better results



BATCH SIZE	128
GAMMA	0.999
TARGET UPDATE	10
EPSILON START	0.9
EPSILON END	0.05
EPSILON DECAY	1500
LR	0.001
Optimizer	rmsprop
Training time (s)	45780
Max peak	348