Machine learning based middle-layer for autonomous accelerator operation and control

THAL03

21/10/2021

INFN-LNF :: **S. Pioli**, B. Buonomo, F. Cardelli, P. Ciuffetti, L. G. Foggetta, C. Di Giulio, D. Di Giovenale, G. Piermarini

INFN-LNL :: V. Martinelli

ICALEPCS 2021

Outline

Singularity project aim to develop automated middle-layer to control accelerator operation through machine learning (ML) algorithms.

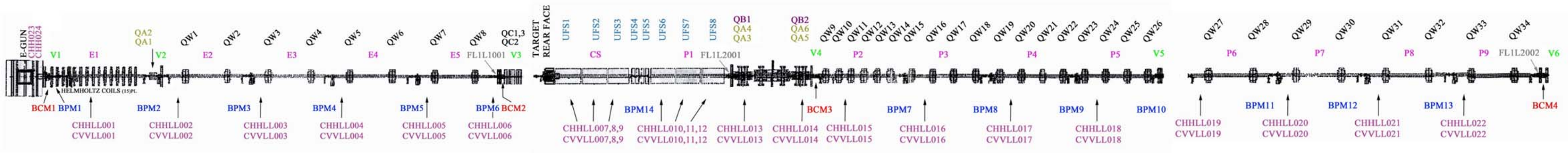
General capabilities:

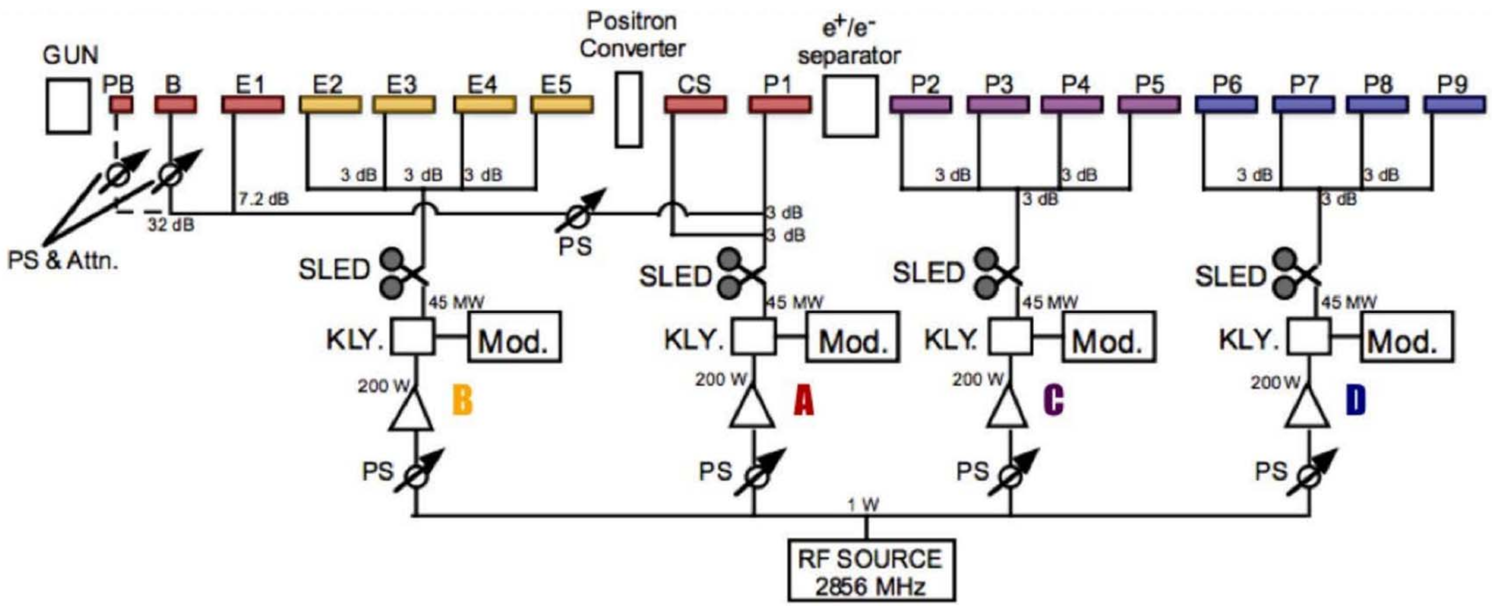
- Machine indipendent
- Device misalignment and performance drift indipendent

Reinforcement Learning based tools for autonomous:

- Particle beam energy tuning
- Particle beam charge optimization

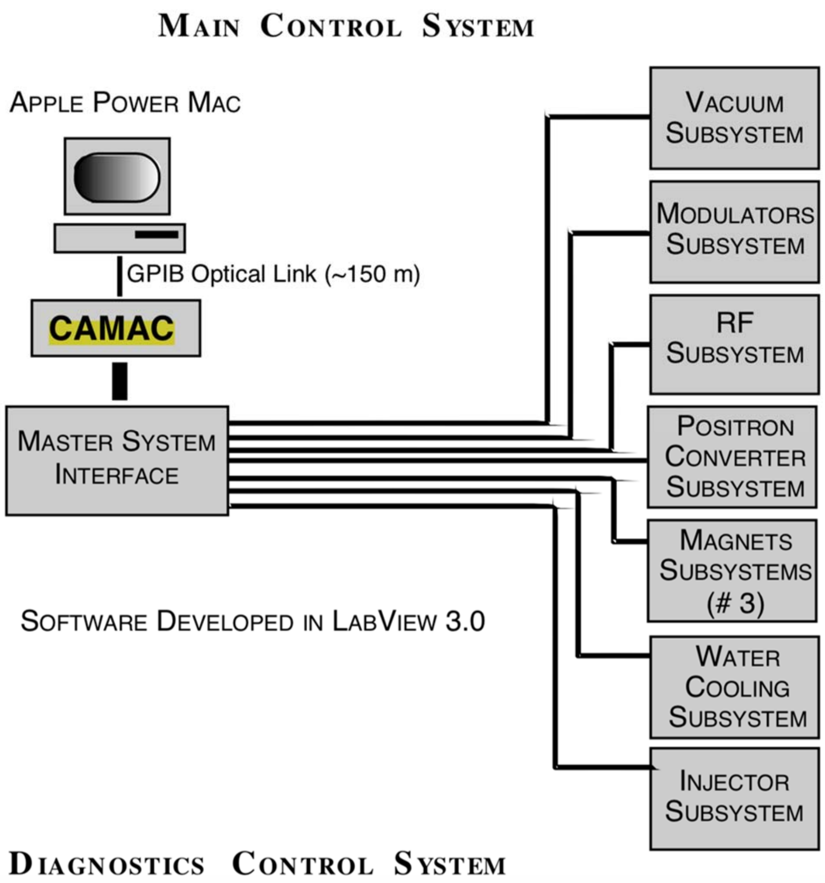
Dafne LINAC



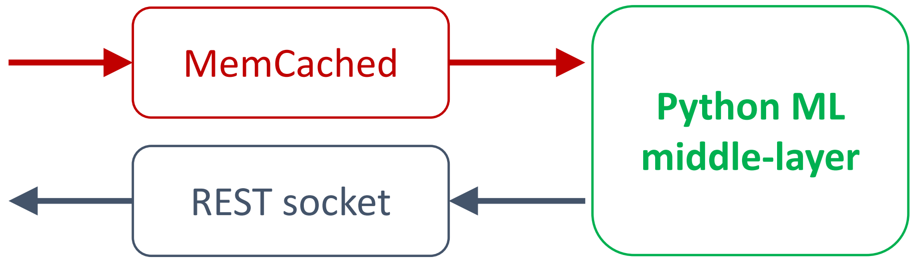


	Design	Operational
Electron beam final energy	800 MeV	510 MeV
Positron beam final energy	550 MeV	510 MeV
RF frequency	2856 MHz	
Positron conversion energy	250 MeV	220 MeV
Beam pulse rep. rate	1 to 50 Hz	1 to 50 Hz
Beam macropulse length	10 nsec	1.4 to 300 nsec
Gun current	8 A	8 A
Beam spot on positron converter	1 mm	1 mm
norm. Emittance (mm. mrad)	1 (electron) 10 (positron)	< 1.5
rms Energy spread	0.5% (electron) 1.0% (positron)	0.5% (electron) 1.0% (positron)
electron current on positron converter	5 A	5.2 A
Max output electron current	>150 mA	500 mA
Max output positron current	36 mA	85 mA
Trasport efficiency from capture section to linac end	90%	90%
Accelerating structure	SLAC-type, CG, 2π/3	
RF source	4 x 45 MWp sledded klystrons TH2128C	

Dafne LINAC control system



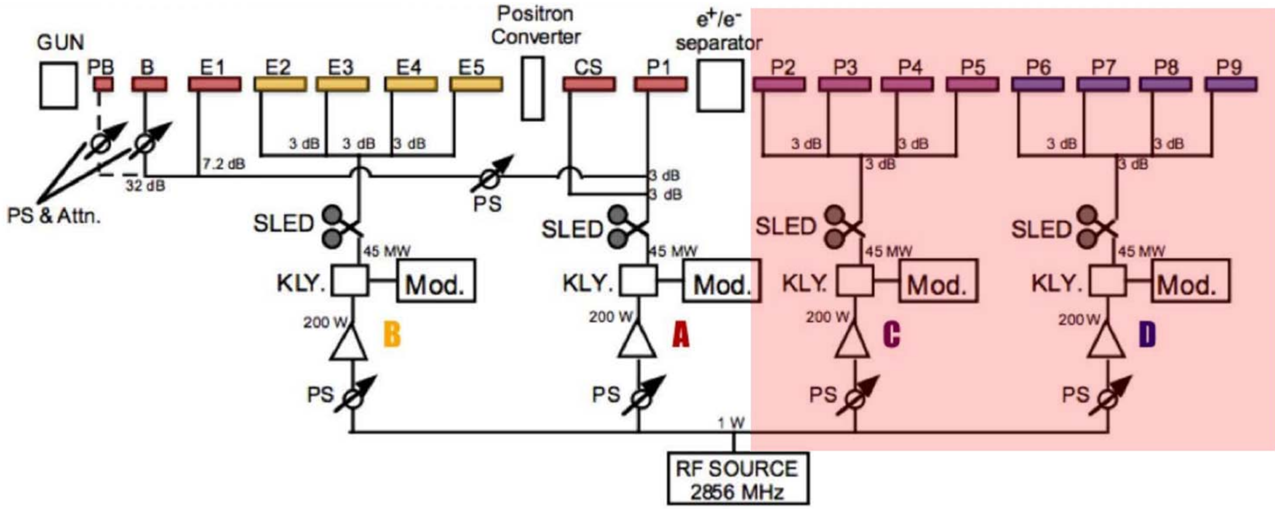
Machine Learning middle-layer acquire LINAC data to process RL tools and interlocks to break operations.



Currently running with off-line data.

Dafne LINAC beam energy tuning tool

Automate accelerator beam energy tuning.



Environment:

- 2 RF Sources:

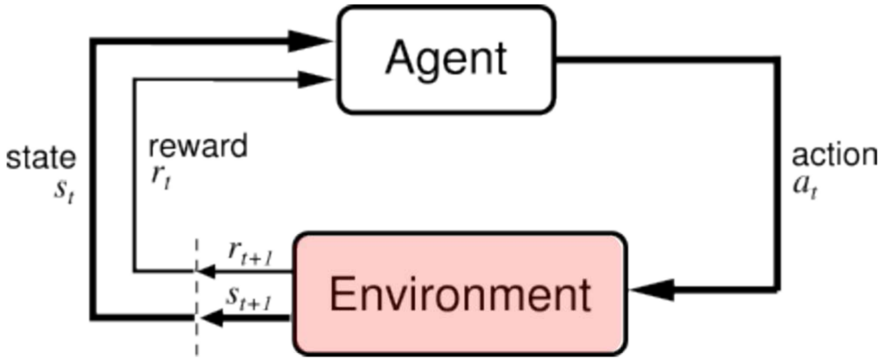
  Phase-C SetPoint [20, 70] deg +/- 0.1 deg

  Phase-D SetPoint [150, 245] deg +/- 0.1 deg

  Power-C SetPoint [0, 35] MW +/- 1 MW

  Power-D SetPoint [0, 55] MW +/- 1 MW

- 1 Odoscope [400, 650] MeV +/- 3 MeV



Markov Decision Process (MDP)

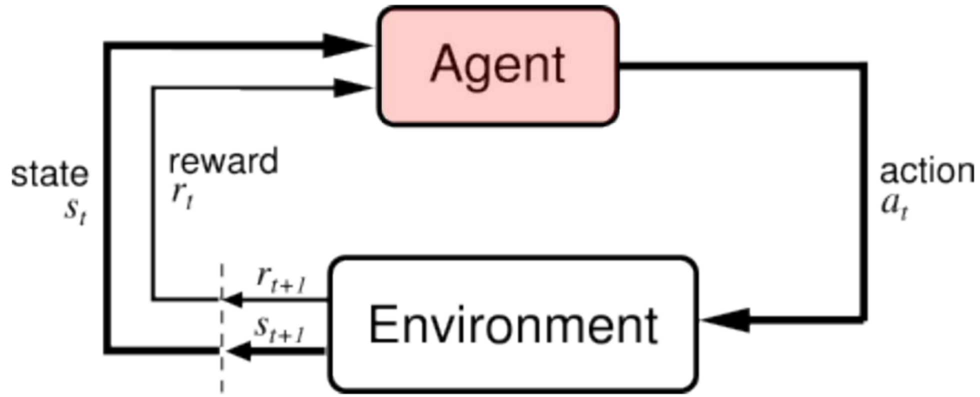
Dafne LINAC beam energy tuning tool

Agent:

- Q-learning to estimate optimal policy Q* to maximize scoring:

$$max(energy_{target} - energy)$$

- Exploration vs Exploitation balancing through decaying ε-greedy strategy.



Markov Decision Process (MDP)

Q* (State, Action)	ph-c,Pot-c, ph-d,Pot-d, …	ph-c,Pot-c, ph-d,Pot-d, …	ph-c,Pot-c, ph-d,Pot-d, …
energy			
…			
Energy working point			goal

Huge Dimension !!!

Bellman transition function:

$$\boldsymbol{Q^{new}(s_t, a_t)} \leftarrow (1-\alpha)\cdot Q(s,a) + \alpha\left(r + \gamma \cdot max_a Q(s_{t+1}, a)\right)$$

$$\alpha = learning\ rate$$

$$\gamma = discount\ factor$$

Dafne LINAC beam energy tuning tool

Training on 300 simulated episodes:

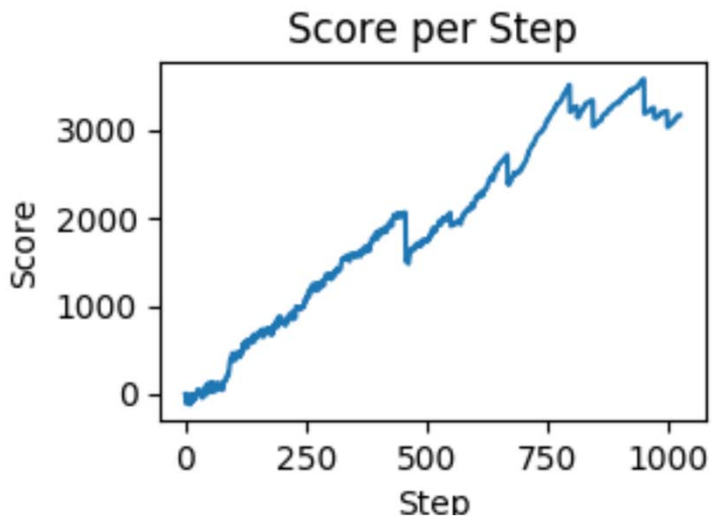
- Instance linac lattice

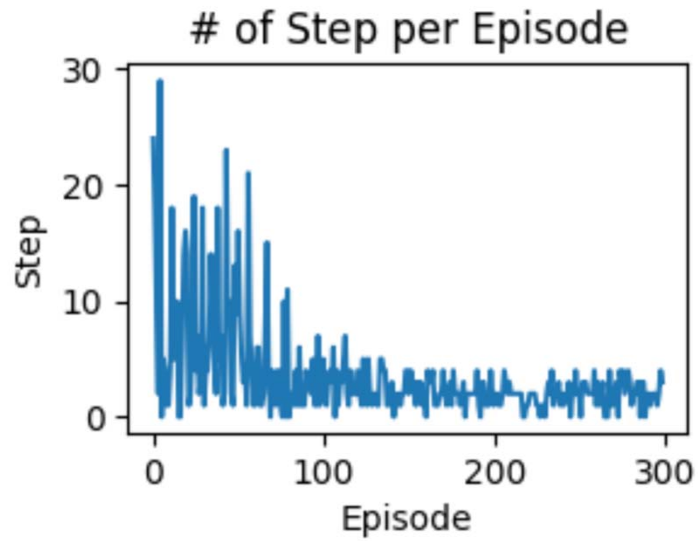
1. generate random [phase, power] setpoint
2. Move RF sources updating $Q(s,a) \rightarrow Q^*$
3. Final Goal:

   $energy_{target} - energy < \ energy_{tollerance}$

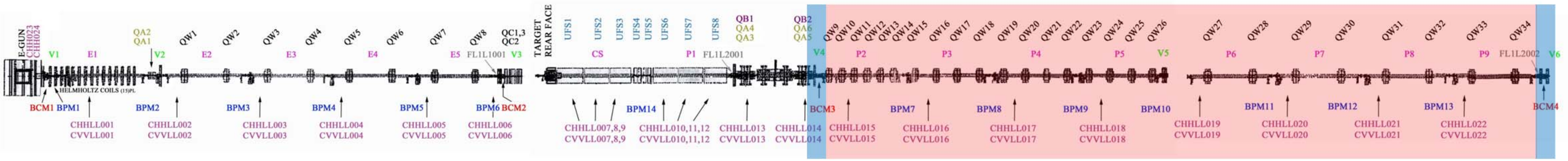
Results:

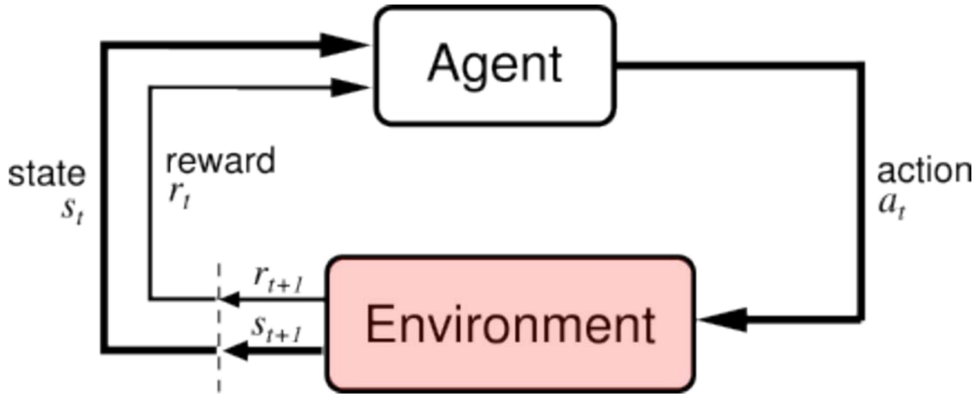
- RL algorithm tested with simulated data:
  - AI trained in 1ksteps (expected 1 week of beam shift).
  - Capable of training on moltiple working point in parallel.
  - Machine, misalignment and performance drift indipendent.





Dafne LINAC beam charge optimization tool





Markov Decision Process (MDP)

Automate accelerator beam charge optimization.

Environment:

- 8 Quadrupole Magnets

  Current PSU SetPoint [0, 10] A +/- 0.1 A

- 2 BCMs [0, 1.2] mA +/- 1 uA

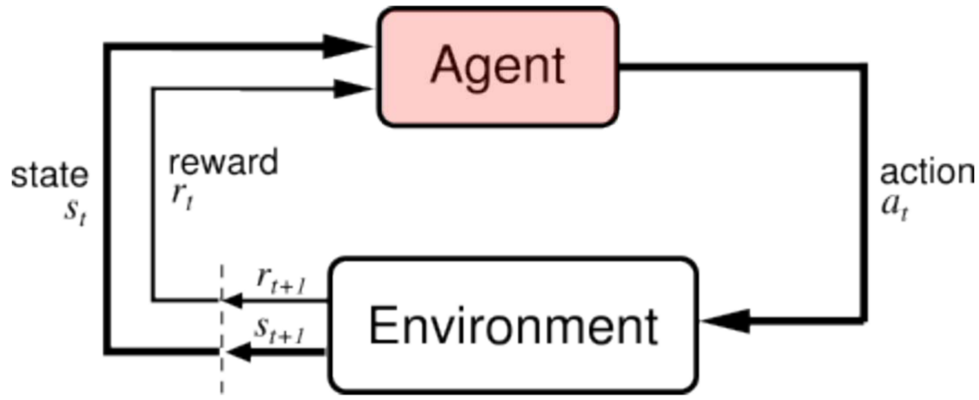
Dafne LINAC beam charge optimization tool

Agent:

- Q-learning to estimate optimal policy Q* to maximize scoring:

  $$max\left(charge\ ratio_{target} - charge\ ratio\right)$$

- Exploration vs Exploitation balancing through decaying ε-greedy strategy.



Markov Decision Process (MDP)

Q* (State, Action)	Q1,Q2,Q3…	Q1,Q2,Q3…	Q1,Q2,Q3…
Charge ratio			
…			
Optimal charge ratio			goal

Huge Dimension !!!

Bellman transition function:

$$\boldsymbol{Q^{new}(s_t, a_t)} \leftarrow (1-\alpha)\cdot Q(s,a) + \alpha\left(r + \gamma \cdot max_a Q(s_{t+1}, a)\right)$$

$$\alpha = learning\ rate$$

$$\gamma = discount\ factor$$

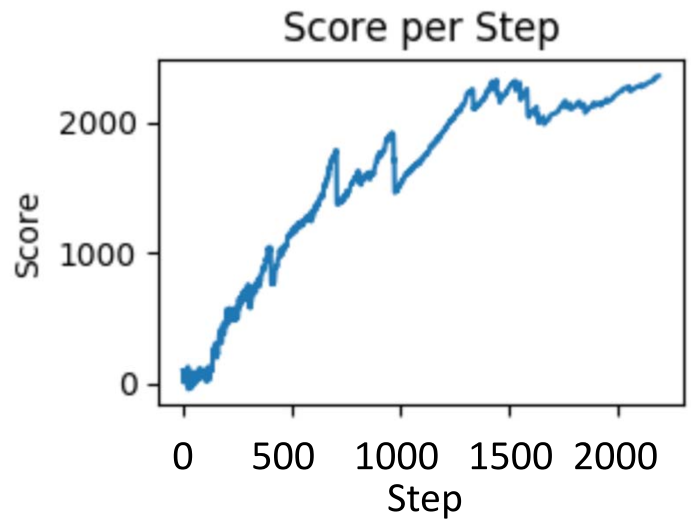
Dafne LINAC beam charge optimization tool

Training on 300 simulated episodes:

- Instance linac lattice

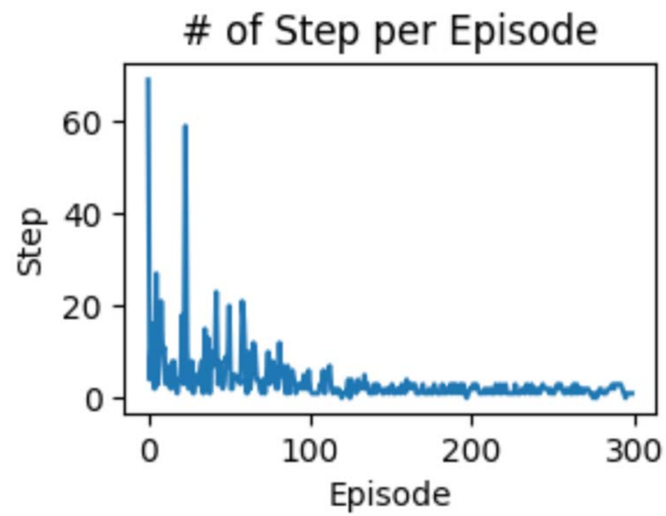
1. generate random quads setpoint
2. Move quads updating $Q(s,a) \rightarrow Q^*$
3. Final Goal:

   $charge\,ratio_{target} - charge\,ratio < \; charge\,ratio_{tollerance}$



Results:

- RL algorithm tested with simulated data:
  - AI trained in 2ksteps (expected 2 weeks of beam shift).
  - Machine, misalignment and performance drift indipendent.



Conclusions and next steps

- Reinforcement Learning tools validated as suitable for autonomous operation on accelerator facilities in operation, in commissioning or in old complex.
- RL tools developed to be easily configured on different lattice with safe operation limit.
- Training period reasonable to schedule dedicated beam shifts.
- First test shift at Dafne LINAC scheduled for next weeks.
- Upgrade to Deep Q-Networks should be done to improve memory managment.

Thank you for the attention