

Randomly select action
and observe response

History buffer full

Train Explore, Exploit
networks



Update ϵ



Choose random
number x from $U(0,1)$



Is $x > \epsilon$?

Yes,
Exploit

Feed last observed
evaluation metrics into
Exploit NN

No,
Explore

Feed every possible
action into Explore NN

Returns expected fitness
values for each action

Choose action where
 $\text{fit_pred} \geq \text{threshold}$

Guided

Choose whether
guided or unguided
explore

Unguided

Choose action where
 $\text{fit_pred} < \text{threshold}$

Returns action
to be taken

Transmit Action to ISS



Add response to
history buffer

History buffer not full

History buffer full

