

Create the Initial Solution
state = initial state condition
it = 0

WHILE
it < n or $f(\text{solution}) \neq LB$

No

Returns the *best solution* found

Yes

PPO chooses a Neighborhood Search Heuristics for current *state*.

Runs Neighborhood Search Heuristics chosen by the PPO, update *state* and, to apply **Action Masking**.

IF
 $f(\text{solution}) < f(\text{best_solution})$

No

Run perturbation of solution.

Yes

Update *best solution* found.
best solution = solution

Increment the Iteration variable.
it = it + 1

