Step 1:

Taking Hidden State Position Step 2:

Decide to Patch by SVM Classifier

Step 3:

Patching

Step 4:

Running Execution on Patched Target

