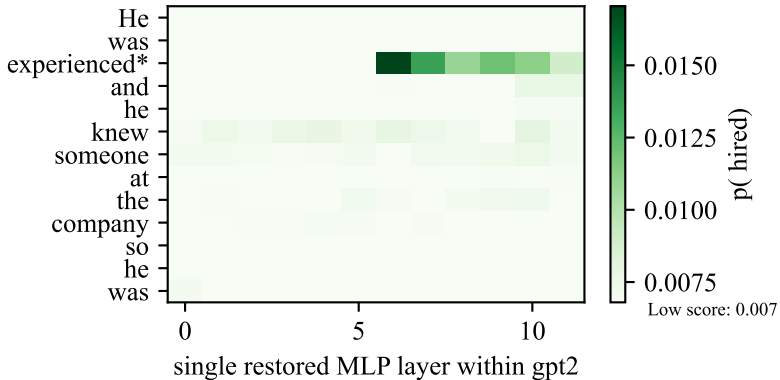Impact of restoring MLP after corrupted input

single restored MLP layer within gpt2

p( hired )

Low score: 0.007

High score: 0.002