DeepLocker – I malware basati su Intelligenza Artificiale sono già realtà

E se l'intelligenza artificiale fosse utilizzata per scopi offensivi? La simulazione

Ogni qual volta si discute di **Intelligenza Artificiale** applicata al settore cyber security di ipotizza la sua introduzione per scopi difensivi. Sistemi basati su AI posso aiutare a identificare tempestivamente una minaccia ed elaborare una risposta per la sua mitigazione in realtime

> *E se l'intelligenza artificiale fosse utilizzata per scopi offensivi?*

Non è fantascienza, sistemi di questo tipo potrebbero coadiuvare un essere umano nelle fasi di attacco scalando in questo modo il livello di complessità dell'offensiva.

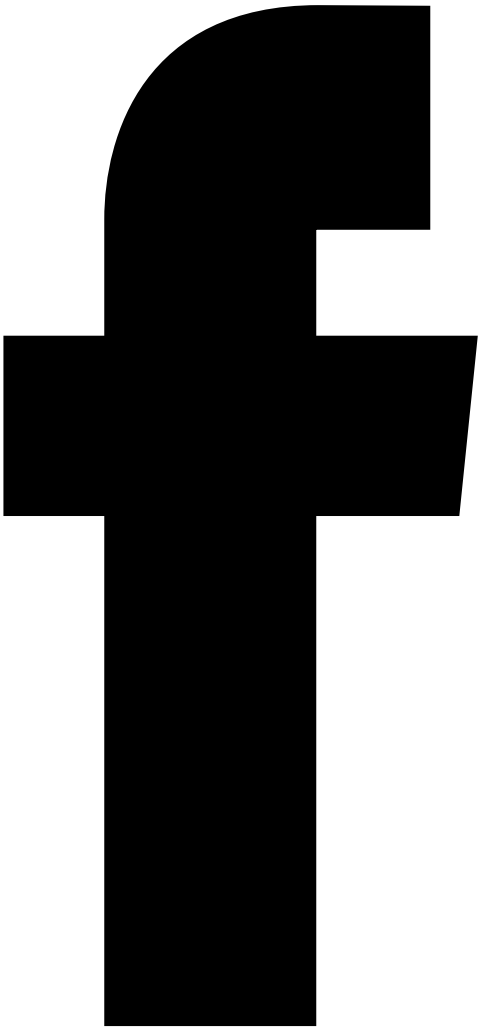
Pensiamo ad esempio al caso dell'adozione di AI nel settore dello sviluppo malware.

Una tipologia di malware che è già tra noi

Gli attaccanti potrebbero usare un malware equipaggiato con un motore di AI in grado di eludere difese sofisticate attivandosi solo in presenza di un particolare obiettivo ed al verificarsi di determinate condizioni.
Se pensate che siano discorsi da rimandare ad un futuro prossimo vi sbagliate, questa tipologia di malware è già tra noi.
I ricercatori di sicurezza del gruppo **IBM Research** hanno sviluppato uno strumento di attacco "altamente mirato ed evasivo" basato su tecnologia AI, "soprannominato DeepLocker che è in grado di nascondere il suo intento malevolo fino a quando non ha identificato il target specifico."
"IBM Research ha sviluppato DeepLocker per capire meglio come diversi modelli di AI esistenti possono essere combinati con le attuali tecniche di sviluppo malware per creare una nuova generazione di codici malevoli particolarmente attivi". Si legge in un post pubblicato dagli esperti.
"Questa classe di malware evasivi basati su AI sono in grado di nascondere la loro finalità fino a quando non raggiungono una vittima specifica. Attivano il loro comportamento dannoso non appena il modello di intelligenza artificiale identifica il bersaglio attraverso indicatori come riconoscimento facciale, geolocalizzazione e riconoscimento vocale".
Secondo i ricercatori di IBM, DeepLocker è in grado di evitare il rilevamento e di attivarsi solo dopo aver soddisfatto un insieme di condizioni specifiche.
Il malware attivato da AI rappresenta una opzione privilegiata in attacchi mirati come quelli condotti da attori nation-state.
Il codice dannoso **potrebbe essere nascosto in applicazioni apparentemente innocue** e selezionare il target in base a vari

indicatori quali riconoscimento vocale, riconoscimento facciale, geolocalizzazione e altre funzionalità a livello di sistema.
"DeepLocker nasconde il suo codice malevolo in applicazioni legittime ed apparentemente innocue, come un software per videoconferenze, per evitare il rilevamento da parte della maggior parte degli scanner antivirus e malware". Continua IBM.
"Ciò che rende unico DeepLocker è che l'utilizzo dell'AI che rende quasi impossibile l'individuazione delle condizioni di innesco per sferrare l'attacco. Il codice malevolo verrà attivato solo se viene raggiunto il target previsto. Tutto ciò è possibile attraverso l'implementazione di un modello di rete neurale profonda (DNN)".

Una simulazione



I ricercatori hanno presentato alla scorsa edizione della conferenza di hacking Black Hat una simulazione di come sia possibile condurre un attacco con un malware basato su AI. Gli esperti hanno nascosto il codice del temuto ransomware WannaCry in un'applicazione di videoconferenza e sono riusciti a mantenerlo invisibile alle soluzioni di sicurezza fino al raggiungimento della vittima predestinata identificata attraverso il riconoscimento facciale. Gli esperti hanno sottolineato che il bersaglio può essere identificato anche utilizzando un riconoscimento attraverso foto disponibili pubblicamente.

"Per dimostrare le implicazioni delle funzionalità di

DeepLocker, abbiamo progettato un PoC in cui camuffiamo un noto ransomware (WannaCry) in un'applicazione di videoconferenza benigna, in modo che non venga rilevata dagli strumenti di analisi del malware, inclusi motori antivirus e sandbox malware. Come condizione di attivazione, abbiamo addestrato il modello di intelligenza artificiale a riconoscere il volto di una persona specifica per sbloccare il ransomware ed eseguirlo sul sistema".
"Immagina che questa applicazione di videoconferenza sia distribuita e scaricata da milioni di persone, il che è uno scenario plausibile al giorno d'oggi su molte piattaforme pubbliche. Una volta avviata, l'applicazione analizzerebbe le istantanee catturate della fotocamera grazie al modello di intelligenza artificiale incorporato, ma si comporterebbe diversamente per tutti gli utenti tranne che per l'obiettivo prestabilito ", hanno aggiunto i ricercatori.
"Quando la vittima si siede davanti al computer e utilizza l'applicazione, la telecamera rileva il suo volto ed il codice malevolo sarà attivato di nascosto grazie al riconoscimento facciale della vittima, che era la condizione di innesco programmata."
Quanto descritto **è solo una piccola parte delle capacità che un sistema di AI può fornire ad un codice malevolo**, basti pensare che ulteriori evoluzioni potrebbero vedere malware in grado di rispondere ai vari tentativi di difesa del sistema obiettivo ed in grado di mutare comportamento in relazione al sistema ospite.
È chiaro che i modelli di attacco e difesa saranno profondamente influenzati da tecnologie basate su intelligenza artificiale, sistemi che sono già tra noi.