NOVA24 Il Sole 24 Ore 06 LUGLIO 2023

COME DIFENDERCI GUIDA

Così i criminali usano ChatGpt per truffarci

Gianni Rusconi

Cynet. Secondo l'azienda israeliana bastano poche decine di dollari per ottenere malware malevoli, scaricarli all'interno di canali Telegram, pagando con PayPal

Sono ormai passati alcuni mesi dal suo lancio ma la discussione è quanto mai "calda": ChatGpt e le tecnologie di intelligenza artificiale generativa in genere portano più benefici o rischi, più vantaggi o criticità? Non c'è solo il tema del lavoro e delle professioni al centro del dibattito ma anche questioni altrettanto critiche come la sicurezza dei dati macinati dagli algoritmi e la protezione della privacy degli utenti. I cybercriminali, e questa è cronaca recente, hanno già approfittato del fenomeno per sperimentare e ingegnerizzare nuove modalità di attacco, con buona pace di OpenAi e del suo dichiarato impegno a garantire un uso adeguato della propria tecnologia. Le capacità dell'Ai generativa, insomma, possono essere utilizzate anche in maniera negativa e si aprono di conseguenza preoccupanti scenari anche per le aziende, chiamate a bilanciare i tanti potenziali benefici in termini di maggiore produttività derivanti dall'adozione di questi strumenti con la necessità di affinare i propri sistemi di protezione (e di prevenzione) per evitare possibili ingenti danni imputabili alla vulnerabilità dei propri sistemi e ai comportamenti poco virtuosi dei propri dipendenti.

Ma da dove arrivano le principali minacce legate a ChatGpt e simili? L'azienda israeliana di cybersecurity Cynet ha rilevato un aumento nella creazione di e-mail di phishing (da cui scaturisce poi l'installazione di ransomware) con modelli di

Generative Pre-trained Transformer in grado di generare attacchi più sofisticati e su larga scala, stimando come la percentuale di vittime di questi attacchi sia fino a cinque volte superiore rispetto alla norma. Tra i principali vettori di phishing sono stati identificati i cosiddetti “Infostealer”, e cioè una specifica tipologia di malware silenti di tipo Trojan, difficilmente intercettabili e in grado di rubare informazioni personali alle vittime quali utente e password, sessioni di navigazione, dati bancari, cookies e applicazioni di messaggistica. La peculiarità che li contraddistingue? Stando alle evidenze riscontrate da Cynet, nonostante questi malware abbiano una forma sempre più ricercata, la facilità con cui è possibile acquistarli è estrema: bastano poche decine di dollari per scaricarli all'interno di canali Telegram, pagando con PayPal e senza la necessità di accedere al darkweb.

In linea generale, la convinzione degli esperti della società israeliana è che i cybercriminali stiano iniziando a sfruttare in serie le potenzialità di soluzioni come ChatGpt 4 per creare testi molto accurati in grado di produrre un ritorno dell'investimento diverse volte superiore alle precedenti tecniche di attacco. La strategia di “go-to-market” del cybercrime non è destinata a cambiare repentinamente ma con i sistemi di Ai generativa è palese che la sofisticatezza (in termini di precisione linguistica e di integrazione di codice maligno) degli attacchi possa sensibilmente e ulteriormente evolvere. I modelli Gpt, questa l'essenza dell'analisi di Cynet, non sono in grado di rivoluzionare il panorama delle minacce ma offrono più velocità e automazione ai criminali della Rete e una maggiore accessibilità anche a persone prive di competenze tecniche. Se, in altre parole, prima era solitamente un tecnico a ideare e lanciare l'attacco, ora questa possibilità è a disposizione di chi escogita la truffa e di tutti coloro dotati di talento criminale. E si tratta, facile intuirlo, di un cambio di paradigma non da poco e dagli effetti ancora assolutamente imprevedibili. Tema caldo, quindi, sul quale si esprimono anche gli specialisti della cybersicurezza made in Italy come CybergON, business unit di Elmec Informatica, secondo cui le principali tecniche di attacco condotte tramite AI generativa sono il phishing, l'installazione di malware e la creazione di campagne DeepFake a scopo di frode e furto di identità. Dalla start up romana Cyber Guru, invece, arriva un suggerimento più organico per difendersi dai cyberattacchi architettati da bot di intelligenza artificiale di nuova generazione come ChatGpt, un suggerimento che invita le imprese a un cambiamento radicale della cultura aziendale. Occorre cioè lavorare sull'organizzazione e sulle persone per ostacolare i cybercriminali nel trarre vantaggio dai bias cognitivi utilizzati per architettare truffe basate su dinamiche di gerarchia all'interno di un ambiente di lavoro. Agendo sul fattore umano, che ricordiamo essere l'agente responsabile di circa l'80% delle violazioni registrate a livello globale nel 2021, è dunque il primo passo per creare una linea di difesa più efficace.