I *deepfake detectors*: un'opportunità di investimento

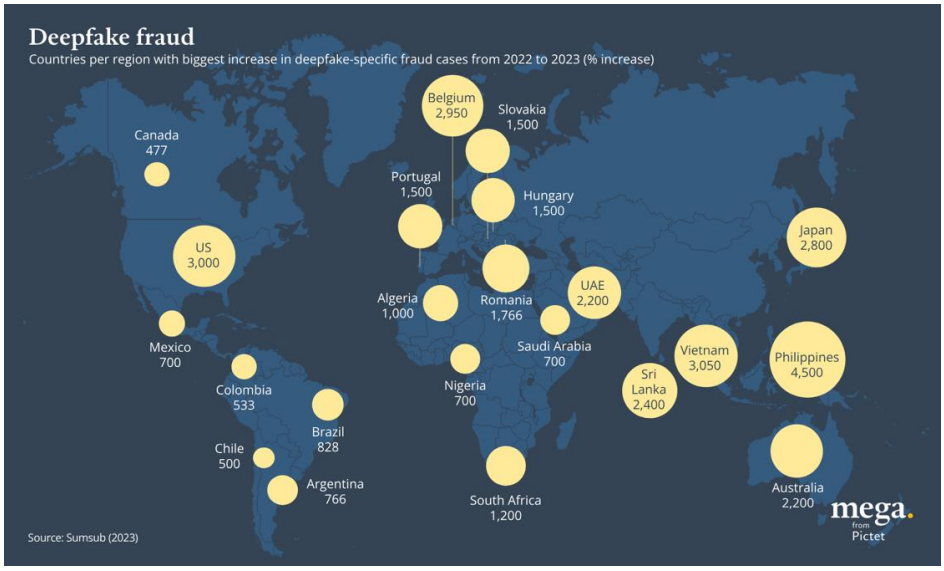
A cura di **Yves Kramer**, Senior Investment Manager, Thematic Equities di Pictet Asset Management

27.05.2024

- *I deepfake rappresentano oggi una sfida importante per il settore della sicurezza informatica, ma anche una grande opportunità. Le aziende di cybersicurezza in grado di adottare l'IA nella progettazione di sistemi specializzati di difesa digitale avranno forti prospettive di crescita nei prossimi anni.*
- *Per contrastare la crescente minaccia, gli investimenti in sicurezza informatica stanno aumentando considerevolmente, con un focus su soluzioni come Zero Trust e Secure Access Service Edge (SASE), adattate per contrastare le minacce dell'IA generativa.*

Il 2024 passerà alla storia come il più grande anno elettorale di sempre, con ben quattro miliardi di persone chiamate a votare. Purtroppo, questa serie di appuntamenti decisivi coincide con l'avvento dei deepfake generati dall'IA, che potrebbero diffondere disinformazione per influenzare gli elettori. Il *deepfake* è una tecnica di sintesi dell'immagine che riproduce le sembianze di una persona attraverso l'uso di Intelligenza Artificiale, in particolare del "deep learning". I deepfake vengono creati addestrando un modello di apprendimento automatico su un ampio set di immagini o video di una persona, per esempio un candidato politico, per imparare e replicarne le sue peculiarità. In totale, in un solo anno, tra il 2019 e il 2020, la quantità di contenuti online deepfake è aumentata del 900%[1].Entro il 2026 la maggioranza dei contenuti online potrebbe essere generata in modo artificiale, rendendo ancora più difficoltoso distinguere i contenuti autentici da quelli fraudolenti generati con il supporto dell'AI. Più dati sono disponibili, infatti, più il falso è realistico e quindi la minaccia cresce.

Fig.1 Frodi tramite deepfake
Paesi per regione con il maggior aumento di casi di frode tramite deepfake dal 2022 al 2023, in %



Fonte: Sumsub, 2023.

[1] "Deepfakes 2020: the tipping point".

Investire in cybersicurezza

La vulnerabilità dell'IA, ma soprattutto dell'IA generativa, nei confronti di cyberattacchi, violazioni e manipolazioni negative di dati sta alimentando la domanda di misure di prevenzione sempre più avanzate. Le autorità politiche statunitensi, per esempio, hanno riconosciuto la necessità di implementare investimenti in sicurezza informatica avanzata; un'azione avviata già nel 2020 quando sono stati stanziati 400 milioni di dollari per la sicurezza elettorale. Secondo Gartner, una società americana di ricerca e consulenza tecnologica, nel 2024 la spesa per la sicurezza informatica dovrebbe aumentare di circa il 14%. Con la crescente diffusione dell'intelligenza artificiale, riteniamo che questa crescita potrebbe essere ancora più rapida in futuro. Alla luce di ciò, è probabile che assisteremo ad un aumento delle soluzioni "Zero Trust", che controllano l'accesso degli utenti verificando continuamente le credenziali delle persone che interagiscono con un'organizzazione, sia internamente che esternamente. Prevediamo inoltre una crescita del Secure Access Service Edge (SASE), una trasformazione della rete progettata per il cloud, che sfrutta il monitoraggio dell'identità e del comportamento degli utenti per guidare le continue modifiche dei criteri.

È interessante notare che la stessa intelligenza artificiale generativa può essere parte della soluzione, in quanto il settore della sicurezza informatica adatta modelli linguistici di grandi dimensioni (LLM) per rilevare più rapidamente gli attacchi e contrastare le potenziali minacce provenienti da un codice pericoloso generato da altre macchine, rappresentando un'opportunità enorme, sia per il settore della sicurezza che per gli investitori. Le aziende di cybersicurezza in grado di adottare l'IA nella progettazione di sistemi di difesa digitale, al centro della nostra strategia Pictet-Security, avranno forti prospettive di crescita nei prossimi anni. Questo vale soprattutto per quelle realtà che danno priorità agli investimenti in infrastrutture e per quelle che sviluppano software e applicazioni di sicurezza specializzate.

Difesa dai deepfake

In questa particolare fase storica, i protagonisti dei deepfake restano comunque un passo avanti, poiché la quantità di risorse dedicate allo sviluppo di nuove forme di contenuti generativi è significativamente superiore a quella investita nello sviluppo di tecniche di rilevamento avanzate. Un'altra limitazione da non trascurare è che questi strumenti sono stati addestrati su dati di provenienza occidentale. Molti strumenti di rilevamento dei deepfake ad oggi mirano ad individuare piccoli difetti in audio e video per identificarli come contenuti manipolati. Mentre una delle tecniche più efficaci finora impiegate adotta l'approccio opposto, ricavando qualità uniche dei filmati reali che i deepfake non possono cogliere. FakeCatcher di Intel, ad esempio, è uno strumento all'avanguardia che identifica le immagini false di persone utilizzando la foto-pletismografia (PPG), una tecnica che rileva i cambiamenti del flusso sanguigno nel viso. Un'altra iniziativa promettente è quella proposta dalla Coalition for Content Provenance and Authenticity (C2PA): utilizzando una firma digitale crittografata si potrebbe garantire maggiore trasparenza circa le modalità di creazione di un contenuto multimediale e gli strumenti utilizzati. Se da un lato questo approccio presenta dei vantaggi, dall'altro potrebbe sollevare questioni etiche relative alla privacy, con il timore che lo standard imposto possa rivelare troppi dati sulla provenienza dell'immagine, come il luogo o l'ora in cui è stata scattata.

La tecnologia da sola non potrà mai risolvere completamente il problema dei deepfake, ma è necessario che i consumatori imparino ad osservare e valutare in maniera più critica i contenuti che consultano, invece di fidarsi di tutto ciò che si vede senza verificarne origine e credibilità.

Dubai, Francoforte, Hong Kong, Londra, Losanna, Lussemburgo, Madrid, Milano, Montreal, Monaco di Baviera, Nassau, New York, Osaka, Parigi, Principato di Monaco, Roma, Shanghai, Singapore, Stoccarda, Taipei, Tel Aviv, Tokyo, Torino, Verona e Zurigo. Pictet Asset Management ("Pictet AM") comprende tutte le controllate e le divisioni del Gruppo Pictet che svolgono attività di asset management e gestione fondi istituzionali. Fra i principali clienti si annoverano alcuni dei maggiori fondi pensione, fondi sovrani e istituti finanziari a livello mondiale.

Contatti Stampa:
BC Communication
Lucrezia Pisani | Tel. +39 347 6732479 | lucrezia.pisani@bc-communication.it
Carla Parisi | Tel. +39 339 5796751 | carla.parisi@bc-communication.it
Chiara Cattaneo | Tel. +39 340 9597461 | chiara.cattaneo@bc-communication.it