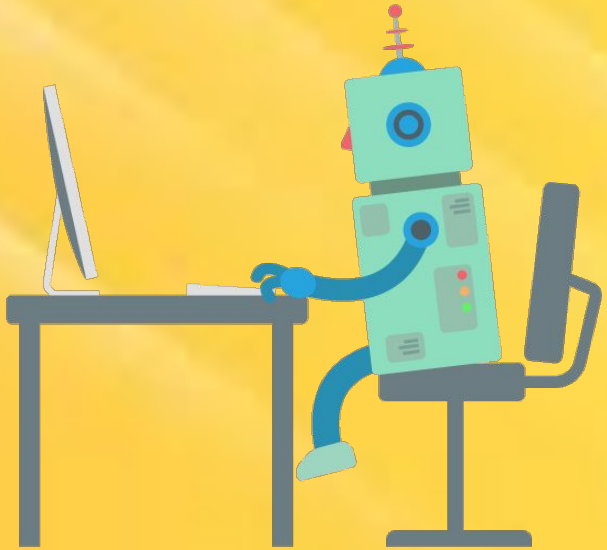
L'INTELLIGENZA ARTIFICIALE NELLA SICUREZZA

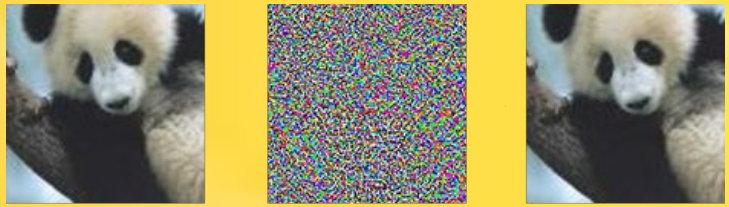
L'intelligenza artificiale, in particolare il deep learning, viene utilizzata in molte applicazioni nel campo della sicurezza.
Alcune di queste sono:

- Rilevamento malware
- Analisi forense di risorse multimediali
- Autenticazione su base biometrica
- Analisi di traffico dati
- Rilevamento di intrusioni in una rete
- Rilevamento di attacchi Denial of service
- Monitoraggio di sistemi fisici (per esempio i metal detector negli aeroporti)
- Steganalisi
- Sistemi di video sorveglianza



LA SICUREZZA DELL'INTELLIGENZA ARTIFICIALE

Solo recentemente la ricerca ha iniziato a dare attenzione alla sicurezza dell'intelligenza artificiale. Ingannare un sistema basato su machine learning, infatti, è possibile.
L'attacco si basa sulla creazione dei cosiddetti adversarial examples, degli esempi disegnati appositamente per indurre in errore la rete neurale. Sono un po' l'equivalente per l'intelligenza artificiale di quello che per gli umani sono le illusioni ottiche.



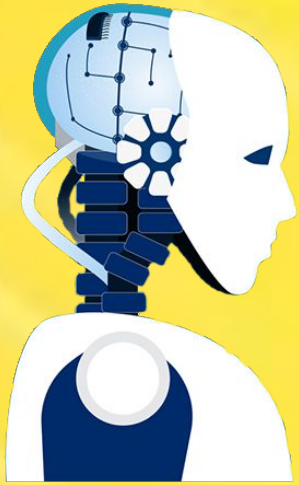
panda scimmia

Introducendo un rumore opportunamente calcolato, il sistema viene indotto in errore. Nonostante le due immagini per un umano siano indistinguibili l'una dall'altra, quella originaria viene classificata correttamente, mentre quella contenente il rumore no.

RIFERIMENTI

- https://www.c3t.it/projects/awareness/articoli&brochure/ai/

POR FESR 2014-2020 obiettivo Crescita e Occupazione (CreO)



INTELLIGENZA ARTIFICIALE E SICUREZZA

Scopri cos'è l'intelligenza artificiale e come è utilizzata nell'ambito della sicurezza

CHE COS' E' L'INTELLIGENZA ARTIFICIALE?

Con l'espressione "Intelligenza Artificiale" si intende quella branca dell'informatica che si occupa di costruire sistemi hardware o software in grado di svolgere compiti considerati "intelligenti", come ad esempio la percezione visiva, l'elaborazione del linguaggio, l'analisi di decisioni.
Oggi l'Intelligenza Artificiale viene abbondantemente utilizzata nella vita quotidiana. Ad esempio, i vari strumenti di riconoscimento vocale che vengono regolarmente utilizzati, dagli smartphone ai sistemi di sicurezza, si basano su algoritmi tipici dell'Intelligenza Artificiale.

IL MACHINE LEARNING

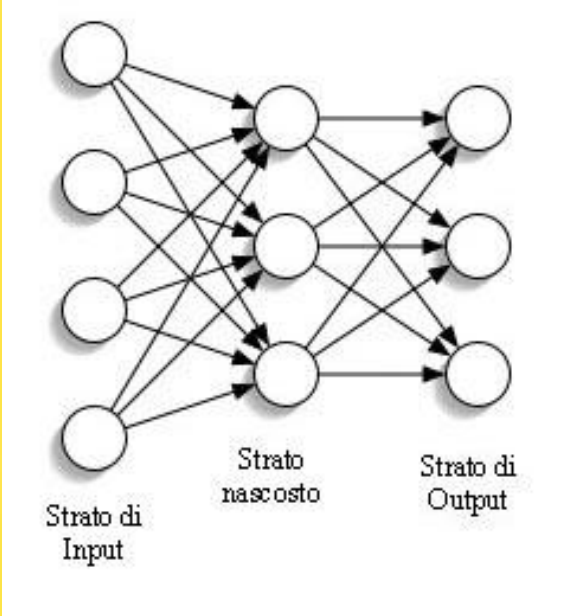
Negli ultimi anni l'intelligenza artificiale ha visto un notevole miglioramento grazie allo sviluppo dell'apprendimento automatico (machine learning). A differenza degli algoritmi utilizzati in precedenza, l'apprendimento automatico, come suggerisce la parola stessa, si basa sul fatto che la macchina impara autonomamente un compito dopo aver osservato degli esempi che le vengono forniti. Questo processo di apprendimento avviene utilizzando le **reti neurali.**

LE RETI NEURALI

Una rete neurale è un modello computazionale composto da neuroni collegati tra loro da archi. I neuroni sono divisi in strati. Nello strato di ingresso viene inserito l'input, negli strati nascosti l'input viene elaborato per poi fornire l'output nello strato di uscita. Ad ogni neurone, inoltre, viene associata una funzione detta funzione di attivazione. Ad ogni arco, invece, viene associato un valore reale detto peso. Il valore di un neurone (esclusi quelli dello strato iniziale che sono dati) viene calcolato nel seguente modo:

- i singoli ingressi del neurone vengano moltiplicati per il peso corrispondente all'arco che li collega. I risultati di queste moltiplicazioni vengono poi sommati
- Si applica la funzione di attivazione al risultato ottenuto al punto precedente

Reti neurali con un grande numero di strati vengono chiamate reti neurali profonde (**deep neural networks**)



IL PROCESSO DI APPRENDIMENTO

Il processo di apprendimento di una rete neurale consiste nel trovare i valori dei pesi ottimali. Ad esempio, nel caso di classificazione di immagini, vogliamo che dati in ingresso i valori dei pixel di una immagine che rappresenta un certo oggetto, vogliamo che l'output della rete sia un codice che rappresenti quell'oggetto.
L'algoritmo su cui si basa il processo di apprendimento viene detto di backpropagation (propagazione all'indietro). Vengono forniti alla rete degli esempi (nel nostro caso, immagine e codici rappresentanti oggetti). La rete, osservando l'errore che commette, **"aggiusta" i valori dei pesi** per ridurre il più possibile l'errore totale sugli esempi forniti. Se l'apprendimento va a buon fine, alla fine la rete sarà in grado di classificare correttamente non solo le immagini che le erano state fornite per "allenarsi", **ma anche immagini nuove che non aveva mai visto!**