



L'integrazione di gateway e di sistemi edge estende la portata dell'AI

Grazie alla capacità di segmentare e classificare dati complessi, la tecnologia dell'AI sta diventando la soluzione vincente per la prossima generazione di sistemi di manutenzione predittiva.

di Cliff Ortmeyer

L'uso dell'apprendimento automatico e dell'intelligenza artificiale (AI) nel controllo industriale e in applicazioni simili è aumentato notevolmente negli ultimi anni.

Ovviamente, ciò ha portato sfide e imposto vincoli all'impiego dell'apprendimento automatico, come le prestazioni in tempo reale e la gestione dei dati commerciali sensibili. Tuttavia, grazie alla capacità di segmentare e classificare dati complessi, la tecnologia dell'AI sta diventando la soluzione vincente per la prossima generazione di sistemi di monitoraggio delle condizioni di funzionamento dei macchinari e per l'ottimizzazione dei processi.



AI E MANUTENZIONE PREDITTIVA

L'utilizzo dell'AI nella manutenzione predittiva si è rivelata estremamente vantaggiosa e può offrire un elevato profitto sugli investimenti; infatti, il monitoraggio dello stato dei macchinari non solo riduce il numero di ispezioni necessarie, ma anche i costi di manutenzione. Ovviamente, è importante evitare i tempi di inattività, in quanto la misurazione precisa della durata di un componente specifico fa sì che quest'ultimo possa essere sostituito in un momento di bassa produttività per eliminare qualsiasi possibilità di guasto durante il processo di lavorazione. Invece di affrontare i problemi solo in seguito al verificarsi di un guasto, il che comporterebbe inevitabilmente costosi tempi di inattività e richiederebbe anche un team di risposta rapida costoso da implementare, gli utenti della manutenzione predittiva sono noti per aumentare l'efficienza produttiva fino al 25 per cento.

Uno dei principali vantaggi dell'utilizzo dell'AI in applicazioni come la manutenzione predittiva consiste nel riscontrare pattern ricorrenti nei dati che potrebbero essere ignorati da algoritmi deterministici. Ad esempio, le variazioni di temperatura e vibrazione, spesso normali nel corso di una lavorazione a macchina, potrebbero indicare un problema per certe combinazioni e profili di serie temporali.

I modelli basati sull'intelligenza artificiale possono analizzare grandi quantità di dati per le serie temporali contribuendo a una migliore comprensione delle situazioni che causano guasti a componenti specifici. Per molti anni, i dati prodotti dalle macchine utilizzate nelle fabbriche sono stati utilizzati per generare semplici segnalazioni, come avvisi nei casi in cui la lettura di un sensore di temperatura supera un livello prestabilito. Un'installazione IoT più complessa è dotata di numerosi sensori che monitorano vari aspetti operativi tra cui: temperatura, vibrazione, consumo di corrente e così via. Questo approccio basato su grandi quantità di dati prevede l'invio di tutti i dati al cloud per ottenere un maggiore livello di analisi, e una segnalazione dei dati storici che possono contribuire alla capacità decisionale automatizzata delle macchine in tempo reale.

Negli ultimi tempi, il sistema di controllo delle macchine di fabbrica si è spostato dai sistemi in siti locali a strutture remote e centralizzate basate sul cloud. Tuttavia, in futuro si prevede il trasferimento dell'elaborazione dei dati e di alcune altre funzioni cloud ai sistemi edge, in modo da risultare il più vicino possibile alle fonti dati. Ciò significa che, sulla base delle analisi in tempo reale delle prestazioni di una macchina e di una cronologia aggregata delle prestazioni su un gran numero di macchinari (anche in diverse fabbriche), un sistema di intelligenza artificiale può identificare rapidamente i modelli di comportamento all'origine di possibili guasti.

I sistemi di manutenzione predittiva odierna possono prevedere quando un componente subirà guasti, se non

viene sottoposto a manutenzione, e suggerire provvedimenti per prevenirli. Con un sistema basato sull'intelligenza artificiale, sarà anche possibile identificare come utilizzare le macchine per massimizzarne la durata, offrendo un compromesso tra prestazioni e longevità. Gli algoritmi di apprendimento automatico rendono più intelligente il sistema di analisi con il passare del tempo e sono disponibili più serie di dati e modelli. Ovviamente, la qualità di un sistema basato sull'intelligenza artificiale dipende dai dati ricevuti, quindi maggiore sarà la quantità di dati immessi nel sistema sul funzionamento del dispositivo, guasti e manutenzione, più accurato diventerà il sistema di analisi preventiva.

CONSERVAZIONE LOCALE DEI DATI AI O SUL CLOUD

Con il termine 'inferenza' si intende il processo di applicazione di nuovi input a un modello addestrato, in modo che questo modello possa determinare il significato degli input. Esiste ancora una forte domanda per l'elaborazione locale dei modelli di intelligenza artificiale, in particolare per l'inferenza, anche perché l'AI sul dispositivo fornisce la minore latenza di comunicazione e la larghezza di banda delle comunicazioni potrebbe semplicemente essere insufficiente a trasmettere dati sufficienti al cloud per avere un'inferenza attendibile. Esistono due approcci per rendere possibile l'elaborazione locale. Uno consiste nell'utilizzare la potenza di calcolo disponibile sul dispositivo stesso, anche se potrebbe non esserci sufficiente capacità residua di elaborazione per eseguire un modello sufficientemente complesso. La seconda opzione consiste nell'utilizzare un altro dispositivo, in parte o per intero, per l'elaborazione. Ad esempio, il dispositivo esegue un modello AI semplificato che svolge un'analisi iniziale dei dati, ma un gateway o un computer a livello di scheda nelle vicinanze si occupa di eseguire un modello più adatto. Il gateway locale può persino aggiornare il modello eseguendo da sé l'addestramento anziché trasferirlo su un server cloud. Anche se spesso è logico scaricare i processi di calcolo più complessi relativi all'addestramento dei modelli sui potenti server cloud, per gli utenti industriali questa esecuzione remota può rappresentare un problema per diversi motivi, tra cui la tempestività e la riservatezza dei dati operativi sensibili. Per molti sistemi di intelligenza artificiale basati sul cloud, che possono eseguire servizi come la traduzione di testi da una lingua all'altra, sono accettabili latenze variabili e persino lunghe; tuttavia, ciò non vale per i sistemi industriali in cui l'AI è utilizzata per il controllo operativo.

QUANDO LA POTENZA DEL CLOUD NON È SEMPRE SUFFICIENTE

Diverse aziende decidono di ospitare quantità significative di dati nel cloud pubblico per la disponibilità di Amazon Web Services (AWS), Microsoft Azure, Google Cloud

Platform e una serie di altri fornitori di servizi cloud. Attualmente, la maggior parte dei dati provenienti dai dispositivi connessi nell'ecosistema IoT viene raccolta e trasmessa al cloud per essere sottoposta a elaborazione e analisi.

La potenza di calcolo nel data center del cloud si concentra laddove vengono aggregati i dati e prese decisioni basate sull'intelligenza artificiale. Sebbene questo approccio si sia rivelato affidabile, la quantità di tempo necessaria per completare il trasferimento dei dati da e verso il cloud induce a problemi di latenza, che possono influire sul processo decisionale in tempo reale. Maggiore è la distanza geografica rispetto al data center del cloud, maggiore è la latenza introdotta.

È stato dimostrato che per ogni 100 miglia percorse dai dati, si perdono circa 0,82 millisecondi di velocità. Il cloud computing può essere agile, ma non può sostenere la crescente domanda di carichi di lavoro ad alta intensità di dati, gestiti da applicazioni IoT utilizzate nell'assistenza sanitaria, nella produzione e nei trasporti. Un sistema di apprendimento automatico senza supervisione può riscontrare pattern nelle letture dei sensori e, invece di passare i dati grezzi dei sensori a un gateway o un server locale, i pattern rilevati possono essere inviati, riducendo notevolmente il carico delle comunicazioni. Un modello addestrato che utilizza tecniche di AI supervisionate può quindi analizzare i pattern e determinarne il significato.

AI SULL'EDGE

Il cloud computing rimarrà indubbiamente una parte importante dell'ecosistema IoT per l'elaborazione di dati complessi e storici, poiché il numero e la praticità delle soluzioni IoT basate su AI continua a crescere. Tuttavia, se l'applicazione deve alimentare il processo decisionale in tempo reale, l'edge computing è un approccio decisamente migliore e più veloce grazie alla capacità di fornire funzionalità continue di calcolo e analisi ai dispositivi terminali.

La capacità di risposta in tempo reale offerta dall'edge

computing è garantita dal fatto che non è più necessario trasferire dati nel cloud per l'elaborazione; in tal modo, si eliminano i problemi di latenza che rappresentano ovviamente un problema importante nelle applicazioni soggette a vincoli temporali, in cui le decisioni basate sull'AI possono influenzare le prestazioni in tempo reale delle macchine IoT.

L'edge computing aumenta anche l'affidabilità delle operazioni, in quanto le decisioni relative a processi, stato della macchina e operazioni vengono eseguite localmente con meno preoccupazioni sulla connettività. Inoltre, i parametri che identificano le situazioni in cui è necessaria la manutenzione predittiva sono integrati nella soluzione IoT, quindi le informazioni in tempo reale possono garantire processi ininterrotti, evitando eventuali guasti alle apparecchiature o malfunzionamenti improvvisi. Un altro vantaggio dell'edge computing consiste nel mantenimento dei dati sensibili nell'ecosistema IT locale, evitando quindi i problemi di sicurezza causati dal cloud pubblico. Le soluzioni rese possibili dall'AI sono in grado di rilevare anomalie ai confini della rete se i cyber attacker tentano di accedere alla rete tramite dispositivi IoT, attivando rapidamente tattiche difensive. L'analisi dei rischi determina tutti i possibili punti di ingresso per gli aggressori ed eventuali problemi alla sicurezza vengono attenuati dalla creazione automatica di piani proattivi. Inoltre, non risulta più necessaria una larghezza di banda dispendiosa per garantire la connettività, poiché l'edge computing aggrega i dati localmente anziché inviarli al cloud.

LA PORTA VERSO IL SUCCESSO

Un dispositivo gateway esegue l'elaborazione completamente o parzialmente. Ad esempio, il dispositivo può eseguire un modello AI semplificato responsabile dell'analisi iniziale dei dati. Nel caso dei sistemi di sor-

Figura 1 - Gateway IoT Avnet SmartEdge.





veglia destinati a rilevare le attività, tale modello può semplicemente stabilire se il segnale è costituito da un rumore di fondo, come il vento, oppure se è da ascrivere a un intruso che passa nelle vicinanze o al rumore improvviso di una finestra rotta.

È possibile trasferire le funzioni più complesse a un gateway che raccolga i dati provenienti da diverse videocamere o altri dispositivi di sicurezza. I dispositivi possono svolgere una sorta di pre-elaborazione al fine di ottimizzare il lavoro del modello del gateway. Integrando soluzioni edge e gateway, è possibile estendere la portata dell'AI. Nonostante, l'AI sull'edge offra la più bassa latenza di comunicazione, la velocità di elaborazione più rapida di un gateway locale è in grado di superare un processore dalle prestazioni ridotte e offrire i migliori parametri di latenza e produttività.

L'integrazione di dispositivi gateway comporta altri vantaggi, quale l'enorme potenza di elaborazione fornita per supportare più dispositivi locali. Gateway, come la nuova scheda di sviluppo Ultra96 di Avnet e il nuovo gateway industriale Avnet IoT basato sul Raspberry Pi, possono essere utilizzati insieme ai dispositivi edge per aumentare la potenza di elaborazione e intraprendere lavori più intensi. Un gateway collegato a una serie di dispositivi periferici o edge fornirà la massima flessibilità per l'elaborazione, laddove necessario e dove può essere intrapresa in modo efficace.

Esaminando con attenzione il gateway IoT di Avnet SmartEdge, questa unità combina semplicità, funzionalità e accessibilità mai raggiunta prima in questo settore. Per gli sviluppatori che hanno familiarità con la prototipazione su un dispositivo Raspberry Pi sarà semplice portare i propri progetti in un ambiente di produzione senza dover mai cambiare piattaforma. Il gateway IoT industriale Avnet SmartEdge consente una connettività continua e sicura al cloud, attraverso la piattaforma IoT Connect di Avnet in esecuzione su Microsoft Azure. Ciò lo rende estremamente utile nello sviluppo di applicazioni di automazione industriale come il monitoraggio re-

moto, la manutenzione predittiva, il controllo dei processi e l'automazione.

In grado di soddisfare i requisiti chiave per i clienti in una vasta gamma di applicazioni industriali, il gateway include porte aggiuntive per gestire protocolli industriali e un modulo di calcolo fidato per una maggiore sicurezza sull'hardware. Le opzioni di connettività comprendono un set completo di periferiche industriali tra cui le interfacce CAN-BUS e RS-232/485 (Modbus e DeviceNet), un I/O digitale isolato e Ethernet 10/100 doppio, con connettività wireless tramite WiFi a 2,4 GHz e BLE 4.2 con antenna integrata e connettività esterna, nonché un'interfaccia mPCIe per i modem cellulari. L'alimentazione è fornita tramite un ingresso CC da 12 V-24 V, è presente un'uscita del display HDMI e un EMMC da 8 Gbyte sostituisce la scheda di memoria SD standard, con la sicurezza garantita da un TPM 2.0 integrato. Il dispositivo ha anche un'intestazione di espansione a 40 pin per collegarsi a una gamma di dispositivi dall'ecosistema di HAT e altri accessori. Infine, il contenitore ha spazio per un accessorio interno extra che fornisce capacità di comunicazione Zigbee, Z-Wave o LoRa, oppure per accessori multipli tramite espansione del contenitore. Le opzioni di montaggio includono guida DIN, a parete o free-standing.

TROVARE LA GIUSTA SOLUZIONE DI APPRENDIMENTO

Mentre l'intelligenza artificiale può rappresentare una potente tecnologia per il controllo industriale, gli utenti devono affrontare la questione della necessità di una conoscenza ed esperienza approfondite per sfruttarla al massimo. Lo sviluppatore industriale dovrà decidere come determinare le forme di apprendimento automatico da implementare nella propria applicazione di destinazione.

Octonion ha sviluppato una soluzione a questo problema con un'architettura che riduce enormemente la curva di apprendimento per gli ingegneri che desiderano

sviluppare soluzioni AI per sistemi industriali e IoT. La piattaforma software Brainium di Octonion è una soluzione flessibile che risponde alle esigenze degli utenti industriali grazie alla sua capacità di operare su tre livelli: dispositivo, gateway e cloud. L'utente può decidere il livello migliore in base all'ambiente di distribuzione scelto. A livello di dispositivo, l'utente ha una vasta scelta per la piattaforma hardware, ma il rapido sviluppo è facilitato dal dispositivo SmartEdge Agile con piattaforma software **Brainium**, un dispositivo IoT a basso consumo che supporta un'ampia gamma di sensori per il rilevamento di movimento e ambientali. Per quanto riguarda la distribuzione, i clienti possono usare l'hardware SmartEdge Agile pronto per l'uso o regolare la progettazione da usare nelle loro implementazioni.

Nell'architettura Brainium, l'hardware del dispositivo con sensori intelligenti comunicherà con un gateway locale utilizzando una rete a bassa potenza come BLE, LoRa-WAN, SigFox, cellulare o WiFi.

Questo modulo gateway offre generalmente prestazioni di calcolo più elevate, diventando un obiettivo adatto per i modelli più grandi di intelligenza artificiale. Inoltre, il gateway fornisce l'accesso al cloud per mettere a disposizione ulteriori servizi, come l'addestramento del modello. Un gateway può essere implementato su hardware già pronto come Raspberry Pi o qualsiasi piattaforma in grado di eseguire Android o iOS. Il livello di infrastruttura del cloud di Brainium può essere implementato su soluzioni server AWS, Microsoft Azure o personalizzate. L'ambiente del software Brainium coordina le attività dei livelli di infrastruttura del cloud, del dispositivo e del gateway per formare un ambiente olistico a disposizione

dell'intelligenza artificiale, in grado di supportare varie combinazioni di tecniche di apprendimento automatico senza supervisione e supervisionate. Ad esempio, il clustering può essere applicato a livello di dispositivo per riconoscere modelli di dati comuni. Questi modelli di dati possono essere utilizzati come input in un modello supervisionato che verrà eseguito a livello di gateway; tuttavia, per evitare di appesantire i dispositivi integrati con le elevate esigenze computazionali dell'addestramento, l'apprendimento può avvenire nel cloud. Il flessibile costruttore di modelli integrato nel software AI Studio fornisce gli strumenti per rendere possibile la messa a punto del modello e dell'ambiente di esecuzione per il caso d'uso di destinazione.

L'intelligenza artificiale ha compiuto il passaggio al cloud, nei gateway e sull'edge, e ora che l'integrazione di edge e gateway ha dimostrato di estendere la portata dell'AI, è chiaro che le tecnologie AI sono diventate un elemento indispensabile in una vasta gamma di applicazioni industriali.

La scelta dell'opzione AI migliore per un particolare caso d'uso dipende interamente dallo sviluppatore dell'applicazione; tuttavia, sono disponibili un certo numero di soluzioni in grado di soddisfare i requisiti di gestione e controllo dei dati di oggi, di domani e del prossimo futuro.

Cliff Ortmeier è Global Head of Technical Marketing presso Farnell.

CHECKLIST: CLOUD OR EDGE?		
	AI SUPPORTED BY THE CLOUD	AI AT THE EDGE
Collected data	<input type="checkbox"/> Does not drive real-time decisions	<input type="checkbox"/> Need to collect and analyze data to drive real-time decisions
Latency	<input type="checkbox"/> Slower transmission of collected data is acceptable	<input type="checkbox"/> Near instantaneous data transmission is optimal
Bandwidth	<input type="checkbox"/> Can support bandwidth requirements to send large amounts of data to the cloud for analysis	<input type="checkbox"/> Local network support for data collection is preferred
Autonomy	<input type="checkbox"/> Reliance on the cloud for connectivity is acceptable	<input type="checkbox"/> Network downtime affects productivity of connected devices

Tabella - Confronto tra elaborazione Edge e Cloud nel caso specifico dell'Intelligenza Artificiale.