

Intelligenza Artificiale e Machine Learning nei Sistemi Industriali di Manutenzione

Massimo Guarascio, Giuseppe Manco, Luigi Pontieri, Ettore Ritacco

Istituto di Calcolo e Reti ad Alte Prestazioni ICAR-CNR

massimo.guarascio@icar.cnr.it, giuseppe.manco@icar.cnr.it,

luigi.pontieri@icar.cnr.it, etttore.ritacco@icar.cnr.it

Abstract

L'industria 4.0 mira all'automatizzazione e all'ottimizzazione di tutti i processi produttivi. Tra questi la manutenzione occupa una posizione rilevante poiché è in grado di abbracciare tutti i settori di produzione. In questo ambito il team di ricerca si è concentrato nello sviluppo di una metodologia e tecniche di manutenzione predittiva basata sull'analisi dei dati provenienti da sensori.

1 L'industria alle porte del 2020

L'attuale mondo dell'industria è soggetto ad una serie di requisiti che nascono dalle leggi e dalle dinamiche di mercato. Ci sono due forti spinte alla base della definizione o aggiornamento dei processi di produzione: da una parte c'è una richiesta di miglioramento costante della competitività nei settori industriali aggrediti, dall'altra la completa ottemperanza della legislatura vigente in materia di sostenibilità, sicurezza e diritto del lavoro. La capacità di definire e realizzare processi automatizzati di controllo ed intervento, il cui obiettivo è soddisfare le due suddette direttrici, rappresenta un challenge per il mondo della ricerca ma al tempo stesso una grande opportunità per l'industria di spingersi oltre gli attuali limiti.

La manutenzione è una delle attività di supporto più importanti in tutti i sistemi di produzione industriale. Essa rappresenta al tempo stesso sia un'opportunità che un collo di bottiglia. Infatti, se da una parte consente di evitare l'incorrere di situazioni di guasto dei macchinari, di interruzioni dei servizi ed eventualmente di penali monetarie, dall'altra rischia di rallentare i processi industriali, poiché richiede di avviare attività non strettamente correlate con il business core, impegnando i macchinari di produzione e numerose risorse in termini di tecnici, strumentazione, tempo, costi eventuali.

Il valore di un buon piano di manutenzione dipende dalla sua capacità di adattarsi al processo di produzione, limitando gli interventi ai casi di necessità. Ad oggi esistono tre tipologie di manutenzione: (i) *Correttiva*, le attività di manutenzione vengono avviate solo a guasto avvenuto; (ii) *Programmata*, ad intervalli regolari (di tempo, spazio, numero utilizzi, ...) i macchinari di produzione vengono controllati e qualora necessario riparati o sostituiti; (iii) *Predittiva*, un sistema intelligente cerca di anticipare l'insorgere di guasti analizzando

eventuali sintomi predittori di situazioni anomale. Dal punto di vista del core business, questo tipo di manutenzione è quella ottima poiché elimina gli stop di interventi di controllo (effettuato offline, senza la necessità del macchinario di produzione monitorato) e riduce la riparazione/sostituzione ai soli casi di necessità.

La manutenzione predittiva necessita, dunque, di evolute tecniche di machine learning ed intelligenza artificiale, poiché solo elevate capacità di predizione del guasto corrispondono ad un effettivo miglioramento del processo di produzione.

2 Linee di ricerca emergenti

Il modello di analisi della manutenzione predittiva può essere astratto attraverso semplici step. In primo luogo bisogna considerare la fonte dati: con l'avvento dell'IoT, sui dispositivi di produzione sono installati numerosi sensori, che producono notevoli quantità di dati in base alla loro frequenza di campionamento. Per tal motivo bisogna affrontare due problemi: la riconciliazione delle rilevazioni, poiché provenienti spesso da sistemi eterogenei, e la memorizzazione efficiente e scalabile delle informazioni, necessaria a causa della natura Big Data dei sensor stream.

Il secondo passo è la caratterizzazione del dato. In particolare, è da notare che i sensor stream sono sequenze temporali molto rumorose, ad eccessiva dimensionalità ed affette da particolari problematiche che possono variare da caso a caso (es. effetti raffica). È di fondamentale importanza riuscire a selezionare solo le informazioni strettamente correlate all'analisi di prevenzione, limitando il costo delle risorse e della computazione dei modelli predittivi. Le sequenze temporali ripulite e filtrate sono il carburante dei modelli di previsione del guasto. Questi modelli si fondano sul Machine Learning; in particolare possiamo distinguere gli *Supervisionato* (le sequenze temporali storiche sono etichettate con variabili di notifica di guasto) e *Non supervisionato* (assenza delle notifiche di guasto).

L'ultimo passo è la spiegazione del guasto, ovvero la presentazione della motivazione che ha portato il modello di previsione alla predizione di un guasto. Tale passo è di fondamentale importanza poiché agevola notevolmente il lavoro del decisore del processo di produzione e del team di manutenzione.

3 Risultati di ricerca

Classification vs. Anomaly Detection. La scelta tra classificazione e scoperta di anomalia non dipende soltanto dalla presenza o meno dell'indicatore del guasto. Ci sono numerosi scenari in cui, nonostante la presenza di tale flag, la classificazione non può riuscire in una predizione di qualità. I guasti sono eventi rari e spesso non sono sufficienti a definire dei pattern di predizione che sono necessari per qualsiasi tecnica di classificazione. Perciò, in questi casi, è preferibile utilizzare approcci di anomaly detection [Manco *et al.*, 2017].

Sliding Windows vs. Survival Analysis. La scoperta del momento in cui si può avverare un guasto è un task di notevole complessità, che, per essere applicabile in sistemi di produzione reali, deve garantire una qualità predittiva elevata. La natura sequenziale del sensor stream e la presenza marcata di rumore, spesso, non consentono ai modelli di previsione del tempo di guasto di ottenere capacità previsionali accettabili [Manco *et al.*, 2018]. Un approccio a sliding windows (sia in fase di apprendimento che di predizione), invece, consente di mitigare la rumorosità dei dati, abilitando il calcolo di statistiche aggregate, e consente di effettuare una predizione di presenza o meno di guasto in un intervallo temporale, offrendo margini di tolleranza alla predizione (Figura 1).

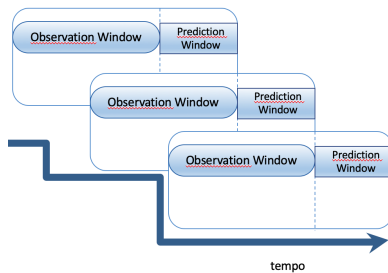


Figura 1: Approccio sliding windows. Si analizza il dispositivo durante una *Observation Window* per predire se si verificherà o meno un guasto durante la *Prediction Window*.

Autoencoded Anomaly Detection. L'utilizzo di una deep neural network nota come Autoencoder può migliorare notevolmente la capacità dell'individuazione delle anomalie. Tale guadagno si ottiene perché le reti neurali effettuano un'approssimazione non lineare della distribuzione che governa la generazione dei dati, senza essere affetta dai bias dei classici modelli di anomaly detection. L'idea di base è piuttosto semplice: se un autoencoder addestrato non è in grado di replicare una sequenza di informazioni, vorrà dire che quest'ultima devia rispetto allo schema dominante che determina la configurazione della rete; in altre parole siamo in presenza di un outlier. Lo schema generale è riportato in Figura 2.

Explanation. Provare a determinare le cause di un guasto è un compito che si può svolgere solo se si è esperti di dominio [Angiulli *et al.*, 2017]. Per tal motivo il team ha adottato una soluzione di supporto all'esperto evidenziando quali siano le feature predominanti che hanno condotto il modello alla predizione del guasto, evidenziando la devianza rispetto al comportamento atteso (vedi Figura 3).

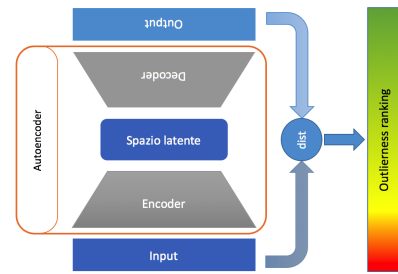


Figura 2: La distanza tra l'input e la replica dell'autoencoder è l'indice di devianza.

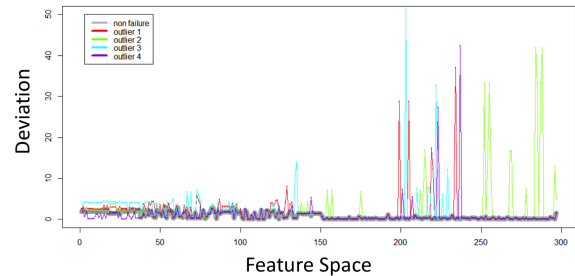


Figura 3: Devianza di un outlier rispetto al comportamento atteso.

4 Scenari d'Uso

Gli scenari d'uso in cui il team sta verticalizzando e testando la metodologia sono: manutenzione di ascensori, treni merce e passeggeri, impianti di produzione di energia eolica, acquedotti e sistemi di gestione degli asset energetici.

5 Attività Progettuali

Il team ha sviluppato la metodologia durante progetti interni, per conto terzi, due POR Calabria e due PON, in alcuni casi in veste di partner in altri in veste di consulente. Il progetto per conto terzi è Cobalt sviluppato in collaborazione con Bombardier Transportation. I due POR Calabria sono "PROSiT: Una piattaforma di Big Data Analytics per la prognostica" e "original Advanced Metering Infrastructure - origAMI", mentre i PON sono "S2BDW: Smarter Solutions in the Big Data World" e "True Detective 4.0: Strumenti e Servizi Intelligenti di Monitoraggio in Tempo Reale per la Manutenzione Predittiva di apparati, per l'Ottimizzazione dei Processi Produttivi e di Automazione Industriale e per la Gestione della Sicurezza Fisica in Ambito Aziendale" (in fase di valutazione).

Riferimenti bibliografici

- [Angiulli *et al.*, 2017] F. Angiulli, et al. Outlying property detection with numerical attributes. *Data Min. Knowl. Discov.*, 31(1):134–163, 2017.
- [Manco *et al.*, 2017] G. Manco, et al. Fault detection and explanation through big data analysis on sensor streams. *Expert Syst. Appl.*, 87:141–156, 2017.
- [Manco *et al.*, 2018] G. Manco, G. Pirrò, e E. Ritacco. Predicting temporal activation patterns via recurrent neural networks. In *ISMIS*, pages 347–356, 2018.