



Come utilizzare l'apprendimento automatico per il rilevamento di anomalie e il monitoraggio delle condizioni

Caso d'uso concreto per l'apprendimento automatico e l'analisi statistica



Vegard Flovik [Seguire](#)

31 dic 2018 · 13 minuti di lettura ★

In questo articolo, introdurrò un paio di diverse tecniche e applicazioni di machine learning e analisi statistiche, e poi mostrerò come applicare questi approcci per risolvere un caso d'uso specifico per il rilevamento di anomalie e il monitoraggio delle condizioni.

Trasformazione digitale, digitalizzazione, Industria 4.0, ecc.

Questi sono tutti termini che probabilmente hai sentito o letto prima. Tuttavia, dietro tutte queste parole d'ordine, l'obiettivo principale è l'uso della tecnologia e dei dati per aumentare la produttività e l'efficienza. La connettività e il flusso di informazioni e dati

tra dispositivi e sensori consente un'abbondanza di dati disponibili. L'attivatore chiave è quindi in grado di utilizzare queste enormi quantità di dati disponibili e di estrarre effettivamente informazioni utili, consentendo di ridurre i costi, ottimizzare la capacità e ridurre al minimo i tempi di fermo. È qui che entra in gioco il recente ronzio dell'apprendimento automatico e dell'analisi dei dati.

Rilevazione di anomalie

Il rilevamento di anomalie (o rilevamento anomalo) è l'identificazione di elementi, eventi o osservazioni rari che destano sospetti differendo significativamente dalla maggior parte dei dati. In genere, i dati anomali possono essere collegati a qualche tipo di problema o evento raro come ad esempio frodi bancarie, problemi medici, difetti strutturali, malfunzionamenti delle apparecchiature ecc. Questa connessione rende molto interessante essere in grado di individuare quali punti di dati possono essere considerati anomalie , poiché l'identificazione di questi eventi è in genere molto interessante dal punto di vista commerciale.

Questo ci porta a uno degli obiettivi chiave: come possiamo identificare se i punti dati sono normali o anomali? In alcuni casi semplici, come nell'esempio seguente, la visualizzazione dei dati può darci informazioni importanti.

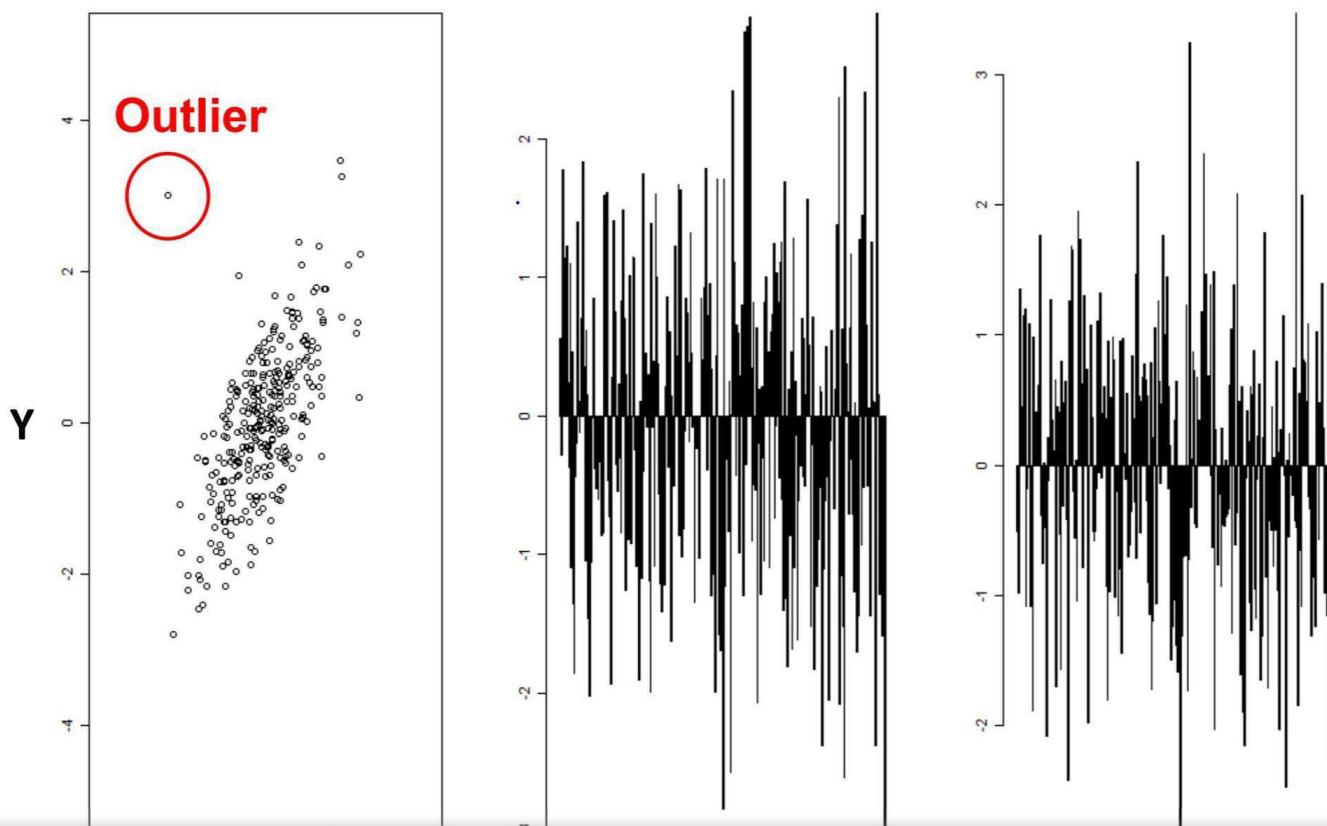




Figura 1: rilevamento anomalie per due variabili

In questo caso di dati bidimensionali (X e Y), diventa abbastanza facile identificare visivamente anomalie attraverso punti dati situati al di fuori della distribuzione tipica. Tuttavia, guardando le figure a destra, non è possibile identificare il valore anomalo direttamente dall'indagine di una variabile alla volta: è la **combinazione** delle variabili X e Y che ci consente di identificare facilmente l'anomalia. Ciò complica sostanzialmente la questione quando si passa da due variabili a 10–100s di variabili, come spesso accade nelle applicazioni pratiche del rilevamento di anomalie.

Collegamento al monitoraggio delle condizioni

Qualsiasi macchina, che sia una macchina rotante (pompa, compressore, turbina a gas o a vapore, ecc.) O una macchina non rotante (scambiatore di calore, colonna di distillazione, valvola, ecc.) Raggiungerà alla fine un punto di cattiva salute. Quel punto potrebbe non essere quello di un guasto o di un arresto effettivo, ma quello in cui l'apparecchiatura non agisce più nel suo stato ottimale. Questo segnala che potrebbe essere necessario un po 'di attività di manutenzione per ripristinare il pieno potenziale operativo. In termini semplici, identificare lo "stato di salute" delle nostre apparecchiature è il dominio del monitoraggio delle condizioni .

Il modo più comune per eseguire il monitoraggio delle condizioni è esaminare ogni misurazione del sensore dalla macchina e imporre un limite di valore minimo e massimo su di essa. Se il valore corrente rientra nei limiti, la macchina è integra. Se il valore corrente non rientra nei limiti, la macchina non è salutare e viene inviato un allarme.

È noto che questa procedura di impostazione di limiti di allarme codificati in modo rigido invia un gran numero di falsi allarmi, ovvero allarmi per situazioni che sono effettivamente stati salutari per la macchina. Ci sono anche allarmi mancanti, ovvero situazioni problematiche ma non allarmate. Il primo problema non solo fa perdere tempo e fatica, ma anche la disponibilità dell'attrezzatura. Il secondo problema è più cruciale in quanto porta a danni reali con i costi di riparazione associati e la perdita di produzione.

Entrambi i problemi derivano dalla stessa causa: la salute di un'apparecchiatura

complessa non può essere valutata in modo affidabile sulla base dell'analisi di ciascuna

misurazione da sola (come illustrato anche nella figura 1 della sezione precedente sul rilevamento delle anomalie). Dobbiamo piuttosto considerare una **combinazione** delle varie misurazioni per ottenere una vera indicazione della situazione

• • •

Sezione tecnica:

È difficile trattare gli argomenti dell'apprendimento automatico e dell'analisi statistica per il rilevamento di anomalie senza entrare anche in alcuni aspetti più tecnici. Eviterò ancora di approfondire lo sfondo teorico (ma fornirò alcuni collegamenti a descrizioni più dettagliate). Se sei più interessato alle applicazioni pratiche dell'apprendimento automatico e all'analisi statistica quando si tratta, ad esempio, del monitoraggio delle condizioni, passa alla sezione "Caso d'uso del monitoraggio delle condizioni".

Approccio 1: analisi statistica multivariata

Riduzione della dimensionalità mediante analisi dei componenti principali: PCA

Poiché la gestione di dati ad alta dimensione è spesso impegnativa, esistono diverse tecniche per ridurre il numero di variabili (riduzione della dimensionalità). Una delle tecniche principali è l'analisi dei componenti principali (PCA), che esegue una mappatura lineare dei dati su uno spazio di dimensione inferiore in modo tale da massimizzare la varianza dei dati nella rappresentazione a bassa dimensione. In pratica, viene costruita la matrice di covarianza dei dati e gli autovettori di questa matrice vengono calcolati. Gli autovettori che corrispondono agli autovalori più grandi (i componenti principali) possono ora essere utilizzati per ricostruire una grande frazione della varianza dei dati originali. Lo spazio delle caratteristiche originali è stato ora ridotto (con una certa perdita di dati, ma si spera conservando la varianza più importante) allo spazio attraversato da alcuni autovettori.

Rilevazione di anomalie multivariate

Come abbiamo notato sopra, per identificare le anomalie quando si ha a che fare con una o due variabili, la visualizzazione dei dati può spesso essere un buon punto di

partenza. Tuttavia, quando si ridimensionano questi dati a dati ad alta dimensione

(come spesso accade nelle applicazioni pratiche), questo approccio diventa sempre più difficile. Questo è fortunatamente dove le statistiche multivariate vengono in aiuto.

Quando si ha a che fare con una raccolta di punti dati, in genere avranno una certa distribuzione (ad es. A Distribuzione gaussiana). Per rilevare anomalie in un modo più quantitativo, calcoliamo innanzitutto la distribuzione di probabilità $p(x)$ dai punti dati. Quindi quando un nuovo esempio, x , entra, ci confrontiamo $p(x)$ con una soglia r . Se $p(x) < r$, è considerata un'anomalia. Questo perché gli esempi normali tendono ad avere un grande $p(x)$ mentre esempi anomali tendono ad avere un piccolo $p(x)$.

Nel contesto del monitoraggio delle condizioni, ciò è interessante perché le anomalie possono dirci qualcosa sullo "stato di salute" dell'apparecchiatura monitorata: i dati generati quando l'apparecchiatura si avvicina al guasto o un'operazione non ottimale, in genere hanno una distribuzione diversa rispetto ai dati di Attrezzatura "sana".

La distanza di Mahalanobis

Considera il problema di stimare la probabilità che un punto dati appartenga a una distribuzione, come descritto sopra. Il nostro primo passo sarebbe quello di trovare il centroide o il centro di massa dei punti campione. Intuitivamente, più il punto in questione è vicino a questo centro di massa, più è probabile che appartenga all'insieme. Tuttavia, dobbiamo anche sapere se l'insieme è distribuito su un ampio raggio o su un piccolo intervallo, in modo da poter decidere se una data distanza dal centro è degna di nota o no. L'approccio semplicistico è stimare la deviazione standard delle distanze dei punti di campionamento dal centro di massa. Inserendolo nella normale distribuzione possiamo ricavare la probabilità del punto dati appartenente alla stessa distribuzione.

Lo svantaggio dell'approccio di cui sopra era che abbiamo ipotizzato che i punti campione fossero distribuiti attorno al centro di massa in modo sferico. Se la distribuzione fosse decisamente non sferica, ad esempio ellissoidale, ci aspetteremmo che la probabilità del punto di test appartenente all'insieme dipenda non solo dalla distanza dal centro di massa, ma anche dalla direzione. Nelle direzioni in cui l'ellissoide ha un asse corto, il punto di prova deve essere più vicino, mentre in quelle dove l'asse è lungo il punto di prova può essere più lontano dal centro. Mettendolo su una base matematica, l'ellissoide che rappresenta meglio la distribuzione di probabilità

dell'insieme può essere stimato calcolando la matrice di covarianza dei campioni. La

distanza di Mahalanobis (MD) è la distanza del punto di prova dal centro di massa divisa per la larghezza dell'ellissoide nella direzione del punto di prova.

Al fine di utilizzare l'MD per classificare un punto di test come appartenente a una delle classi N, si calcola innanzitutto la matrice di covarianza di ciascuna classe, in genere sulla base di campioni noti per appartenere a ciascuna classe. Nel nostro caso, poiché siamo interessati solo a classificare "normale" vs "anomalia", utilizziamo i dati di addestramento che contengono solo le normali condizioni operative per calcolare la matrice di covarianza. Quindi, dato un campione di prova, calcoliamo l'MD nella classe "normale" e classifichiamo il punto di prova come "anomalia" se la distanza supera una determinata soglia.

Nota di cautela: l'uso del MD implica che l'inferenza può essere fatta attraverso la matrice media e covarianza - e che è una proprietà della distribuzione normale da sola. Questo criterio non è necessariamente soddisfatto nel nostro caso, poiché le variabili di input potrebbero non essere distribuite normalmente. Tuttavia, proviamo comunque e vediamo quanto funziona!

Approccio 2: rete neurale artificiale

Reti del codificatore automatico

Il secondo approccio si basa sull'utilizzo di reti neurali autoencoder . Si basa su principi simili a quelli dell'analisi statistica di cui sopra, ma con alcune lievi differenze.

Un autoencoder è un tipo di rete neurale artificiale utilizzata per apprendere codifiche di dati efficienti in modo non supervisionato . Lo scopo di un codificatore automatico è apprendere una rappresentazione (codifica) per un insieme di dati, in genere per la riduzione della dimensionalità . Insieme al lato di riduzione, viene appreso un lato di ricostruzione, in cui l'autocodificatore tenta di generare dalla codifica ridotta una rappresentazione il più vicino possibile al suo input originale.

Dal punto di vista architettonico, la forma più semplice di un autoencoder è una rete neurale feedforward , non ricorrente molto simile ai molti percettori a singolo strato che rende un percettrone multistrato (MLP) - avente uno strato in ingresso, uno strato in

uscita e uno o più strati nascosti che li collegano - ma con il livello di output con lo stesso

numero di nodi del livello di input e allo scopo di *ricostruire* i propri input.

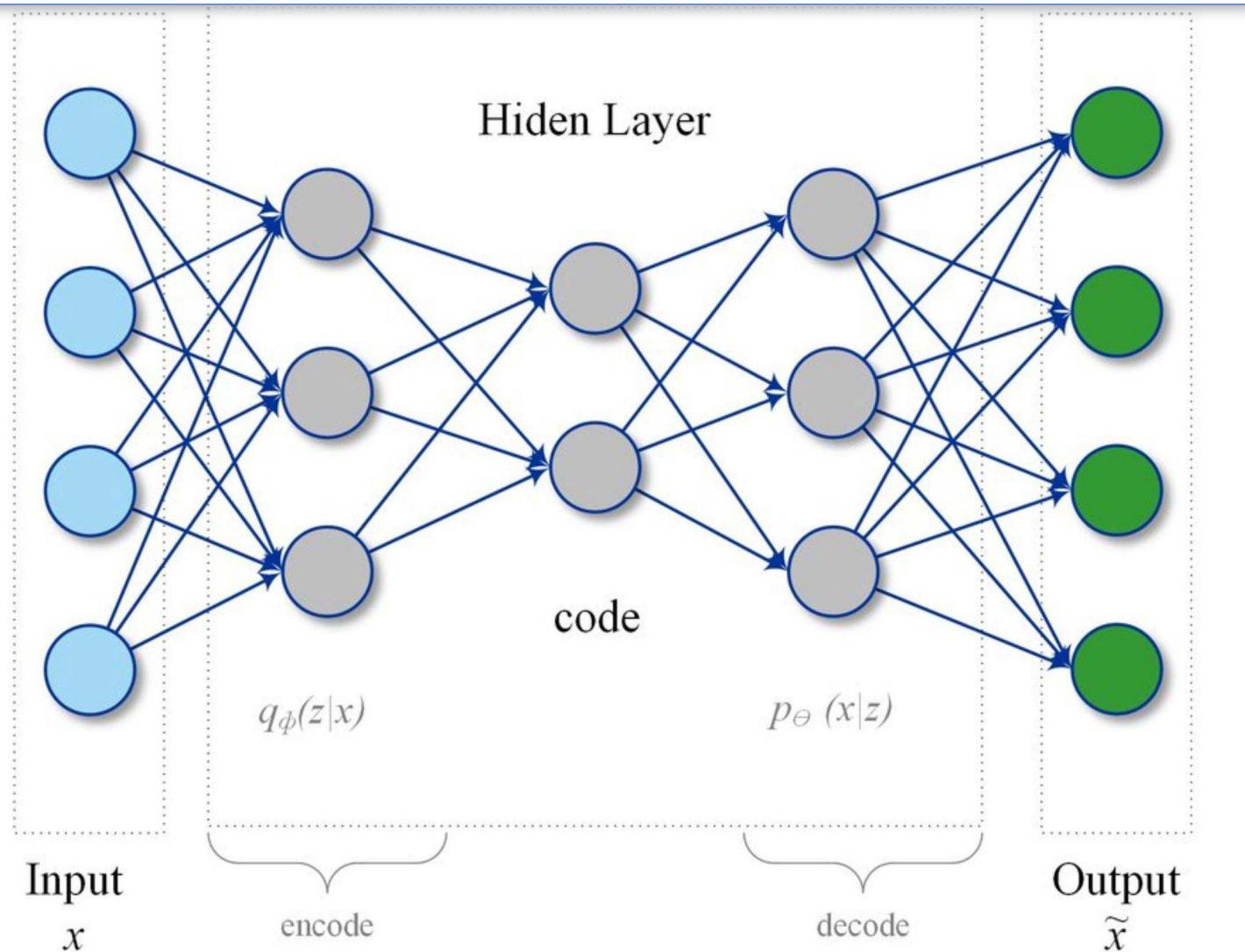


Figura 2: rete del codificatore automatico

Nel contesto del rilevamento delle anomalie e del monitoraggio delle condizioni, l'idea di base è quella di utilizzare la rete del codificatore automatico per "comprimere" le letture del sensore in una rappresentazione di dimensione inferiore, che acquisisce le correlazioni e le interazioni tra le varie variabili. (Essenzialmente lo stesso principio del modello PCA, ma qui permettiamo anche interazioni non lineari tra le variabili).

La rete del codificatore automatico viene quindi addestrata sui dati che rappresentano lo stato operativo "normale", con l'obiettivo di comprimere e quindi ricostruire le variabili di input. Durante la riduzione della dimensionalità, la rete apprende le interazioni tra le varie variabili e dovrebbe essere in grado di ricostruirle nuovamente alle variabili originali all'output. L'idea principale è che quando l'apparecchiatura monitorata si degrada, ciò dovrebbe influire sull'interazione tra le variabili (ad es. Variazioni di

temperatura, pressioni, vibrazioni, ecc.). In questo caso, si inizierà a vedere un errore

maggiori nella ricostruzione della rete delle variabili di input. Monitorando l'errore di ricostruzione, si può quindi ottenere un'indicazione della "salute" dell'attrezzatura monitorata, poiché questo errore aumenta man mano che l'apparecchiatura si degrada. Distanza di Mahalanobis , qui utilizziamo la distribuzione di probabilità dell'errore di ricostruzione per identificare se un punto dati è normale o anomalo.

• • •

Caso d'uso del monitoraggio delle condizioni: guasto del cuscinetto ad ingranaggi

In questa sezione, esaminerò un caso d'uso pratico per il monitoraggio delle condizioni utilizzando i due diversi approcci descritti sopra. Poiché la maggior parte dei dati su cui stiamo lavorando con i nostri clienti non sono apertamente disponibili, ho scelto piuttosto di dimostrare i due approcci sui dati resi disponibili dalla NASA, che possono essere scaricati QUI .

In questo caso d'uso, l'obiettivo è rilevare il degrado dei cuscinetti degli ingranaggi su un motore e fornire un avviso che consenta l'adozione di misure predittive al fine di evitare un guasto agli ingranaggi (che potrebbe essere ad esempio una manutenzione / riparazione pianificata dell'attrezzatura) .

Dettagli sperimentali e preparazione dei dati:

Tre serie di dati composte ciascuna da quattro cuscinetti sono state guastate in condizioni di carico e funzionamento costanti. I segnali di misurazione delle vibrazioni vengono forniti per i set di dati per tutta la durata dei cuscinetti fino al guasto. Il fallimento si è verificato dopo 100 milioni di cicli con una crepa nella razza esterna (vedi il documento readme dalla pagina di download per ulteriori informazioni sugli esperimenti). Poiché l'attrezzatura è stata utilizzata fino a guasti, i dati dei primi due giorni di funzionamento sono stati utilizzati come dati di addestramento per rappresentare le apparecchiature normali e "sane". La parte rimanente dei set di dati per il tempo che ha portato al cedimento del cuscinetto è stata quindi utilizzata come dati di prova, per valutare se i diversi metodi fossero in grado di rilevare il degrado del cuscinetto prima del guasto.

Approccio 1: distanza PCA + Mahalanobis

Come spiegato più dettagliatamente nella "Sezione tecnica" di questo articolo, il primo approccio consisteva nell'eseguire prima un'analisi del componente principale e quindi calcolare la distanza di Mahalanobis (MD) per identificare i punti di dati come normali o anomali (segno di degrado dell'apparecchiatura). La distribuzione dell'MD per i dati di allenamento che rappresentano attrezzature "sane" è illustrata nella figura seguente.

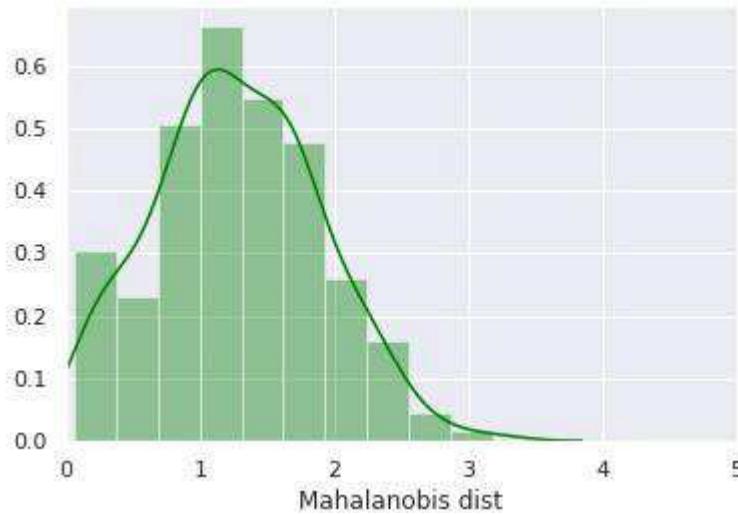


Figura 3: distribuzione della distanza di Mahalanobis per apparecchiature "sane"

Utilizzando la distribuzione di MD per apparecchiature "sane", possiamo definire un valore di soglia per cosa considerare un'anomalia. Dalla distribuzione sopra, possiamo ad esempio definire un $MD > 3$ come un'anomalia. La valutazione di questo metodo per rilevare il degrado dell'apparecchiatura ora consiste nel calcolare l'MD per tutti i punti di dati nel set di test e confrontarlo con il valore di soglia definito per contrassegnarlo come un'anomalia.

Valutazione del modello sui dati di test:

Utilizzando l'approccio di cui sopra, abbiamo calcolato l'MD per i dati di prova nel periodo di tempo che ha portato al cedimento del cuscinetto, come illustrato nella figura seguente.



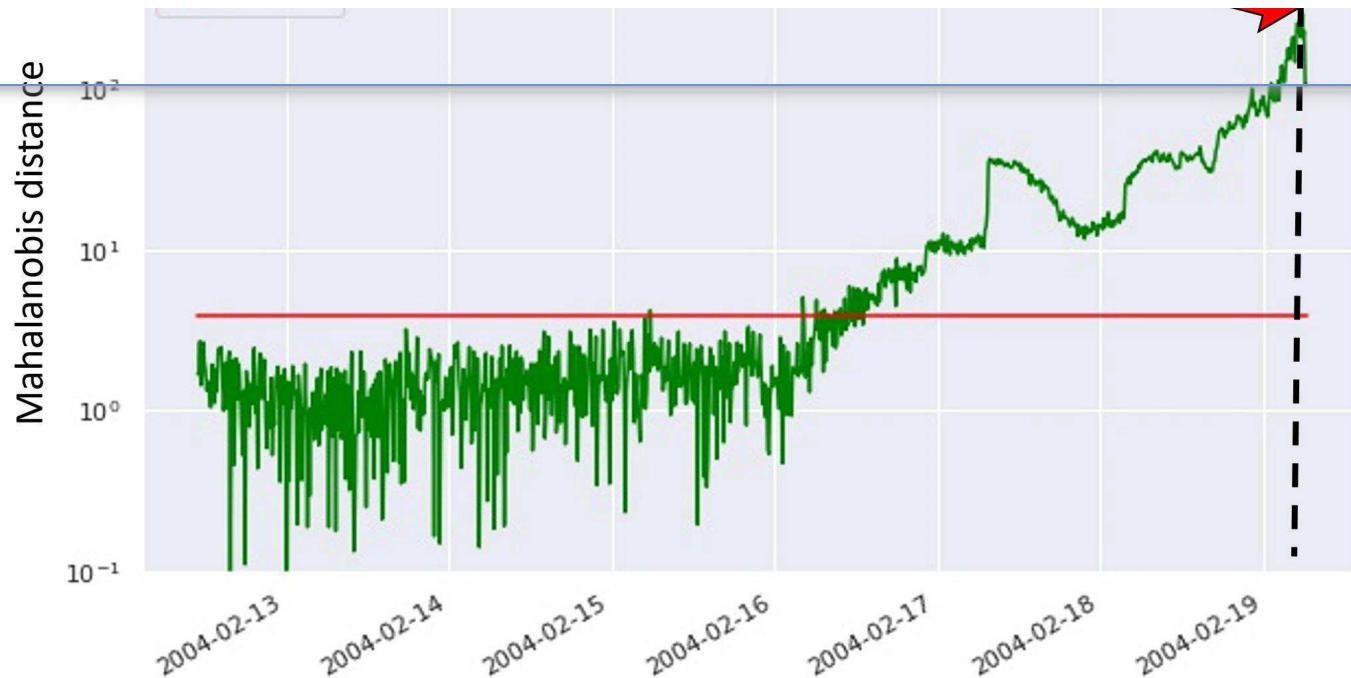


Figura 4: Previsione del guasto del cuscinetto mediante l'approccio 1

Nella figura sopra, i punti verdi corrispondono all'MD calcolato, mentre la linea rossa rappresenta il valore di soglia definito per la segnalazione di un'anomalia. Il guasto del cuscinetto si verifica alla fine del set di dati, indicato dalla linea tratteggiata nera. Ciò dimostra che il primo approccio alla modellazione è stato in grado di rilevare il guasto dell'apparecchiatura imminente circa 3 giorni prima del guasto effettivo (in cui il MD supera il valore di soglia).

Ora possiamo passare attraverso un esercizio simile usando il secondo approccio di modellizzazione, al fine di valutare quale dei metodi funziona meglio dell'altro.

Approccio 2: rete neurale artificiale

Come spiegato più dettagliatamente nella "Sezione tecnica" del documento, il secondo approccio consisteva nell'utilizzare una rete neurale autoencoder per cercare anomalie (come identificato attraverso una maggiore perdita di ricostruzione dalla rete).

Analogamente al primo approccio, qui utilizziamo anche la distribuzione dell'output del modello per i dati di addestramento che rappresentano apparecchiature "sane" per rilevare anomalie. La distribuzione della perdita di ricostruzione (errore assoluto medio) per i dati di allenamento è mostrata nella figura seguente:

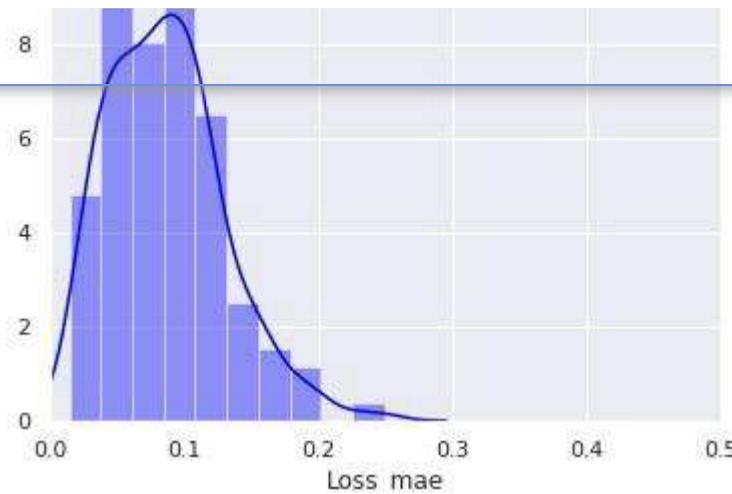
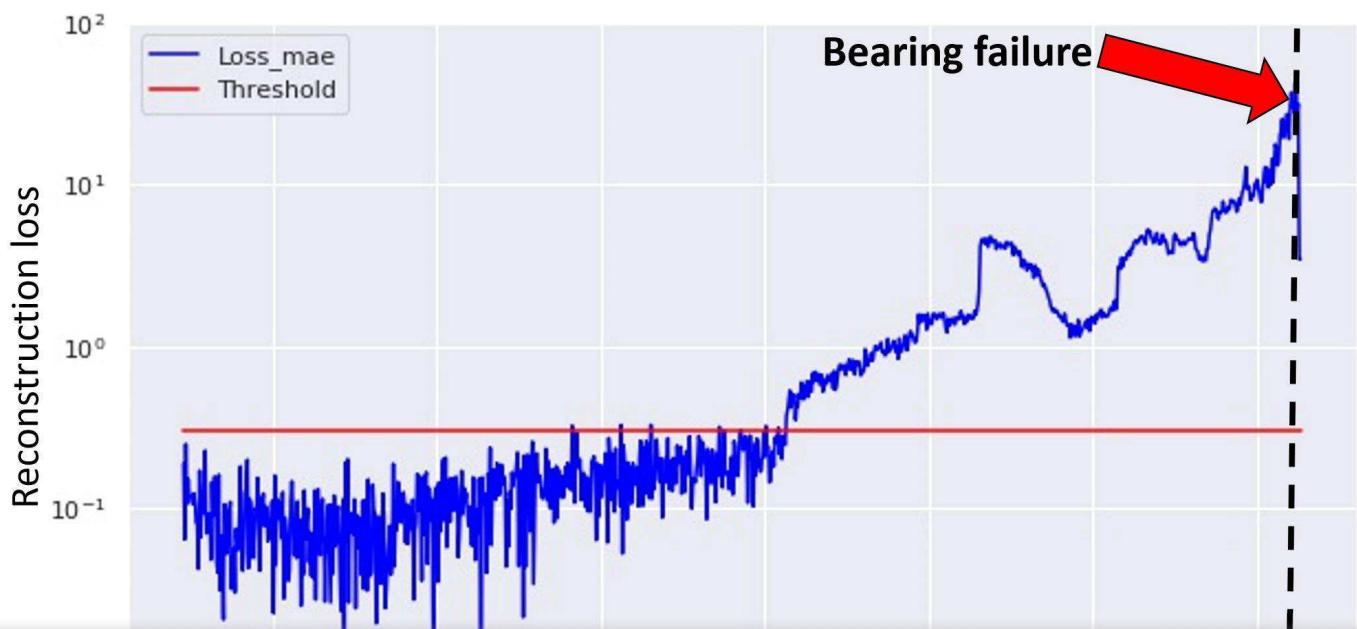


Figura 5: distribuzione della perdita di ricostruzione per apparecchiature "sane".

Utilizzando la distribuzione della perdita di ricostruzione per apparecchiature “sane”, ora possiamo definire un valore soglia per cosa considerare un'anomalia. Dalla distribuzione sopra, possiamo ad esempio definire una perdita $> 0,25$ come un'anomalia. La valutazione del metodo per rilevare il degrado delle apparecchiature ora consiste nel calcolare la perdita di ricostruzione per tutti i punti di dati nel set di test e nel confrontare la perdita con il valore di soglia definito per contrassegnarla come un'anomalia.

Valutazione del modello sui dati di test:

Utilizzando l'approccio sopra, calcoliamo la perdita di ricostruzione per i dati di prova nel periodo di tempo che porta alla rottura del cuscinetto, come illustrato nella figura seguente.



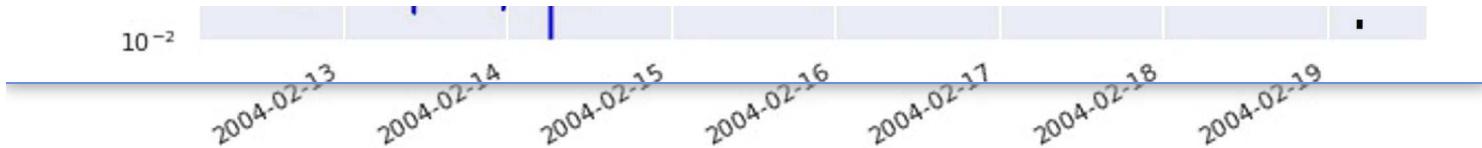


Figura 6: Previsione del guasto del cuscinetto mediante l'approccio 2

Nella figura sopra, i punti blu corrispondono alla perdita di ricostruzione, mentre la linea rossa rappresenta il valore di soglia definito per segnalare un'anomalia. Il guasto del cuscinetto si verifica alla fine del set di dati, indicato dalla linea tratteggiata nera. Ciò dimostra che anche questo approccio alla modellazione è stato in grado di rilevare il prossimo guasto dell'attrezzatura circa 3 giorni prima del guasto effettivo (in cui la perdita di ricostruzione supera il valore di soglia).

Riepilogo dei risultati:

Come si è visto nelle sezioni precedenti sui due diversi approcci per il rilevamento delle anomalie, entrambi i metodi sono in grado di rilevare con successo il prossimo guasto dell'apparecchiatura diversi giorni prima dell'effettiva rottura. In uno scenario di vita reale ciò consentirebbe di adottare misure predittive (manutenzione / riparazione) in anticipo rispetto al guasto, il che significa sia risparmi sui costi che la potenziale importanza per gli aspetti HSE del guasto delle apparecchiature.

Prospettiva:

Con il costo ridotto dell'acquisizione dei dati attraverso i sensori, nonché la maggiore connettività tra i dispositivi, essere in grado di estrarre informazioni preziose dai dati sta diventando sempre più importante. Trovare modelli in grandi quantità di dati è il regno dell'apprendimento automatico e delle statistiche e, a mio avviso, ci sono enormi possibilità di sfruttare le informazioni nascoste in questi dati per migliorare le prestazioni in diversi domini. Il rilevamento delle anomalie e il monitoraggio delle condizioni, come trattato in questo articolo, sono solo una delle tante possibilità.
(L'articolo è disponibile anche [QUI](#))

In futuro, credo che l'apprendimento automatico verrà utilizzato in molti più modi di quanti non possiamo nemmeno immaginare oggi. Quale impatto pensi avrà sui vari settori? Mi piacerebbe sentire i tuoi pensieri nei commenti qui sotto.

Modifica: questo articolo sul rilevamento delle anomalie e il monitoraggio delle

condizioni ha ricevuto molti feedback. Molte delle domande che ricevo riguardano gli

aspetti tecnici e come impostare i modelli, ecc. A causa di ciò, ho deciso di scrivere un articolo di follow-up che copra in dettaglio tutti i passaggi necessari, dalla pre-elaborazione dei dati alla costruzione di modelli e visualizzazione dei risultati.

Altri articoli:

Se hai trovato interessante questo articolo, potrebbero interessarti anche alcuni dei miei altri articoli:

1. Apprendimento automatico: dall'hype alle applicazioni del mondo reale
2. Il rischio nascosto di AI e Big Data
3. AI per la gestione della catena logistica: analisi predittiva e previsione della domanda
4. Come (non) utilizzare l'apprendimento automatico per le previsioni sulle serie storiche: evitare le insidie
5. Come utilizzare l'apprendimento automatico per l'ottimizzazione della produzione: utilizzo dei dati per migliorare le prestazioni
6. Come si insegna la fisica ai sistemi di intelligenza artificiale?
7. Possiamo costruire reti cerebrali artificiali usando magneti su scala nanometrica?