



Anomaly detection nei processi industriali

Federico Trotta



Descrizione del progetto

IL PROGETTO:

Questo progetto è la mia tesi di Laurea. Mi sono laureato in Ingegneria Meccanica con tesi "Analisi dei dati di produzione con tecniche di Data Science" ed è un'analisi di dati di produzione reale. Ne è venuto fuori un metodo per l'anomaly detection nei processi industriali.

L'IDEA:

Nell'industria, un prodotto nasce, tipicamente, a seguito di diverse fasi di produzione. Nelle linee di assemblaggio, in generale, si cerca di fare in modo tale che il tempo per concludere una fase produttiva (tempo ciclo) sia uguale per tutte le fasi produttive di quel prodotto. Nelle PMI, però, capita spesso che un operatore svolga, in giorni diversi, diverse fasi di assemblaggio. La domanda che mi sono posto è: possiamo capire quale operatore sia il più adatto a svolgere una fase produttiva, analizzando lo storico dei suoi tempi ciclo?

Descrizione dell'analisi

Ho scelto un prodotto che viene realizzato in tre fasi produttive: due di assemblaggio ed una di collaudo funzionale.

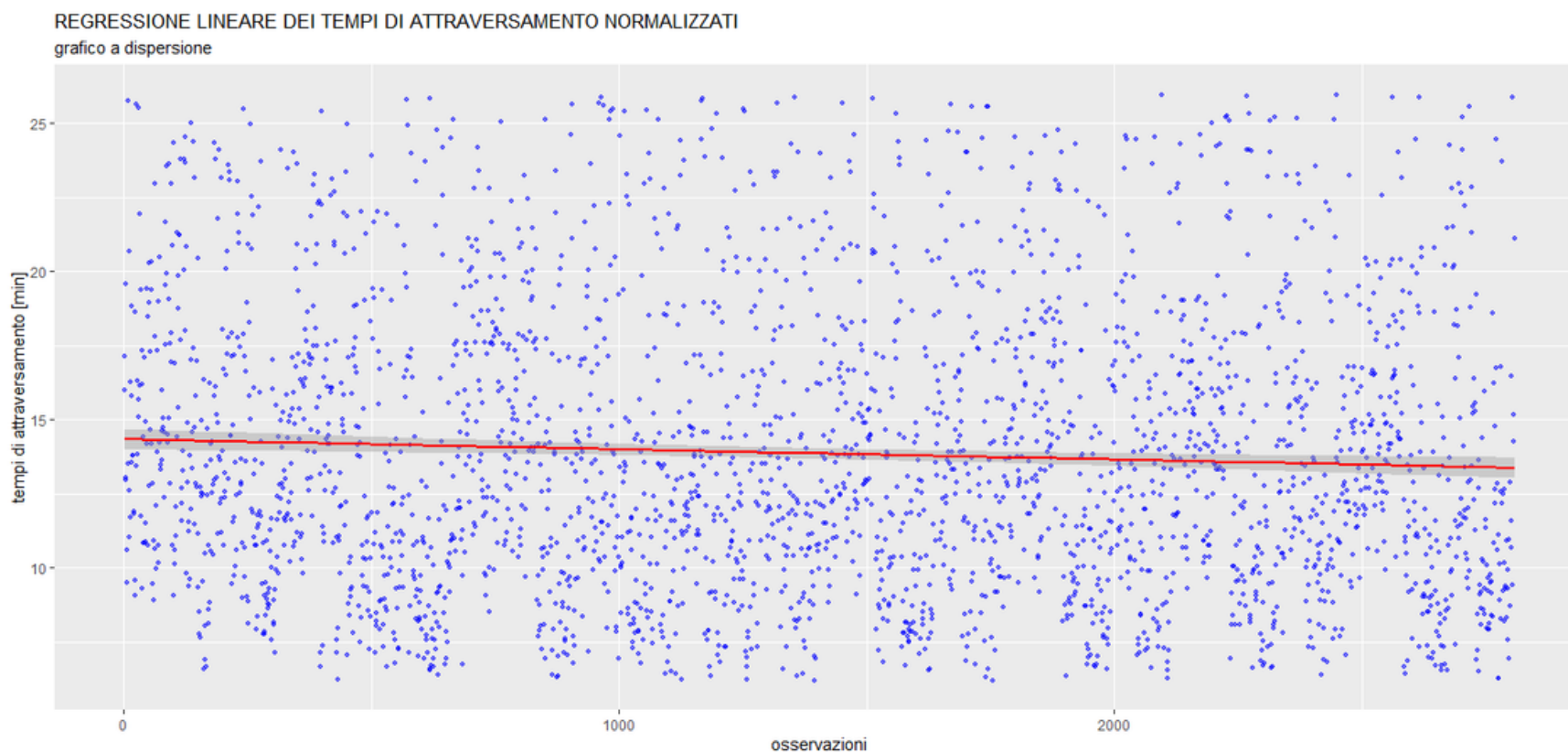
Nei mesi precedenti era stato fatto un lavoro per diminuire al massimo il tempo ciclo della fase di collaudo; per cui, una prima verifica da fare, era quella sul tempo totale per realizzare il prodotto (tempo di attraversamento).

Dopo di che, ho preso in considerazione le prime due fasi di produzione che, essendo di assemblaggio, possono essere variabili in base all'operatore e ne ho studiato i tempi, per vari operatori.

Per ogni operatore ho studiati i tempi prima con la regressione lineare. poi con quella polinomiale, poi con le spline.

Analisi del tempo di attraversamento

Come era lecito aspettarsi, dall'analisi della regressione lineare dei tempi di attraversamento (tempi totali, registrati nei mesi, per realizzare il prodotto) ho ottenuto una retta di regressione lineare con coefficiente negativo; questo significa che, nei mesi, il tempo di attraversamento è diminuito.

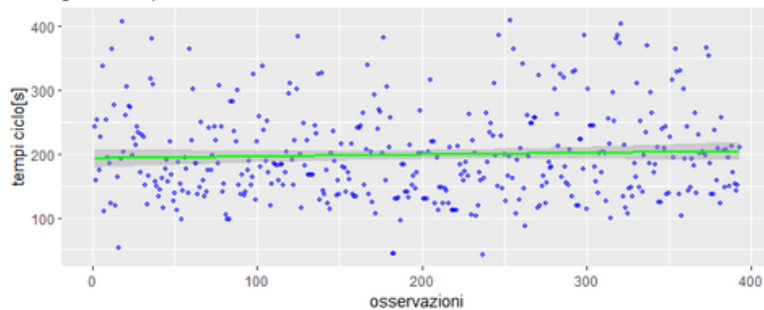


Analisi della prima fase di produzione: regressione lineare

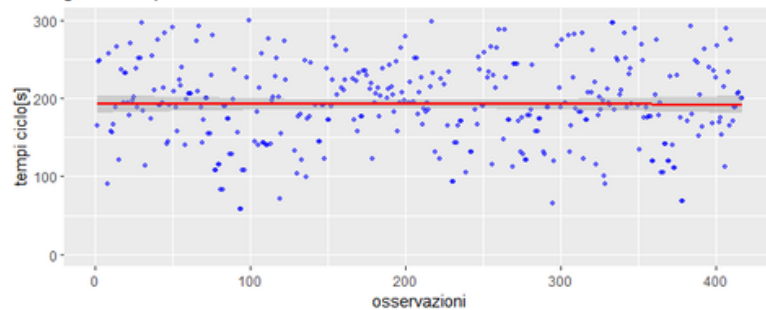
Ho individuato gli operatori che, nel tempo, producono il maggior numero di osservazioni (sono quelli che, più degli altri, lavorano in questa fase produttiva) e sono 4.

Ho calcolato le regressioni lineari dei loro tempi ciclo. Tutti e quattro hanno dei valori di R^2 prossimi allo zero, e le rette sono, infatti, tutte orizzontali. Sembra, quindi, che restino tutti costanti nel loro lavoro (la regressione è orizzontale), che è quello che si vorrebbe nell'industria; ma questa analisi non è sufficiente. Ho, quindi, provato la regressione polinomiale.

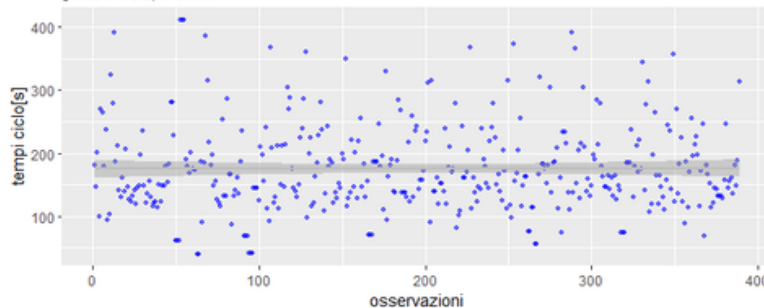
A1 REGRESSIONE LINEARE DEI TEMPI CICLO DI A1
grafico a dispersione



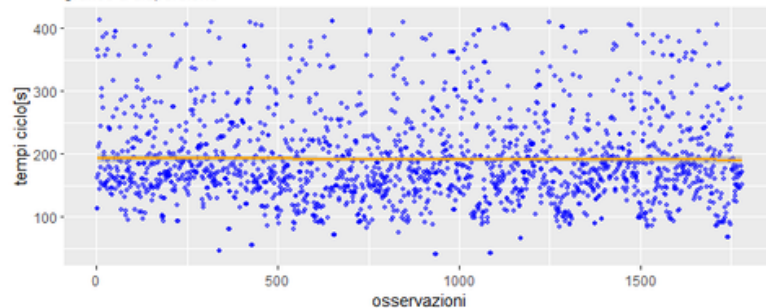
D1 REGRESSIONE LINEARE DEI TEMPI CICLO DI D1
grafico a dispersione



N1 REGRESSIONE LINEARE DEI TEMPI CICLO DI N1
grafico a dispersione



O1 REGRESSIONE LINEARE DEI TEMPI CICLO DI O1
grafico a dispersione

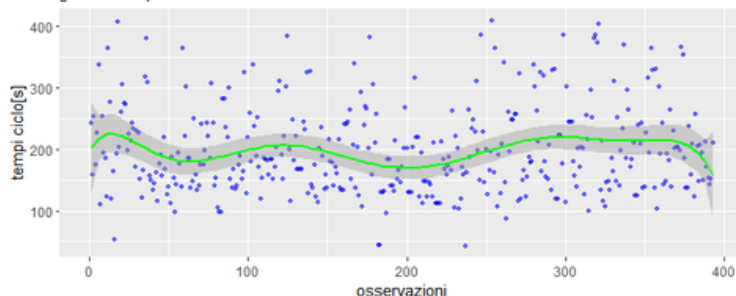


Analisi della prima fase di produzione: regressione polinomiale

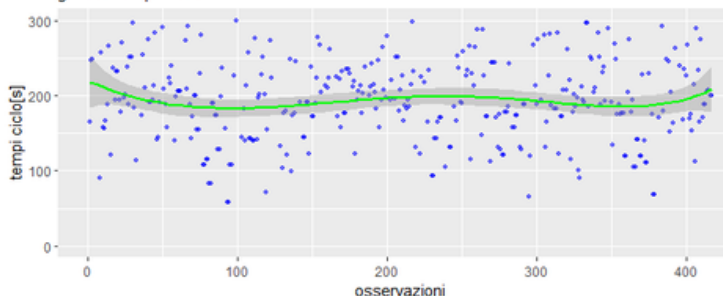
Ho provato diversi gradi di polinomio, fino ad arrivare a quello ottimale, utilizzando anche R^2 come parametro.

Come si vede, per gli operatori N ed O la polinomiale diventa praticamente una retta; D presenta qualche oscillazione; A ha palesemente delle oscillazioni nelle registrazioni dei tempi ciclo. Per essere più sicuro, ho utilizzato le splines che matematicamente sono "più gestibili" a piacere.

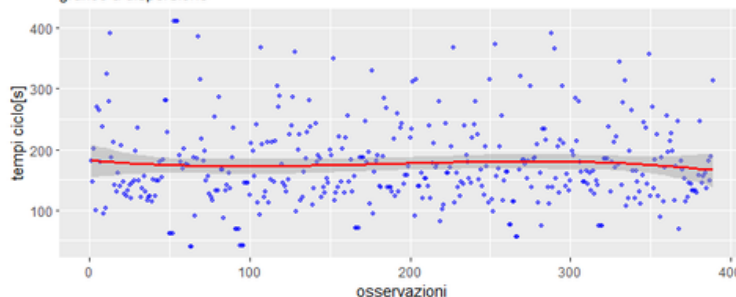
A1 REGRESSIONE POLINOMIALE DEI TEMPI CICLO DI A1
grafico a dispersione



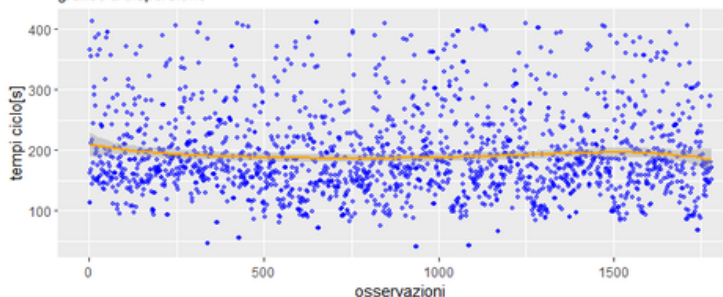
D1 REGRESSIONE POLINOMIALE DEI TEMPI CICLO DI D1
grafico a dispersione



N1 REGRESSIONE POLINOMIALE DEI TEMPI CICLO DI N1
grafico a dispersione



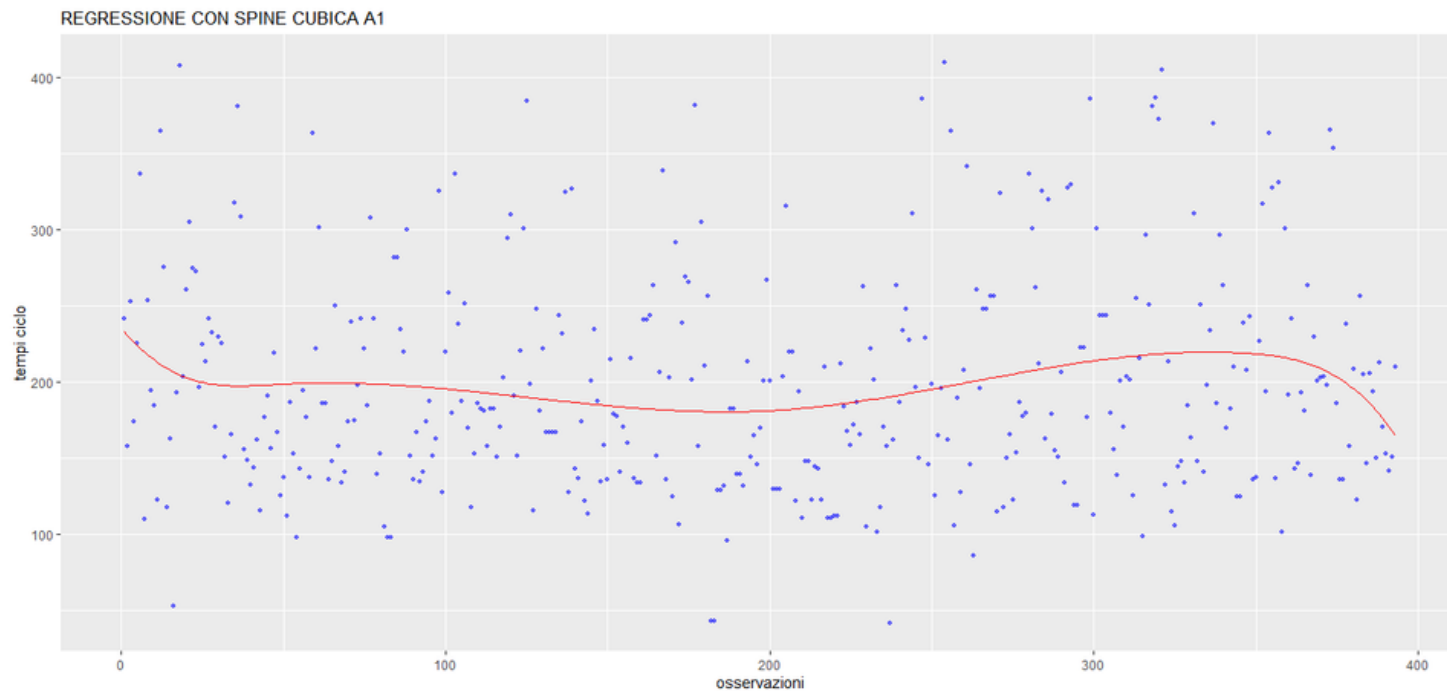
O1 REGRESSIONE POLINOMIALE DEI TEMPI CICLO DI O1
grafico a dispersione



Analisi della prima fase di produzione: regressione con spline

Applicando una spline cubica ai tempi dell'operatore A abbiamo adesso maggior certezza che il suo tempo ciclo abbia delle oscillazioni, il che significa che questo operatore non è rimasto costante nel tempo nello svolgere questa fase produttiva.

Studiando, invece, le spline degli altri operatori otteniamo praticamente delle rette, per cui possiamo considerarli come costanti nello svolgere la loro fase di lavorazione.



Conclusioni e possibili applicazioni

La seconda fase di assemblaggio è stata studiata con la stessa modalità, per cui non è stata riportata qui.

Questo studio non è nient'altro che un'Anomaly Detection nei processi industriali perché ci permette di individuare quegli operatori che, nello svolgere una determinata fase produttiva, non rimangono costanti a livello di tempo ciclo.

Infatti, una retta di regressione orizzontale significa che non c'è correlazione tra l'andamento temporale delle osservazioni ed il tempo ciclo registrato; e se anche i polinomi e le spline tendono ad essere rette significa che quell'operatore è veramente (abbastanza, entro una certa tolleranza) costante nello svolgere quella fase produttiva, che è esattamente quel che si vuole a livello industriale.

Questo tipo di analisi, quindi, unitamente ad un'analisi qualitativa del risultato, permette di individuare gli operatori più veloci ed i più costanti; potendo confrontare i risultati ottenuti in termini di tempo con quelli qualitativi (il miglior operatore è il più veloce e costante, e contemporaneamente genera un prodotto senza difetti qualitativi), si possono individuare gli operatori più "adatti" a svolgere un certa fase produttiva, aggiungendo queste informazioni, per esempio, in una skill matrix (la skill matrix mi dice le qualità dell'operatore, ma non le misura)

