

Tre applicazioni concrete di modelli deep per gestire le ispezioni visive, le telemetrie e la documentazione tecnica a supporto delle operations

Luca Paganelli, Riccardo Zanella, Mirco Nani, Luca Montanari, Alex Magrini, Simone Scala

Injenia SRL

luca.paganelli@injenia.it, riccardo.zanella@injenia.it, mirco.nani@injenia.it,
luca.montanari@injenia.it, alex.magrini@injenia.it, simone.scala@injenia.it

Abstract

Abbiamo sperimentato l'applicazione di 3 classi di modelli DNN con l'obiettivo di ottimizzare la gestione delle operazioni sugli impianti lavorando su ispezioni visuali (con classificazione automatica dello stato di usura degli elementi dell'infrastruttura) sulle fonti telemetriche (identificando allarmi e costruendo un trend previsionale sui successivi 30 minuti) e sull'ottimizzazione della ricerca sulla documentazione tecnica di impianto.

1 Introduzione

L'automazione in ambito industriale può essere descritta attraverso molteplici declinazioni cui corrispondono anche tipicamente diversi gradi sulla scala che congiunge due estremi operativi, il primo rappresentato da processi completamente manuali, il secondo caratterizzati dall'assenza quasi totale di intervento umano.

Le soluzioni realizzate a partire da tecniche di machine learning o deep learning si affacciano con crescente penetrazione in questo campo e sono in generale accomunate dal principio primo che ne fonda le basi teoriche: l'utilizzo estensivo ed efficace dell'esperienza pregressa accumulata e tracciata nell'ambito di un processo già digitalizzato.

In questo documento sono descritte 3 soluzioni complementari che indirizzano 3 scenari applicativi nel campo delle gestione delle operazioni su impianti attraverso l'applicazione di modelli di deep learning allo stato dell'arte per l'elaborazione delle immagini, l'interpretazione del testo l'analisi delle serie temporali.

2 Automazione sull'ispezione visiva

Gli impianti industriali sono tipicamente caratterizzati da elementi operazionali esposti all'usura indotta dalle

condizioni atmosferiche. Quando la distribuzione sul territorio è significativa o la "sensorizzazione" realizzabile insufficiente, l'ispezione visiva diventa la via più accurata per il monitoraggio. Le tecniche di acquisizione sono molteplici: missioni con elicottero, attraverso droni, da scansioni satellitari, nello spettro del visibile, dell'infrarosso dell'interferometria.

L'interpretazione di questo materiale e la sua automazione è oggetto della soluzione che abbiamo consolidato.

2.1 Soluzione

L'applicazione dei modelli deep in ambito di object detection ed object classification a partire da un dataset proprietario finemente annotato grazie anche al contributo esperto delle unit di manutenzione del cliente ha permesso di implementare un censimento estensivo dell'intera infrastruttura ed automatizzare la ricognizione delle componenti con criticità.

Sono stati addestrati un modello DNN su architettura Faster-RCNN-ResNET101 con la tecnica del fine-tuning su un dataset di 350.000 immagini con labeling di tipo bounding-box rettangolare per l'identificazione delle componenti ed un modello su architettura Inception ResNET V2 per l'attribuzione della sotto-classe specifica ed il livello di degrado con la tecnica del transfer-learning con 1M+ crop a partire dal primo dataset.

2.2 Risultati

L'architettura descritta permette di ottenere un'accuratezza del 75% end-to-end (composta da accuratezza nell'identificazione di tutti i componenti e nella corretta attribuzione del livello di criticità per ognuno di essi) con un tasso di recall superiore all'85%.

L'architettura di advanced analytics complessiva, basata su strumenti ML ed MPP di nuova generazione (Beam, Tensorflow, Airflow all'interno di Apache Foundation e Dataflow, BigQuery, CMLE su Google Cloud Platform) permette di sostenere un throughput complessivo di 100.000 immagini per 6 ore di elaborazione.

3 Predittore su rilevazione telemetrica

Quando la ricchezza di sensoristica permette di monitorare i sistemi e gli agenti esterni con completezza è possibile costruire una soluzione di advanced analytics di carattere predittivo che possa supportare l'operatore nella scelta degli interventi da effettuare in operations.

3.1 Soluzione

Abbiamo sfruttato il trend storico (2 stagioni, primavera ed estate) per costruire un predittore in grado di fornire in *near-real-time* un'indicazione sull'evoluzione (successivi 30 minuti) a partire dalle evidenze verificatesi nelle precedenti 24H sulle grandezze in telemetria (sovra-campionate ad 1 minuto) e sulle previsioni meteorologiche emesse a diversa granularità: a grana grossa l'allerta governativa ed a grana fine su griglia di 1 km², un'area geografica di 450 km² e orizzonte 20 minuti per la previsione della precipitazione delle acque piovane.

Il modello di ML implementato combina layer fully-connected e convoluzionali per ridurre la dimensionalità dei dati (meteorologici e telemetrici) e sintetizzare uno stato iniziale a partire dalla prima parte della finestra temporale; i successivi layer ricorrenti LSTM con architettura encoder/decoder analizzano la seconda parte della finestra temporale e predicono i successivi 30 minuti.

3.2 Risultati

L'architettura descritta permette di ottenere un MAE tra lo 0,01 e lo 0,02 dopo 100.000 step di addestramento che si traduce in un errore di stima al 95 percentile misurato per il livello dell'acqua sui nodi della rete di raccolta in un valore compreso tra i 20 cm (su un orizzonte dei successivi 5 minuti) ed i 50 cm (su un orizzonte dei successivi 30 minuti).



Figura 1: Esempio del Predittore su Serie Temporal. Su sfondo grigio le telemetrie, a destra le previsioni meteo

4 Ricerca documentale assistita con interfaccia in linguaggio naturale

La complessità impiantistica si traduce solitamente nella necessità di dover operare su un corpus documentale tecnico

ed operativo non banale. La ricerca all'interno di esso rappresenta un dispendio significativo di tempo.

L'ottimizzazione del tempo di ricerca si traduce in un beneficio complessivo per l'azienda, in primis perché è un fattore chiave nella velocità di intervento in operations.

4.1 Soluzione

In ambito NLP l'interpretazione del testo è indirizzata attraverso la costruzione di una rappresentazione geometrica del testo in uno spazio n-dimensionale (embedding).

L'approccio è stato quello di partizionare i documenti, un corpus complessivo di quasi 800 istanze, in nuclei elementari (frasi di senso compiuto) mantenendo il riferimento degli stessi all'interno del documento, "allenare" gli embedding su questi nuclei, applicare il modello alle interrogazioni in linguaggio naturale ed implementare una ricerca nearest-neighbor basata sulla cosine-similarity tra il vettore dell'interrogazione ed i vettori calcolati su tutto il corpus.

Abbiamo adottato "Gensim" (un'implementazione dell'architettura doc2vect di Le e Mikolov) attestandolo su uno spazio degli embedding ottimale pari a 300 dimensioni.

4.2 Risultati

Abbiamo ottenuto una significatività dei risultati che su un test set di circa 96 domande ha confermato la "risposta corretta" a livello puntuale di precisa nel 55% dei casi e mediamente nell'ambito dei primi 10 risultati.

Ammettendo corretti anche i risultati in cui la risposta si discosta di una sola pagina rispetto all'atteso la misura sale al 62% sul test set. Questa metrica condivisa, pur oggettiva, è tuttavia troppo rigida e severa e non valorizza pienamente le performance in termini di *recall*.

Riferimenti bibliografici (principali)

- [Mikolov *et al.*, 2014] Q. Le, T. Mikolov. 2014. *Distributed Representations of Sentences and Documents*. In Proceedings of ICML 2014.
- [Evermann *et al.*, 2016] Joerg Evermann, Jana-Rebecca Rehse and Peter Fettke. 2016. *Predicting Process Behaviour using Deep Learning*. CoRR 2016.
- [Sutskever *et al.*, 2014] Ilya Sutskever, Oriol Vinyals and Quoc Le. 2014. *Sequence to Sequence Learning with Neural Networks*. CoRR 2014.
- [Badrinarayanan *et al.*, 2015] V. Badrinarayanan, A. Kendall and R. Cipolla. 2015. *SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation*. CoRR 2015.