

TEXT MINING: VALORIZZARE I DATI NON STRUTTURATI

È UN RAMO DELL'INTELLIGENZA ARTIFICIALE CHE CONSENTE DI ANALIZZARE E NORMALIZZARE INFORMAZIONI TESTUALI. MOLTO UTILIZZATA SUI MERCATI PIÙ AVANZATI, QUESTA TECNICA È UTILIZZABILE AD ESEMPIO NEL PROCESSO VALUTATIVO E NELL'AUTOMATIZZAZIONE DELLA GESTIONE SINISTRI

Tra le più grandi sfide del settore assicurativo vi è senza dubbio quella relativa a una maggiore valorizzazione delle informazioni e dei dati raccolti. Le imprese assicurative, infatti, sono continuamente impegnate a migliorarne la qualità e la completezza per poterne sfruttare il prezioso valore intrinseco.

I modelli predittivi comunemente utilizzati nel settore assicurativo italiano si basano esclusivamente su dati strutturati, ovvero dati raccolti e memorizzati secondo uno schema predefinito, tipicamente in forma tabellare, facilmente fruibili per qualsiasi analisi quantitativa. Ad esempio, i modelli utilizzati per valutare l'evoluzione attesa dei futuri costi dei sinistri si basano su dati storici strutturati, come gli importi storicamente risarciti o da risarcire, il numero dei sinistri denunciati, il costo medio, etc.

Tuttavia, moltissime informazioni e dati raccolti dalle imprese assicurative sono in forma non strutturata, cioè raccolti e memorizzati senza uno schema predefinito e pertanto difficilmente fruibili nei modelli predittivi. Ne consegue che non sono ancora del tutto sfruttati in modo efficiente. Molto del loro potenziale valore implicito è ancora inutilizzato.

Oggi, grazie agli sviluppi della tecnologia e del digitale, nuove opportunità stanno modificando rapidamente il contesto di riferimento e la competitività tra imprese. Il forte impulso statistico-informatico consente al mercato assicurativo di analizzare dati e informazio-

ni attraverso nuove potenti tecniche, come l'*intelligenza artificiale* (IA) e le sue relative branche, in primis il *text mining* e le *machine learning*.

Il text mining è un insieme di approcci di IA, tra i più sofisticati attualmente disponibili, che analizza i dati non strutturati trasformandoli in dati strutturati e normalizzati. Le machine learning sono un insieme di complessi metodi di statistica computazionale che potenziano e migliorano la performance di un algoritmo nell'identificare *pattern* nei dati.

UNA TECNOLOGIA PER PREVENIRE IMPATTI FUTURI

Lo sviluppo delle tecniche di text mining e machine learning creano nuove opportunità nel settore assicurativo, consentendo l'uso di modelli più potenti e con maggiore capacità predittiva proprio perché calibrati sull'analisi del valore implicito dei dati non strutturati.

In Italia la gran parte dei dati non strutturati non è ancora valorizzata, ma nei mercati più avanzati queste informazioni iniziano a fornire un prezioso contributo a tutte le analisi previsionali, rafforzando e migliorando gli attuali modelli di business con notevoli vantaggi competitivi.

Studiare le connessioni, quantificare il grado di correlazione o modellare i legami tra particolari fenomeni, considerata la mole di dati non strutturati disponibili, permetterebbe di comprendere meglio, e soprattutto in



anticipo, gli impatti attesi nel futuro di diversi fenomeni e grandezze, oltre ai vantaggi derivanti dall'automatizzazione di alcuni processi, con la relativa riduzione di costi a fronte di una migliore allocazione ed efficientamento.

Sulla base dei risultati di alcuni *case study* che abbiamo effettuato nell'ambito della valutazione delle riserve sinistri, abbiamo riscontrato considerevoli benefici nell'utilizzo di dati non strutturati attraverso lo sviluppo di modelli di text mining e machine learning. Alcuni esempi delle macro categorie di dati non strutturati (elenco non esaustivo) utilizzati nel processo valutativo sono:

- stato dell'ambiente e dinamica del sinistro: tipologia di strada, condizione meteo (nebbia, ghiaccio, pioggia, etc.), dichiarazioni sulla dinamica, informazioni da mobile e dispositivi connessi;
- intervento delle autorità e verbali: carabinieri/polizia stradale, ambulanza, referti ospedalieri/pronto soccorso, etc.;
- riscontro peritale: riscontri prodotti dai periti e informazioni dei testimoni, etc.;
- stato di salute e referti post evento: interventi chirurgici, risonanza magnetica, riabilitazione, lesioni pregresse, restrizioni lavorative, etc.;
- aspetti legali: coinvolgimento di avvocato, azioni legali in corso, ricezione lettera di contestazione, etc..

Grazie al text mining si è potuto analizzare e sintetizzare moltissime informazioni non catturate nei dati strutturati, e pertanto si è potuto arricchire il set di dati solitamente utilizzato nei modelli predittivi.

AMPLIFICARE LE POTENZIALITÀ DELL'IA

Nel settore assicurativo, i modelli previsionali maggiormente diffusi sono i modelli di regressione calibrati su dati strutturati. L'innovazione consente di ampliare il *range* degli attuali modelli previsionali, offrendo alle imprese assicurative l'opportunità di introdurre nuovi e

potenti algoritmi di IA, in grado di sfruttare il prezioso valore intrinseco delle informazioni provenienti sia dai dati strutturati sia da quelli non strutturati.

Ci sono diversi ambiti di applicazione del text mining nel settore assicurativo, ma uno degli ambiti di maggior interesse è certamente quello legato alla capacità di prevedere in anticipo l'evoluzione attesa del costo dei sinistri, soprattutto quelli particolarmente complessi caratterizzati da una bassa frequenza ma con un elevato costo medio.

In conclusione, le imprese di assicurazione posseggono informazioni e dati che non sono ancora del tutto utilizzati e pertanto non valorizzati in modo efficiente. Informazioni e dati il cui prezioso valore implicito potrebbe essere estrapolato attraverso l'utilizzo di modelli di text mining e sfruttato con algoritmi di machine learning, offrendo alle imprese assicurative nuove opportunità di innovare e migliorare le loro strategie di business in un contesto sempre più *data-driven*.

Milliman, tra le società leader nella consulenza attuariale e strategica a livello internazionale, ha una profonda conoscenza del settore assicurativo ed è pioniera nel campo della *data scientist* e dell'*advanced analytics*. Milliman ha un team dedicato di consulenti e data scientist nello sviluppo e nell'implementazione degli approcci innovativi di IA in ambito assicurativo, come ad esempio il text mining e modelli di machine learning, e supporta numerose compagnie in tutto il mondo. L'applicazione effettiva degli algoritmi predittivi di IA sviluppati dal team di Milliman trova riscontro in diversi ambiti, compresi quelli legati alle valutazioni delle riserve sinistri, della tariffazione dei prodotti assicurativi e della competitività del prezzo. I traguardi ottenuti nelle applicazioni pratiche confermano il potenziale valore derivante dall'uso dell'IA e l'opportunità di potenziare sensibilmente la capacità predittiva degli *advanced analytics model*.

Milliman ha sviluppato Milliman Nodal™, la piattaforma di analisi predittive leader nel mercato assicurativo che usa il text mining e altre tecniche di IA nella gestione dei sinistri. Per approfondimenti visitare il sito www.milliman.com/en/products/nodal.

Per approfondimenti e ulteriori informazioni: nicola.biscaglia@milliman.com o italy.info@milliman.com.