

Immobili

Il settore immobiliare visto alla luce dei Big data

27 Gennaio 2022

Quelli del mondo della “casa” sono molteplici e non strutturati, mostrano limiti e criticità, comunque superabili attraverso adeguati strumenti tecnologici e competenze in grado di estrarne la conoscenza



Tecniche di *Data Mining* per estrarre conoscenza dai *Big data* del settore immobiliare: su questo tema, Maurizio Festa, responsabile dell'ufficio Statistiche e Studi sul mercato immobiliare dell'Agenzia delle Entrate, ha pubblicato un'interessante **studio** sull'ultimo numero di **Territorio Italia**. Una riflessione sapiente e documentata sull'equilibrio necessario all'uso efficace e responsabile di

questi strumenti innovativi.

Si tratta di un importante contributo alla comprensione del fenomeno dei *Big data*, che offre un quadro d'insieme sullo stato dell'impiego di metodologie di *Data mining* e *Machine learning* nel settore immobiliare, non esente da limiti e criticità, e sulle nuove figure professionali, che, in quest'ambito, occupano un ruolo chiave: i *Data scientist*, tra i più ricercati, oggi, nel mercato del lavoro.

Big data e il valore delle “V”

Nell'ultimo decennio, abbiamo assistito a un incremento esponenziale dei volumi dei dati raccolti. Una mole crescente: in volume (si affaccia il termine *zettabyte*), in varietà (molte le “fonti”, *human generated* - da *social network* e *blogging*, eccetera, *machine generated* - da sensori e strumenti scientifici, *business generated* - acquisiti nei processi di produzione o consumo di beni e servizi; e i

“formati”, testo, video, immagini, audio, eccetera; dati strutturati, semi-strutturati o non strutturati); in velocità dei flussi dei dati (generati ed elaborati), in *veracity* (veridicità e accuratezza dei dati) e nel valore che ne discende.

I *Big data*, soprattutto grazie alle interazioni di miliardi di utenti Ict, sono arricchiti dagli attributi descrittivi di ciascun dato, producendo un risultato statistico più ricco e più tempestivo.

I cosiddetti sensori intelligenti, collegati alla rete, generano flussi continui di dati che, opportunamente lavorati, diventano una risorsa economica vera e propria.

Questa ricchezza di conoscenza offre la possibilità di processi decisionali legati proprio ai *Big data* (Ddd - *Data driven decision making*).

Data science e Machine learning

Nuovi database, nuove tecniche di analisi, ma anche nuove tecniche di selezione e metodologie interpretative dei risultati, sono il risultato evolutivo dell'estrazione dai dati della conoscenza dei fenomeni socio-economici. Un processo che prende il nome di *Knowledge discovery in data* (Kdd) e si compone di cinque fasi: *Selection* (dove si definiscono le origini dei dati), *Data cleaning e pre-processing* (dove si migliora la qualità del set di dati), *Transformation* (dove si analizzano gli attributi e le relazioni dei dati), *Datamining* (dove si estraggono strutture, correlazioni, modelli e regole all'interno dei dati), *Evaluation/Interpretation* (dove si valuta la validità dei modelli, si interpretano per raggiungere l'efficacia del processo).

Dunque, accanto al classico modello statistico, si affaccia una più ampia disciplina che riguarda la Scienza dei dati (*Data science*) e che permette di “incrociare” tutte le caratteristiche dei *Big data*, analizzarle e metterle in relazione, elaborando modelli complessi.

Le modalità di lavoro cambiano nettamente e richiedono nuove professionalità: una squadra di *Data scientist* che abbia variegate competenze e attitudini.

Le tecniche di analisi, infatti, prevedono l'uso di algoritmi, con cui rispondere a interrogazioni precise.

In questo ambito, il *Machine learning* si configura con un approccio induttivo e, nel tempo, la grande mole di dati permette all'algoritmo di “imparare” da essi e collocarsi nell'ambito dei modelli previsionali.

Le *Reti neurali* usano altri algoritmi per estrarre nuove conoscenze grazie all'intelligenza artificiale, in grado di processare in modo proficuo ed efficiente i *Big data*, effettuando classificazioni, associazioni e *clustering*.

L'applicazione all'ambito immobiliare: vantaggi e criticità

I *Big data* del settore immobiliare sono molteplici e non strutturati. Derivano dalle interazioni tra i soggetti che entrano in relazione in questo settore, come ad esempio: proprietari, costruttori, acquirenti, locatari, intermediari d'affari, intermediari finanziari, agenzie fiscali, periti, enti territoriali.

Sono emerse, così, piattaforme con finalità differenti, che orbitano intorno al mondo immobiliare. Alcune sono orientate a capire i "gusti" dell'utente, in modo tale da offrire a ciascuno l'immobile che soddisfi meglio le sue esigenze; altre si muovono all'interno delle dinamiche del mercato e offrono, ai potenziali investitori, informazioni sul momento più favorevole per comprare o vendere; altre ancora riguardano modelli di valutazione automatizzati, permettendo agli investitori di minimizzare i rischi.

Tuttavia, non mancano criticità e limiti che è doveroso evidenziare.

L'accesso a questa mole enorme di dati/informazioni/conoscenze/elaborazioni è riservata a pochi, con il rischio di nuove disuguaglianze. A questo proposito, la Pubblica amministrazione è impegnata nel rendere fruibili con libero accesso *data-set* delle informazioni raccolte, sia in forma grezza che elaborata (*open data*).

Vi è, inoltre, il tema della tutela della *privacy*: in questo ambito i dati sono messi in relazione come in una centrifuga, per cogliere le connessioni a qualsiasi livello. Infatti, non è semplice creare un insieme di dati realmente anonimo a partire da un ampio insieme di dati personali e, al tempo stesso, mantenere le informazioni sottostanti necessarie agli scopi per cui sono state raccolte. D'altronde, rendere anonimi i dati può inficiare l'utilità del dato stesso, rendendolo "incerto" (in quanto non riferibile a nessuno) e non idoneo a qualsiasi tipo di analisi. Tuttavia, nel rispetto del nuovo regolamento europeo, la filosofia della *privacy by design* consente la tutela fin dalla progettazione e per l'intero ciclo di vita dei *Big data*, integrando da subito misure per l'adeguata protezione dei dati personali.

Inoltre, i processi di estrazione di conoscenza attraverso tecniche di *Data mining*, ci spiega l'autore, rappresentano una sorta di "nuova forma di empirismo" che si allontana dalla tradizionale formulazione di ipotesi, costruzione di modelli, verifica sperimentale su dati e simulazioni, infine formulazione di una teoria, prediligendo talvolta modelli "opachi", cosiddetti *black box*, che possono sfuggire anche agli stessi programmatori e sviluppatori.

Un altro concetto da esplorare con attenzione è quello della rappresentatività dei *Big data*, che potrebbe apparire indiscutibile, vista l'enorme quantità di informazioni che contengono, eppure la loro esaustività non è scontata.

Infine la presenza di errori nei dati non è da sottovalutare: l'enorme numero di dati processati potrebbe mostrare correlazioni forti ma, allo stesso tempo, spurie, cioè non necessariamente associate a un rapporto causa-effetto.

Quanto al settore immobiliare, vi sono grandi potenzialità di sviluppo che possono derivare da modelli *data driven* di nuova generazione, potendo apportare quella conoscenza necessaria su cui fondare quelle decisioni politiche, economiche e sociali determinanti in quest'ambito. L'urbanistica, l'industria turistica, il *social housing*, il consumo energetico degli edifici, il contenimento di consumo di suolo, i nuovi progetti di sviluppo edilizio sono solo alcuni dei temi nei quali i *Big data* del settore immobiliare potranno essere proficuamente analizzati.

Gli strumenti tecnologici necessari a gestirli e le competenze necessarie per estrarne la conoscenza e comprendere i fenomeni e la loro causalità sono le priorità. È inoltre fondamentale comprendere che questa nuova frontiera non sostituisce i modelli classici, bensì, con l'opportuna interpretazione, va a integrarsi a essi.

Dalla descrizione dell'autore emerge come i *Data scientist* debbano maturare un corretto approccio ai dati, scrupoloso e consapevole dei punti di forza e di debolezza del metodo, per restare protagonisti della conoscenza. Il panorama futuro promette l'accesso a conoscenze sempre più raffinate, purché gestite da professionisti consapevoli che un'overdose informativa non è il risultato atteso.

I *Big data* rappresentano senza alcun dubbio uno strumento di ricerca rivoluzionario, ma, mette in guardia l'autore, per costruire processi di *Data mining* efficienti, bisognerà continuare ad affrontare i temi ancora aperti: affinare le competenze interpretative e di analisi, ricercare il giusto equilibrio alla trasparenza e alla *privacy*, spingere sull'espansione degli *open data*, migliorare le tecniche di selezione dei dati, eliminare l'opacità dei modelli *black box*, gestire il ciclo di vita dei dati.

continua

La prima ***puntata*** è stata pubblicata giovedì 13 gennaio

di

Rosa Colucci

URL: <https://www.fiscooggi.it/rubrica/immobili/articolo/settore-immobiliare-visto-alla-luce-dei-big-data>