

Con i big data il CRM è morto. O no?

Giuseppe Zavattoni
UniCredit

Viene descritto il percorso evolutivo dei processi CRM nel tempo, partendo dai primordi, per arrivare al campaign management in real time mediante tecnologie di ultima generazione quali i big data.

In generale ho la necessità di fermarmi e “fare il punto”. Amo capire i fatti del passato che mi hanno accompagnato lungo un percorso. Ho la personale esigenza di capitalizzare, raccogliere le energie e guardare avanti, in modo consapevole. La sfera professionale non sfugge a questa mia forma mentis. Prima di affrontare un nuovo progetto, mi volto, guardo quanta strada ho percorso, cosa ho messo nel mio zaino professionale e quanto di quel bagaglio mi aiuterà ad affrontare una nuova salita.

È stato questo anche il mio approccio molto particolare verso il mondo dei cosiddetti big data.

Ho incontrato addetti ai lavori che mi suggerivano un colpo di spugna. Volevano spazzare via tutto: “Giuseppe, il CRM è defunto”, dicevano, “adesso ci sono i big data, faranno tutto da soli. Ripartiamo da zero. Start from scratch”.

Calma. Fermiamoci, pensiamo.

Mi sono chiesto se l’esperienza di questi vent’anni di CRM fosse proprio da cancellare, o piuttosto rappresentasse una solida base da traghettare, in modo da valorizzarne le eccellenze e migrarle verso la verde vallata dei big data.

Penso a quell’insieme di esperienze, risorse e professionalità senza le quali un sistema di CRM non regge, non sta in piedi. Questo blocco di risorse umane e tecnologiche è in grado di orientarsi verso un nuovo percorso evolutivo, senza per questo generare una separazione dalla propria storia.

Fin dall’inizio (fig. 1) abbiamo puntato sugli analytics, il data mining, sul targeting sofisticato che porta a costruire campagne ipersegmentate. Un certo grado di successo delle iniziative ha provocato la crescita nella domanda interna. Questa ha fatto nascere l’esigenza di un’automazione dei processi di campaign management. Non solo, abbiamo puntato anche all’industrializzazione del machine learning, trasformando temi considerati di nicchia, o accademici, per spostarli verso un approccio di produttività industriale, pur garantendone un’elevata qualità.

Automazione quindi, che ci ha trascinato a progettare campagne ricorrenti giornaliere, gli event trigger. Stavamo passando da un approccio che assegnava ai clienti probabilità di acquisto di un prodotto, a un sistema che iniziava a spostare l'attenzione verso l'ascolto dei bisogni del cliente. Spostarci quindi da una situazione dove in maniera statisticamente presuntuosa affermavamo che il cliente Marco Bianchi, molto probabilmente aveva necessità di una carta di credito, al più umile ascolto di Marco che ci stava dicendo che era alla disperata ricerca di una buona carta.

Ci sembrava di aver raggiunto un traguardo di assoluto rilievo, quando un Direttore Marketing mi disse che era tutto molto interessante, ma che ci stavamo perdendo delle grosse opportunità perché i contatti non erano in tempo reale (fig. 2).

Avremmo dovuto intercettare i bisogni dei clienti nel momento in cui si fossero palesati. In tal modo avremmo identificato i momenti chiave sul loro ciclo di vita e selezionato il migliore canale per contattarli.

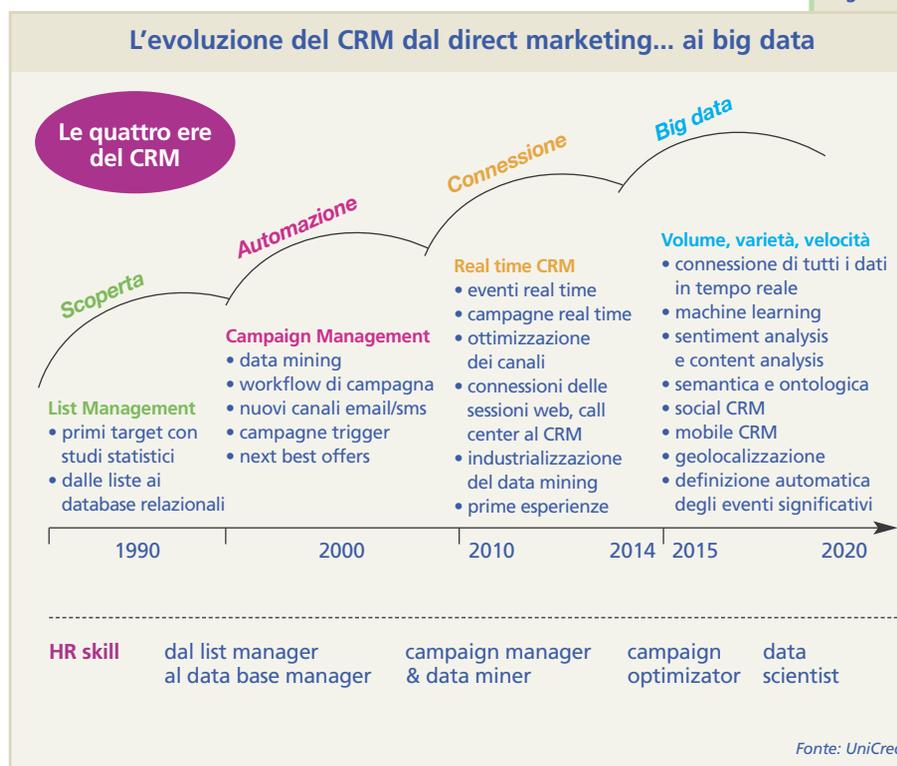
Quel commento nascondeva qualcosa di più dell'invito a creare un sistema di contact management in real time, multicanale. Era la richiesta di una svolta, di un cambio di mentalità. L'abbiamo realizzato, il **real time c'è**. Da tre anni è una realtà che funziona e porta risultati.

Da subito però ne abbiamo intuito i limiti. Due i più importanti.

Il primo riguarda la **definizione di un "evento"**, inteso come la ragione per la quale scatenare un contatto in tempo reale con un cliente al fine di proporre un'offerta orientata a soddisfare il bisogno che l'evento stesso fa emergere. Abbiamo costruito una ricca libreria di eventi partendo da regole derivate dall'esperienza del business bancario, commerciale, dal buon senso e dal mestiere. Ci siamo chiesti se esistesse la possibilità di migliorare, automatizzare le definizioni, industrializzare. Forse sì. Un altro limite è rappresentato dai costi e dal time to market della realizzazione dei processi informatici per l'alimentazione di ogni singolo evento. I processi automatici di campaign management real time devono passare dai sistemi gestionali della banca, risultano quindi onerosi in termini di costo realizzativo e nei tempi di rilascio. Ci chiedevamo come abbattere i costi, essere più flessibili e veloci nei rilasci. Possibile? Forse sì.

Intercettare i bisogni del cliente in tempo reale per identificare i momenti chiave sul suo ciclo di vita

fig. 1



**Nuove figure professionali
con skill specifiche
al passo con la tecnologia
dei big data**

Siamo dunque di fronte a metodologie e tecnologie cresciute e evolute nel tempo. Come del resto anche gli skill delle risorse umane, i professionisti del CRM. In origine c'erano i list manager, supportati dagli analisti, poi diventati data miner, che lavoravano assieme ai campaign manager, trasformati in campaign optimizer. E ora, all'apice di questo processo evolutivo ci troviamo di fronte a un nome altisonante: i "big data!". Gestiti e utilizzati da una nuova figura professionale, il "Data scientist", che dovrebbe riunire nella stessa persona un informatico, uno statistico e esperto di business (ma è davvero una professionalità così nuova?).

I big data, i big data. Prendiamoci un paio di minuti per fissarne alcuni aspetti. Innanzitutto dobbiamo mettere subito in chiaro che si tratta di una tecnologia, molto furba, ma che non sconfinava nella magia, o nell'esoterismo. Molti li

hanno descritti con la calzante etichetta delle "tre V": volumi, velocità e varietà. Per quanto riguarda i primi due aspetti, ho toccato con mano. Grazie al calcolo in parallelo e "in memory" trattano volumi incredibili di dati con una velocità sorprendente. Non ho trovato lo stesso entusiasmo per gli algoritmi di **machine learning**. Malgrado molti ne decantassero doti miracolistiche, nella realtà non portano discontinuità nei risultati in termini statistici. Sono "solo" maledettamente veloci e (questo fa il caso nostro) funzionano molto bene se applicati ai dati transazionali, allo streaming carte, POS, Atm, bonifici, log internet, mobile banking, contact history, agenzie e call center.

fig. 2

L'evoluzione in sintesi

Ieri

- Calcolo delle probabilità d'acquisto e estrazione liste clienti
- Invio prevalente di target list in agenzia e contact center

Oggi

- Intercettare i bisogni quando si palesano e invio in tempo reale ai canali
- Identificare i momenti chiave della vita del cliente
- Scegliere i migliori canali di contatto ottimizzati per il cliente
- Fornire maggiori informazioni sulla relazione banca/cliente ai colleghi in rete



Fonte: UniCredit



Insomma è possibile costruire modelli di machine learning in real time, utilizzando i dati transazionali della banca, associati anche a dati più statici come i saldi, il cross selling, la profittabilità. Inoltre, grazie alla potenza di calcolo dei big data si aprono opportunità nel trattare anche dati non strutturati, come le causali di bonifico, le note di contatto degli operatori di rete e call center, i merchant dei pagamenti POS. In questo caso ci viene in aiuto una ricca famiglia di algoritmi per il natural language processing (NLP) che ci consentono di estrarre semantica dai dati, effettuare una sentiment analysis dei nostri clienti, comprenderne il mood. Se ci pensiamo bene, tutte queste transazioni rappresentano le informazioni che utilizzavamo per creare i nostri eventi real time.

Dall'approccio deduttivo alla logica induttiva

A questo punto avevamo tutto il materiale necessario per un nuovo cambio di paradigma, per rendere più economica, veloce e di sicuro ancora più efficace la creazione degli eventi real time.

Bisognava passare dalla logica deduttiva, o top down, fin qui utilizzata, alla logica induttiva, o bottom up, normalmente utilizzata per lo studio dei modelli di machine learning (fig. 3). Eravamo e siamo convinti che la combinazione dei risultati portati dai contatti in tempo reale, associati a una definizione degli eventi algoritmica, e non più di buon senso, avrebbe generato una forte ulteriore discontinuità. Ragionando in modo deduttivo, l'esperienza e la logica ci dicono che se un cliente effettua un prelievo su un conto non capiente, molto probabilmente necessita di un finanziamento. Partiamo dunque da un'idea, ne cerchiamo il riscontro nei dati e una volta che il fatto si presenta, inviamo un'offerta. Ma quante regole ci perdiamo perché fuori dalla nostra conoscenza, dal nostro dominio umano che non può gestire l'associazione multivariata di centinaia di variabili? La logica induttiva vuole risolvere questo problema. Appunto parte dal basso, dai dati e, a fronte di un obiettivo (es. necessità di un prestito), impara dai dati, scopre regolarità che identificano il bisogno all'interno dei database.

Ecco la novità: con i big data questo tipo di approccio bottom up è applicabile in tempo reale. Viene dapprima elaborato un algoritmo studiando i dati storici nei quali vengono inseriti i dati transazionali (bonifici, carte, log, ...). L'algoritmo viene alimentato dallo stream di dati in real time e produce come output la probabilità che un cliente sia propenso alla sottoscrizione di un determinato prodotto. Se questa probabilità, o scoring, è superiore a una determinata soglia, entra in azione il tool di campaign management che si occupa di contattare il cliente con un'email o un SMS. Il tutto in un pugno di secondi dopo il verificarsi della transazione che ha provocato l'evento.

Questa prospettiva, dove i big data si sono rivelati essenziali, risolve il problema dell'identificazione degli eventi secondo regole e abbatte costi e lentezza nella produzione dei dati. Infatti, soddisfatto il pre-requisito dell'alimentazione continua dello stream transazionale per la determinazione degli eventi, ciò dipende solo dalla disponibilità dei relativi modelli di machine learning. I modelli sono costruiti al nostro interno dai nostri Data scientist, con le stesse metodologie utilizzate oggi per lo sviluppo degli scoring batch. Siano questi di propensione all'acquisto dei più disparati prodotti, di churn, o anche che si verifichi un determinato accadimento sul ciclo di vita del cliente.

Gli strumenti

Dicevo metodologie, non strumenti d'analisi, o software analitico. In questo caso i big data ci hanno aperto la strada all'utilizzo di software open source. Sì, anche



I modelli di machine learning sono costruiti dai Data scientist della banca

in banca. Devo ammettere che i nomi degli strumenti big data disorientano. Sembra di finire in uno zoo: Pig, Impala, Hive, senza dimenticare l'elefantino

fig. 3

I tempi sono maturi per un altro cambio di paradigma: dall'approccio top down a quello bottom up



Fonte: UniCredit

Hadoop. Una scelta importante è stata quella di utilizzare la libreria di machine learning presente in Spark tramite un software open source di machine learning sviluppato da una società legata all'Università di Costanza. Grazie a questo package è possibile gestire dei workflow che partono dal dato e arrivano allo scoring scrivendo poco codice e lavorando a oggetti che, assemblati in un workflow, rappresentano la logica dei big data e del processo analitico, in modo trasparente all'analista. Sottostante abbiamo Hadoop/Impala, ma per chi ci lavora potrebbe esserci Oracle, DB2 o un foglio excel. Il big data è demitizzato, si conservano velocità, gestione dei volumi di dati e, soprattutto, lo streaming. Anche in questo caso abbiamo privilegiato gli aspetti di industrializzazione dei processi. Vogliamo che i colleghi si concentrino sul problema, sulla statistica e non troppo

Grazie ai **big data** la banca è già proiettata verso il futuro con prodotti tecnologicamente avanzati

sull'informatica. La relativa facilità nella generazione delle analisi le rende replicabili, stabili, distribuibili. In generale preferisco avere lo "zero virgola" in meno in un indicatore statistico, ma processi stabili, veloci e efficienti.

Gli sviluppi

Partiremo prestissimo con due prototipi, già testati in ambiente di sviluppo: uno su un prodotto di finanziamento e un altro su un prodotto assicurativo. Su questi abbiamo sperimentato una serie di algoritmi appartenenti alle librerie Spark quali random forest, logistic regression, naive bayes, k-means. Qui lo sforzo è stato quello di garantire robustezza e qualità nel funzionamento in streaming garantendo una connessione ottimale con i processi di campaign management in tempo reale.

Stiamo già pensando ai prossimi scenari che si stanno aprendo grazie ai big data. Un recommender di prodotti bancari è già in produzione (utilizza lo stesso algoritmo di Spotify per consigliare musica), stiamo riflettendo sul customer journey, sentiment analysis, geolocalizzazione, pricing, black box installate sulle auto grazie alle assicurazioni RCA, ... Ci sarà di che divertirsi.

Forse è il caso di affermare che grazie ai big data il CRM è più vivo che mai. Che dite?

Giuseppe Zavattoni, CRM – UniCredit

