

## Text Mining e Customer Relationship Management

*Il successo di una applicazione di **Customer Relationship Management** basata sulla tecnologia di **Text Mining** nasce dal saper coniugare i vantaggi offerti dai tradizionali sistemi di **CRM analitico** con il patrimonio conoscitivo e informativo disponibile in tutte le nuove forme di comunicazione offerte da Internet.*

“Avanzati tools di **Text Mining** aiutano i sistemi di **CRM** a vedere meglio i clienti”

### Lo scenario

La crescita esponenziale del Web e dei suoi utilizzatori avvenuta negli scorsi anni ha dato accesso a nuove fonti di conoscenza.

Un immenso patrimonio informativo e conoscitivo che è racchiuso all'interno di documenti, articoli, report di mercato, studi e ricerche, etc. o presentato sotto forma di dialogo all'interno degli spazi aggregativi che sul Web prendono il nome di “comunità virtuali”.

Se consideriamo che solo in Italia sono attivi più di 400 newsgroup e più di 60.000 mailing list con più di 2000 iscritti e che il fenomeno è in continua espansione, possiamo comprendere quanto questa forma di interazione sia importante non solo socialmente e culturalmente ma anche come miniera di informazioni quasi infinita.

*E-mail, chat line, newsgroup, forum e blog* rappresentano per le aziende una nuova via ancora da esplorare ricca di indicazioni circa le opinioni, i gusti, le preferenze, le attitudini, dei consumatori non solo rispetto ai propri prodotti e alla azienda, ma anche in relazione ai concorrenti e ai loro prodotti.

### Il problema

L'analisi dei testi richiede strumenti efficienti ed automatici, in mancanza di essi la maggior parte delle comunicazioni non strutturate non è nemmeno letta.

Utilizzare approcci manuali non è pratico richiede molto tempo, può essere molto costoso e i risultati possono essere inconsistenti.

Anche le più evolute applicazioni di *Customer Relationship Management* analitico, che comprendono tecnologie di *Data Mining* per l'analisi dei dati e sistemi *Olap* per la presentazione dei risultati operano con dati strutturati a cui è possibile attribuire un significato univoco predefinito, ed un insieme di valori noti (es. sesso, età, etc. del cliente).

Tale restrizione esclude quindi tutte quelle informazioni non strutturate contenute nei documenti testuali o presentati sotto forma di dialogo dall'infinità di fonti pubbliche disponibili su Internet.

### La soluzione

L'attività di sfruttamento della conoscenza contenuta in tutte le nuove forme di comunicazione offerte dal Web ora è possibile grazie all'applicazione del *Text Mining*.

L'evoluzione delle tecniche e delle metodologie utilizzate in questo campo hanno subito una profonda trasformazione nel corso degli ultimi 10 anni, oggi giorno si basano essenzialmente su metodi per il trattamento del linguaggio naturale (*Natural Language Processing*), *Reti Neurali* e *Intelligenza Artificiale*.

Il *Text Mining* consente di:

- Identificare le entità contenute nei documenti analizzati e le loro relazioni;
- Riassumere e classificare automaticamente i documenti trattati;
- Individuare, attraverso strumenti di clustering, la presenza di pattern o ricorrenze tra i diversi documenti che altrimenti sfuggirebbero alla capacità analitica di strumenti manuali di analisi dei testi.

“Il **Text Mining** per analizzare, classificare e raggruppare in modo automatico migliaia di testi in pochi secondi”

## Text Mining e Customer Relationship Management

“Un approccio semantico per estrarre conoscenza dai testi”

### Il processo di Text Mining

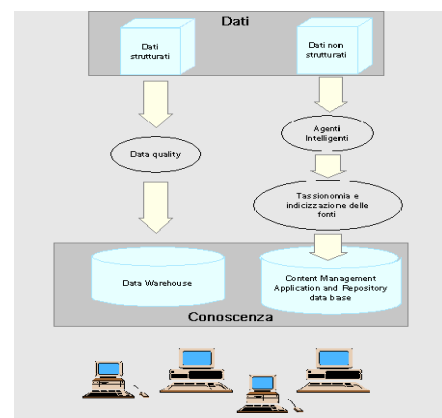
Il processo di *Text Mining* si articola nelle seguenti tre fasi di analisi:

#### Document pre-processing:

- Definizione delle regole necessarie per l'acquisizione e la raccolta dei documenti;
- Definizione e identificazione del formato dei documenti;
- Normalizzazione del testo attraverso la cancellazione degli elementi specifici del formato e riconduzione di date e valute e altre caratteristiche locali ad un unico standard;
- Riduzione e trasformazione dei testi (si eliminano le stop words, si identificano i nomi propri).

#### Lexical processing:

- Scelta delle unità da analizzare (parole, espressioni di più parole e termini);
- Definizione delle regole grammaticali e semantiche per risolvere le ambiguità del testo;
- Esecuzione dell'analisi linguistica e lessicale (lemmatizzazione, ricerca delle parole chiave e altri tagging);
- Definizione delle categorie semantiche specifiche all'argomento trattato per la ricerca delle "parole chiave";
- Classificazione delle informazioni estratte secondo concetti e/o altri metadati.



#### Text Mining processing:

- Classificazione e categorizzazione dei testi trattati;
- Estrazione di conoscenza dai dati (si identificano le entità, per esempio nomi di persone o di società, dati numerici, date, ecc e i loro legami);
- Visualizzazione tecnica dei risultati;
- Integrazione risultati del *Text Mining* con i dati derivanti dal processo di *Data Mining*.

### I vantaggi

Le informazioni ottenute dalle diverse fasi di analisi agevolano le aziende a:

- Pianificare le decisioni e l'implementare attività strategiche;
- Prevenire situazioni future di rischio;
- Conoscere meglio i clienti, i competitor, i fornitori, i partners;
- Captare segnali deboli (annuncio di possibili cambiamenti)
- Identificare nuovi potenziali clienti;
- Valutare la posizione competitiva;
- Identificare i Trend del settore o di mercato;
- Identificare tematiche riguardanti i prodotti evidenziando possibili problemi e cambi di tendenze.

“Il Text Mining per incentivare e sviluppare i Most Growable Customer e fidelizzare i Most Valuable Customer”