

Perché SAS®

- **Organizzazione, accesso, reperibilità e condivisione delle informazioni più rapidi ed efficienti.** Il proliferare della globalizzazione e l'ampliarsi del panorama competitivo costringono le organizzazioni ad abbandonare la reattività a favore della proattività. La globalizzazione, inoltre, impone la gestione, la cooperazione e l'amministrazione di contenuto in più lingue. SAS Text Analytics consente alle organizzazioni di sfruttare appieno i contenuti a disposizione e di garantire il riutilizzo di repository di dati disparati, a prescindere dal proprietario del contenuto e da dove è stato generato.
- **Pulizia del caos a livello di contenuto che coinvolge molteplici repository aziendali con documenti duplicati o leggermente modificati.** SAS Text Analytics consente di effettuare estrazioni da un vasto insieme di documenti e focalizzarsi sul contenuto in modo da riconoscere ciò che è stato duplicato o leggermente modificato e pubblicato nuovamente. Ciò consente un efficace controllo delle versioni e un migliore utilizzo del sistema e delle capacità di memorizzazione per tutto il ciclo di vita del contenuto, dalla creazione o acquisizione fino alla conservazione e alla fruizione.

- **Impiego di tecniche linguistiche e statistiche avanzate per scoprire tendenze e pattern nel testo.** Applicando l'elaborazione del linguaggio naturale, regole linguistiche e tecniche statistiche avanzate, SAS Text Analytics è in grado di trasformare automaticamente dati testuali in informazioni fruibili, individuando relazioni o associazioni esplicite.
- **Capacità di percepire e distinguere i sentimenti positivi da quelli negativi in tempo reale nella moltitudine confusa di informazioni disponibili sul web e in altri social media.** Al ritmo frenetico con cui oggi viene pubblicato nuovo contenuto nel mondo dell'informazione, la necessità di conoscere in tempo reale i trend a livello di sentimenti è fondamentale. SAS Text Analytics consente di distinguere fra sentimenti positivi e negativi in più prodotti estraendo automaticamente i veri sentimenti espressi online.
- **Soltanto SAS offre un framework analitico integrato** che copre l'intero processo di creazione dell'intelligence, dall'integrazione e gestione dei dati, alle analisi avanzate e alla distribuzione efficiente di informazioni e report significativi attraverso soluzioni di business e di settore specializzate.

Chi è SAS

SAS è leader a livello di software e servizi di business analytics e il più grande fornitore indipendente nel mercato della business intelligence. Grazie a soluzioni innovative all'interno di un framework integrato, SAS aiuta i clienti di oltre 45.000 aziende a migliorare le performance e offrire valore, accelerando il processo decisionale. Dal 1976 SAS da ai propri clienti in tutto il mondo THE POWER TO KNOW®.

www.sas.com/italy



SAS® Text Analytics

Massimizzare il valore insito nei dati non strutturati



La crescente domanda di informazioni e l'incremento esponenziale dei volumi dei dati sono problematiche rese ancora più complicate dalla varietà dei tipi e dalla complessità dei dati da gestire, inclusi dati semistruzzurati e non strutturati come e-mail, log di messaggi istantanei, corrispondenza, pagine web, blog, ecc.

Per ricavare il massimo valore dai dati aziendali, è indispensabile poterli analizzare. Tuttavia, a causa dell'ambiguità e della molteplicità dei modi in cui uno stesso testo rappresenta concetti simili, le informazioni insite nei dati testuali non sempre sono facili da distinguere, quantificare, analizzare o sfruttare. Inoltre, numerose organizzazioni non sono in grado di combinare le informazioni testuali con i propri dati strutturati e risulta quindi impossibile ottenere una visione completa e precisa dell'azienda. Ne consegue che decisioni importanti vengono spesso prese senza disporre di informazioni complete.

SAS Text Analytics, che sfrutta la tecnologia Teragram, offre soluzioni all'avanguardia per estrarre informazioni contestualmente pertinenti, analizzarle e strutturarle in modo da renderle più facilmente reperibili e da rivelare pattern, sentimenti e relazioni precedentemente celati.

Sfide

- **Organizzazione delle informazioni e relativo accesso.** I volumi di dati aumentano con un ritmo incessante. Gestirli in modo efficace ed efficiente richiede una struttura organizzativa accessibile sottostante.
- **Proliferazione di social media e di documenti elettronici.** Analizzare le informazioni provenienti dal web è sempre stato un compito particolarmente impegnativo. Se non dispongono di tecniche finalizzate all'analisi di grandi volumi di contenuto digitale in modo automatico, le società tendono a ignorarlo o semplicemente ad accantonarlo, anziché sfruttarne al massimo il valore.
- **Ricerca e reperimento inefficaci.** Dalle ricerche spesso si ottengono quantità enormi di materiali poco significativi che richiedono una revisione manuale per determinarne l'eventuale utilità.
- **Difficoltà di analisi dei dati testuali.** Proprio per l'ambiguità e le molteplici modalità di rappresentazione dei concetti, i dati testuali non sono facili da distinguere, quantificare o analizzare.
- **I dati testuali non sono integrati con i dati strutturati.** Se le informazioni testuali non possono essere integrate con altri dati, è impossibile ottenere una visione completa e precisa della società o di specifiche situazioni.



Estrarre informazioni e creare conoscenza da documenti testuali

Vantaggi

- **Eliminare il caos a livello di contenuti** nella propria organizzazione mantenendo e distribuendo soltanto le informazioni pertinenti, accurate e aggiornate.
- **Scoprire relazioni e tendenze esplicite** in dati non strutturati e semistrutturati per un processo decisionale più informato.
- **Analizzare le informazioni provenienti dal web, da siti di social networking e da documenti elettronici interni** per comprendere i sentimenti espressi e il loro effetto sui profitti.
- **Restare aggiornati su tendenze, ricerca e posizionamento competitivo** del mercato valutando automaticamente i materiali interni ed esterni.
- **Essere proattivi individuando precocemente eventuali problematiche relative a prodotti o servizi** prima che influiscano negativamente su vendite e soddisfazione del cliente.
- **Classificare accuratamente il contenuto** per consentire un'efficiente organizzazione, accesso, distribuzione e reperibilità delle informazioni.
- **Ottenere una visione più completa delle problematiche organizzative** creando una struttura per l'analisi contestuale e l'integrazione con dati strutturati.

Funzionalità

Snellire il processo aziendale che coinvolge grandi volumi di contenuto in più lingue

SAS Text Analytics utilizza tecniche avanzate di elaborazione del linguaggio naturale (Natural Language Processing - NLP) e linguistiche, offerte da Teragram, per classificare e analizzare automaticamente grandi volumi di contenuto per entità, fatti ed eventi. Tali tecniche sono poi utilizzate per creare metadati, sviluppare tassonomie e generare regole e definizioni di concetti applicabili a vaste raccolte di documenti e impiegate per implementare processi aziendali in tempo reale.

Scoprire ed estrarre nuove informazioni da documenti testuali

Una vasta gamma di strumenti consente di trasformare i dati testuali in un formato sistematico e fruibile per facilitare la classificazione dei documenti, individuare associazioni esplicite fra documenti e raggruppare documenti in categorie per nuove analisi.

Ricavare documenti

Vengono impiegate tecnologie linguistiche e di analisi avanzate per definire la competenza in materia disponibile nell'ambito dell'organizzazione. Una volta definiti a livello di sistema, i documenti sono elaborati automaticamente in base a tali regole di conoscenza, semantica e concetti. I documenti sono distribuiti in tempo reale e le informazioni vengono incanalate in flussi esistenti, consentendo l'assegnazione dei documenti alle persone di pertinenza.

Analizzare i social media e catturare informazioni pertinenti ai prodotti

Processi automatizzati individuano e analizzano il contenuto digitale in tempo reale per permettere di scoprire e interpretare le emozioni insite in informazioni qualitative presenti su Internet o in comunicazioni interne. Gli esperti di marketing, brand e logistica sono in grado di superare le limitazioni dettate dalla scoperta, interpretazione e annotazione manuale dei sentimenti mediante strumenti di ricerca o con lunghe revisioni manuali del testo. Non si può fare affidamento su un "engine" di ricerca per trovare il sentimento alla cieca se le domande da porre non sono note.

Collegare repository di documenti a relazioni definite in modo coerente

Integrare repository e risorse di testo esistenti associando testi disparati a definizioni e regole gerarchiche comuni. In tal modo, le organizzazioni sono in grado di massimizzare il valore dei testi a loro disposizione collegandoli con relazioni definite in modo coerente, riuscendo a rispondere rapidamente e in modo preciso alle esigenze di reperimento di informazioni.

Integrare le informazioni testuali con dati strutturati

Una soluzione di text mining completa consente di integrare dati strutturati con dati non strutturati (testo in formato libero) in modo da ottenere una visione completa e analisi significative all'interno di un ambiente integrato di modeling predittivo. L'analisi effettuata utilizzando entrambi i tipi di dati produrrà una visione più precisa e completa dell'organizzazione.

Componenti

SAS® Enterprise Content Categorization

SAS Enterprise Content Categorization è un'applicazione di facile utilizzo in ambiente Windows dotata di potenti tecniche di analisi del testo che permettono di controllare il flusso delle informazioni all'interno dell'organizzazione. Offre la possibilità di definire una tassonomia gerarchica che prevede il raggruppamento di argomenti correlati. Questo software organizza informazioni, documenti o siti web utilizzando una directory gerarchica semplice, intuitiva e navigabile. Classifica in modo automatico e preciso i documenti mediante regole personalizzabili e, nel contempo, associa, in modo coerente e in tempo reale, eventuale nuovo materiale a fonti testuali esistenti. Ciò conduce a una più rapida ed efficiente organizzazione, accesso, reperibilità e condivisione delle informazioni, riducendo l'overhead associato al processo di classificazione del contenuto, come il tagging manuale e l'indicizzazione retrospettiva.

Add-on for SAS® Enterprise Content Categorization

I moduli aggiuntivi di SAS Enterprise Content Categorization applicano tecniche avanzate di elaborazione del linguaggio naturale (NLP) e linguistiche per automatizzare le operazioni di elaborazione testi, consentendo quindi alle organizzazioni di realizzare rapidi e significativi miglioramenti in termini di efficienza, proprio dove servono. Queste funzionalità aggiuntive sono in grado di gestire enormi quantitativi di informazioni, impiegano la stessa velocità di elaborazione per qualsiasi

quantità di dati e codificano i dati comprimendoli. I moduli aggiuntivi, che includono ricerca e indicizzazione automatiche, segnalazioni in tempo reale relative al contenuto, individuazione di documenti duplicati, un workbench interattivo e un web crawler automatizzato, possono essere integrati nella soluzione SAS Enterprise Content Categorization quando risulta necessario estendere le funzionalità correnti. Grazie agli add-on, le organizzazioni sono in grado di personalizzare la propria soluzione di gestione del contenuto con la possibilità di migliorare i processi aziendali incentrati sui documenti e di aggiungere funzionalità all'evolversi delle esigenze.

SAS® Text Miner

SAS Text Miner offre una vasta gamma di strumenti di modeling linguistico e analitico per individuare ed estrarre conoscenza da molteplici documenti di testo. Dopo avere trasformato il testo in modo che possa essere alimentato in engine di data mining, gli argomenti e i temi sono identificati come associazioni esplicite, in modo da poter riunire i documenti in gruppi correlati per il modeling descrittivo o predittivo. L'esplorazione interattiva permette di individuare pattern, precedentemente sconosciuti, in insiemi di documenti e di applicare tali scoperte direttamente ai modelli predittivi, alle tassonomie o ai dashbord aziendali, attribuendo il massimo valore a tutte le fonti di informazioni. Questa soluzione permette di risparmiare denaro e risorse automatizzando le lunghe attività di lettura e comprensione del testo. Integrando origini dati strutturate con informazioni testuali, si ottiene una visione più precisa e completa della propria organizzazione.

SAS® Sentiment Analysis

Grazie a SAS Sentiment Analysis, le organizzazioni sono in grado di monitorare in modo accurato le opinioni dei consumatori (e di altri) con un'unica tecnologia che applica sia l'arte che la scienza alla valutazione di siti web, dati di call center interni, e-mail e altri dati testuali per estrarre i sentimenti espressi. Applicando modelli statistici e regole linguistiche che possono essere ottimizzate dagli esperti di business knowledge, è possibile definire i sentimenti positivi, negativi e neutri associati alle parole usate per descrivere prodotti, servizi, ecc. Di conseguenza, si ha la possibilità di utilizzare i feedback dei consumatori per sviluppare e gestire prodotti e strategie più efficaci.

SAS® Ontology Management

Questa soluzione all'avanguardia permette la gestione dei contenuti e la manutenzione delle ontologie nei repository e nei database a livello aziendale. Crea termini semantici che sono utilizzati per organizzare repository di testo in precedenza dissociati e isolati. Inoltre, crea e gestisce metadati omogenei e centralizzati presenti in molteplici insiemi di testo, in modo che gli engine di ricerca e reperimento delle informazioni identifichino in modo sistematico concetti comuni. In tal modo, le organizzazioni sono in grado di massimizzare il valore dei repository e degli insiemi di testi semantici collegandoli tra loro con relazioni definite in modo coerente, per rispondere in modo rapido e preciso alle esigenze di ricerca e reperimento delle informazioni.