

TECNICHE DI MARKETING:

Segmentazione dei clienti mediante l'Algoritmo K-Means



La segmentazione dei clienti è una pratica cruciale per le aziende che aspirano a capire meglio le necessità e i comportamenti dei loro clienti. Questo processo consiste nel dividere la base dei clienti in gruppi o segmenti che condividono caratteristiche simili. Attraverso la segmentazione, è possibile personalizzare le strategie di marketing, migliorare i servizi offerti e aumentare l'efficienza operativa.

Perché la Segmentazione è Importante?

La segmentazione permette alle aziende di:

- Identificare nicchie di mercato** specifiche all'interno di un panorama più ampio, focalizzandosi su gruppi di clienti con esigenze particolari.
- Personalizzare prodotti e servizi** per rispondere in modo più mirato alle esigenze di ogni segmento, aumentando così la soddisfazione del cliente.
- Distribuire le risorse di marketing** in maniera più efficiente, concentrando gli sforzi e i budget sui segmenti che promettono il maggiore ritorno sull'investimento.
- Incrementare la lealtà dei clienti** attraverso offerte e comunicazioni che risuonino più

direttamente con le loro aspettative e bisogni.

- Adeguare i prezzi** per diversi gruppi di mercato, massimizzando la redditività attraverso strategie di pricing differenziate.

Clustering con K-Means

L'algoritmo k-means è un metodo molto usato per raggruppare dati simili insieme, come ad esempio organizzare i clienti in gruppi basati su comportamenti o preferenze simili. K rappresenta il numero di cluster che otterremo e means indica che useremo le distanze dai centri dei cluster stessi.

KNIME offre un metodo efficace e accessibile per realizzare queste divisioni attraverso l'utilizzo di questo algoritmo.

- Inizio con i Centroidi:** Immaginiamo di avere un grande gruppo di punti sparsi su un foglio. L'algoritmo inizia sce-

gliendo alcuni punti casuali come "centri" iniziali, chiamati centroidi. Questi punti servono come punti di riferimento iniziali per formare i gruppi.

- Assegnazione dei Punti:** Ogni punto nel foglio (che rappresenta un cliente o un dato) viene poi assegnato al centro più vicino a lui. Quindi, se un punto è più vicino al centro A che non al centro B, verrà messo nel gruppo di A.
- Aggiornamento dei Centri:** Una volta che tutti i punti sono stati assegnati a un centro, l'algoritmo guarda dentro ogni gruppo e trova un nuovo centro, che è semplicemente il "punto medio" di tutti i punti in quel gruppo. Questo nuovo punto medio diventa il nuovo centro per il prossimo giro di assegnazioni (si veda Fig.1 sotto).

- Ripetizione:** Questi passi —

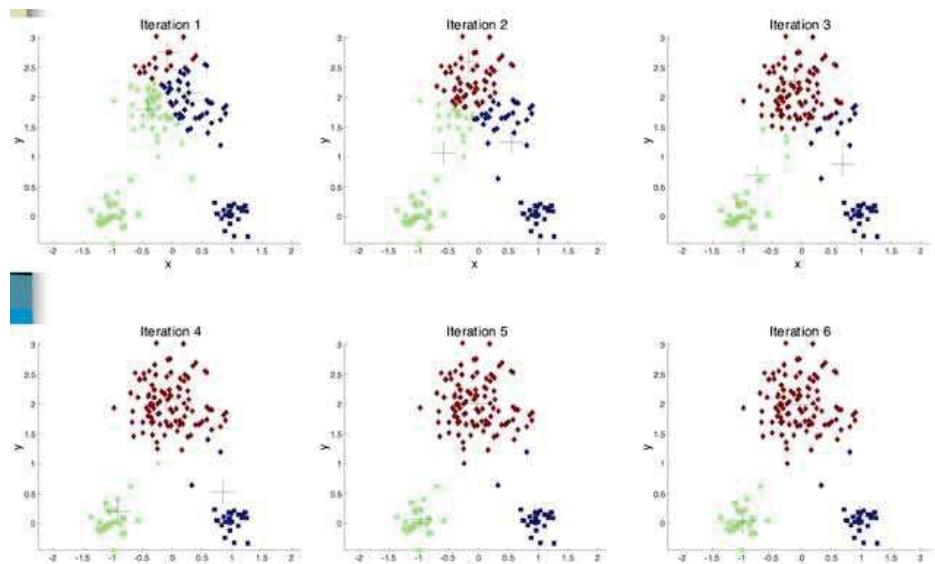


Figura 1 Tratta da: <http://bias.csr.unibo.it/golfarelli/DataMining/MaterialeDidattico/DMISIClustering.pdf>

assegnare i punti ai centri e poi aggiornare i centri — si ripetono più volte. Con ogni ripetizione, i centri si spostano leggermente, basandosi sempre più strettamente su come i dati sono distribuiti realmente.

5. **Finitura:** Il processo continua finché i centri non cambiano più di tanto da un giro all'altro, il che significa che i gruppi sono stati definiti chiaramente e l'algoritmo ha finito il suo lavoro.

k-means è la scelta del numero k di cluster, che non è determinata dall'algoritmo stesso e deve essere definita a priori. La selezione di k può essere guidata da considerazioni pratiche sul contesto di applicazione o può essere determinata attraverso metodi numerici, ad esempio con quello "Elbow", il gomito, come nell'esercizio (il cui link è riportato alla fine di questo articolo).

Il gomito della curva (si veda Fig.2) suggerisce il miglior numero

di cluster, si sceglie in base a dove cambia maggiormente la pendenza.

Analisi e Monitoraggio

Una volta definiti, i cluster si analizzano per comprendere le caratteristiche dominanti di ciascun gruppo. Questo passaggio è fondamentale per sviluppare messaggi e campagne di marketing mirate. Inoltre, è essenziale un monitoraggio continuo per valutare l'efficacia delle strategie di segmentazione e per fare aggiustamenti basati su cambiamenti nel comportamento dei clienti o nelle condizioni di mercato.

Come Implementare la Segmentazione con KNIME

Per lo svolgimento dell'esercizio (si veda il link per scaricarlo riportato alla fine di questo articolo) mi sono liberamente ispirato ai libri dell'amico Andrea De Mauro: "Big Data Analytics" e "Data Analytics per tutti", che vi consiglio vivamente.

Nell'esercizio, oltre al calcolo dei cluster, a titolo di esempio, sono stati calcolati anche il fatturato dei primi cinque prodotti e i prodotti

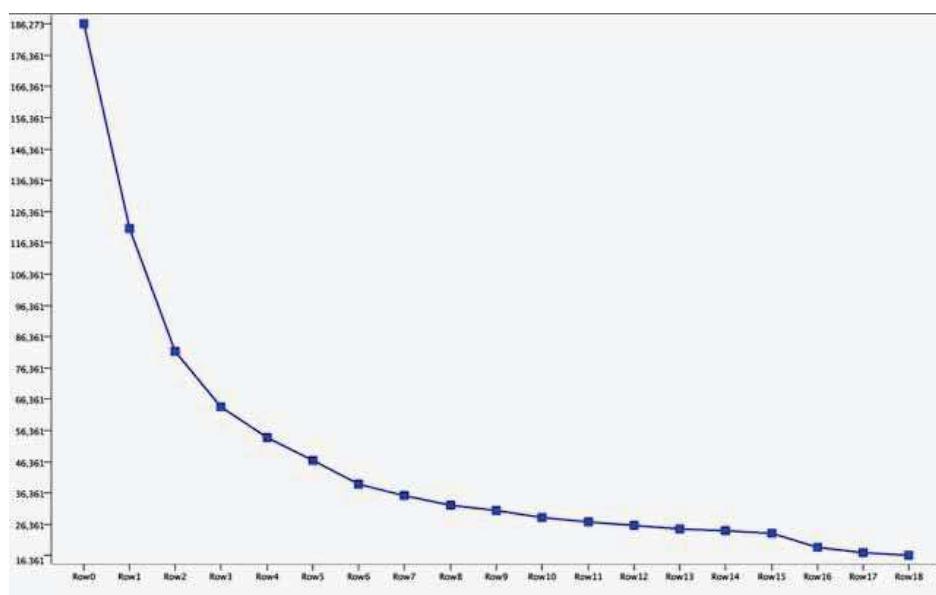


Fig.2 Metodo di Elbow: il gomito della curva suggerisce il miglior numero di cluster

Scegliere il Numero di Gruppi: k

Un punto importante da considerare è quanti gruppi (o cluster) si vogliono formare. Si può decidere basandosi su ciò che si conosce già dai dati, o sperimentando con diversi numeri, per vedere quali forniscono i risultati più utili.

Un aspetto critico dell'algoritmo

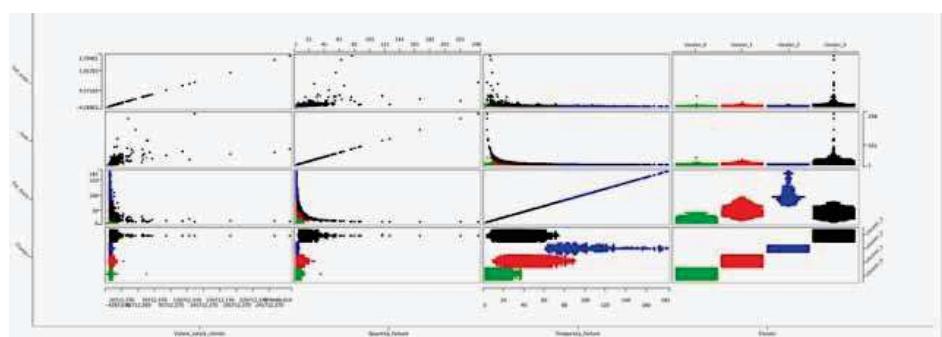


Figura 3 Confronto nei cluster di Valore Totale Cliente, Quantità Fatture, Frequenza Fatture

più venduti in volume.

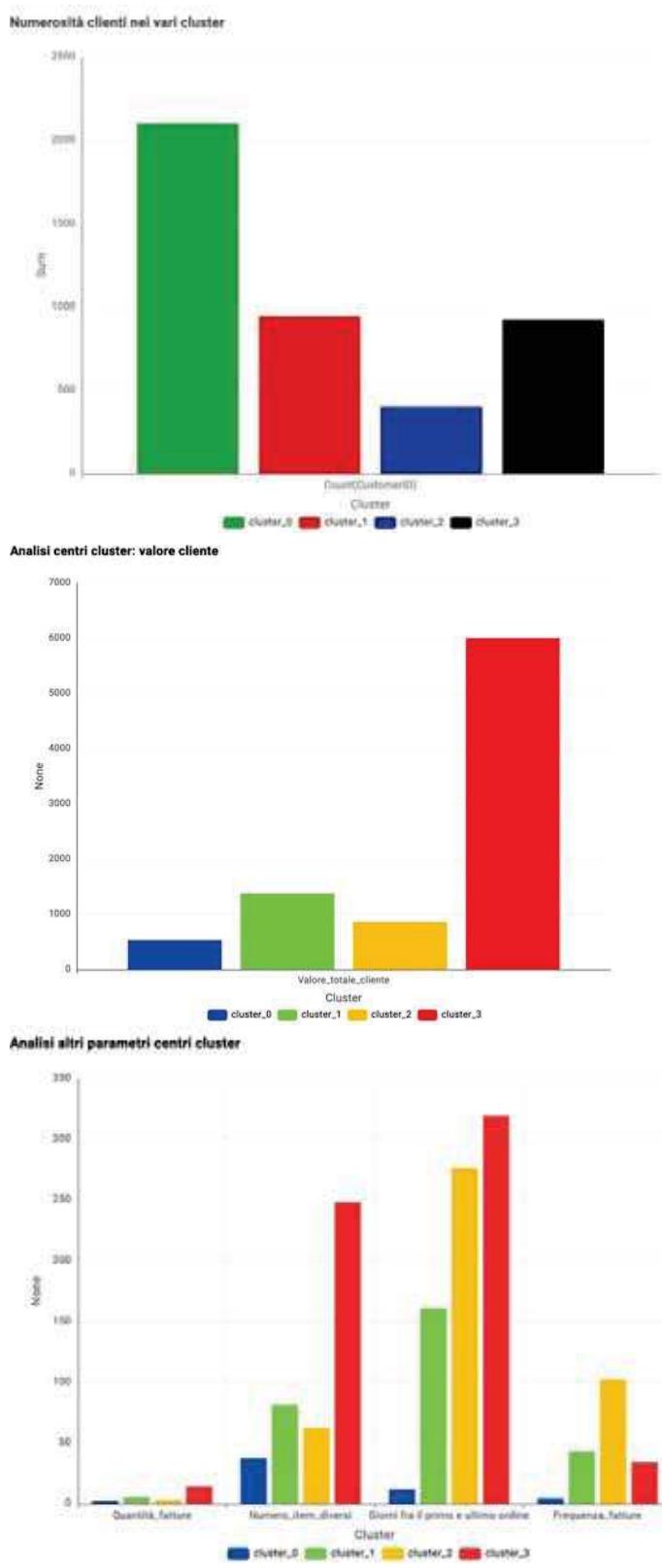
Il processo di segmentazione in KNIME inizia con la raccolta e l'organizzazione dei dati dei clienti, che possono provenire da fonti diverse come transazioni, interazioni sui social media, sondaggi, e dati raccolti da terze parti. Questi dati vengono poi caricati in KNIME, dove attraverso l'uso di nodi specifici per la lettura dei dati (come CSV Reader o Excel Reader), si inizia la preparazione del dataset per l'analisi.

Preparazione dei Dati

Dopo il caricamento, i dati richiedono una serie di trattamenti per garantire che l'analisi sia accurata e significativa:

- Rimozione delle colonne non necessarie:** Utilizzando il nodo “Column Filter”, si eliminano le informazioni irrilevanti per l'analisi.
- Gestione dei valori mancanti:** I dati incompleti vengono trattati con il nodo “Missing Value” per decidere se sostituirli o rimuoverli.
- Normalizzazione delle variabili numeriche:** Per evitare distorsioni dovute a scale diverse, si utilizza il nodo “Normalizer”.

Dopo aver preparato i dati, il passo successivo è l'applicazione dell'algoritmo k-means: configurare il numero di cluster desiderato e specificare alcuni parametri, come il numero di iterazioni e il metodo di inizializzazione dei centroidi.



È importante sperimentare con un diverso numero di cluster per identificare la soluzione ottimale per i nostri scopi.

Una volta eseguito il clustering, è essenziale analizzare i gruppi formati, sia con le tabelle, sia con metodi grafici. Ad esempio, vediamo nei grafici a pag. 26 la numerosità dei clienti nei vari cluster o le varie caratteristiche di interesse.

Dobbiamo però analizzare anche i valori dei centroidi, che ci rappresentano i criteri con cui il sistema ha creato la segmentazione, ad esempio il cluster con il maggiore valore per cliente, oppure altri parametri, come il numero o la frequenza delle fatture.

Conclusione

Con i cluster chiaramente definiti e analizzati, le aziende possono adot-

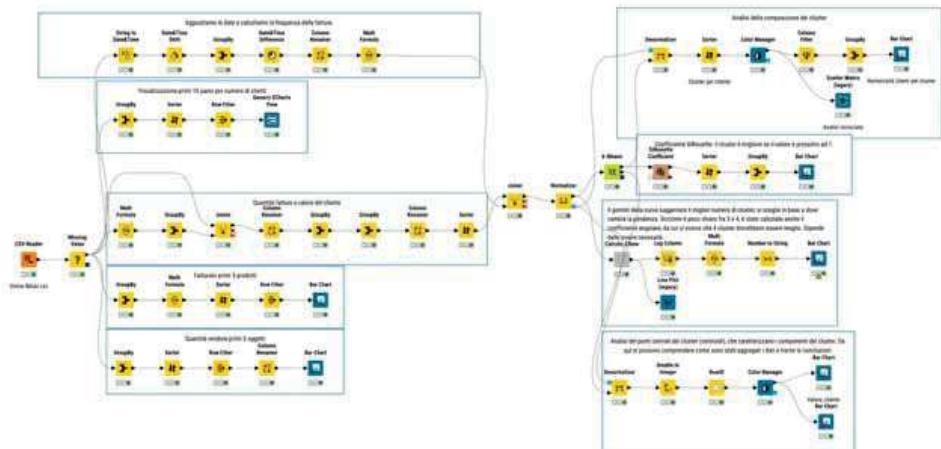


Fig. 7 L'esercizio

tare strategie di marketing mirate per personalizzare le offerte, creando promozioni basate sulle caratteristiche distinctive di ciascun cluster. Questo non solo ottimizza le risorse, ma migliora anche l'engagement dei clienti. KNIME, grazie alla sua interfaccia intuitiva e alle robuste funzionalità di analisi dei dati, rende questo pro-

cesso accessibile anche a chi non possiede una profonda esperienza nel campo dell'analisi dei dati.

L'esercizio (si veda fig.7) è scaricabile al seguente indirizzo:
knime://My-KNIME-Hub/Users/zompazompa/Public/Esercizio_settembre_24_finale

**CAFINDUSTRIA
EMILIA-ROMAGNA S.P.A.**

Sede operativa:
Via Castiglione 124
40136 Bologna

Tel **051.33.09.09**
Fax 051.26.56.90

info@cafindustria.onmicrosoft.com
www.cafindustria.com

**CONTATTACI
PER MAGGIORI
INFORMAZIONI**



TRA I NOSTRI SERVIZI:

- **Modello REDDITI PF, anche con quadri RM - RT - RW**
- **Modello F24 per versamento IMU, IVIE, IVAFE e imposta sostitutiva**
- **Dichiarazioni di successione**
- **Applicazione "cedolare secca"**
- **Invii telematici**
- **Consulenze specialistiche per espatriati ed impatriati**
- **Servizio colf & badanti**
- **ISEE**
- **Servizio di assistenza personalizzato per tutto l'anno**

IL CONTRIBUENTE È GARANTITO NEI SUOI DIRITTI DI RISERVATEZZA