

Dipartimento di Impresa e Management
Corso di Laurea Magistrale in Marketing

Cattedra Business and Marketing Analytics

**Digital Fashion: machine learning a supporto del
marketing nel settore della moda partendo dalla
classificazione di borse nei siti e-commerce secondo le
loro caratteristiche**

Prof. Frontoni Emanuele

RELATORE

Prof. Laura Luigi

CORRELATORE

Gabriele Donà 765601

CANDIDATO

INDICE

INTRODUZIONE: IL CAMBIAMENTO RADICALE NEL SETTORE DELLA MODA	1
1.1 L'IMPORTANZA DELLA MODA AL GIORNI D'OGGI	1
1.2 CRESCITA ECONOMICA DEL SETTORE	5
1.3 E-COMMERCE	12
1.4 AI NEL SETTORE DELLA MODA	17
CAPITOLO 2: LITERATURE REVIEW	20
2.1 P'AI NELL'E-COMMERCE	20
2.2 CLASSIFICAZIONE DEGLI ITEMS DI MODA	27
2.3 GAP DI RICERCA	29
CAPITOLO 3: MATERIALI E METODI	31
3.1 DATASET	31
3.2 TECNOLOGIE UTILIZZATE	41
3.2.1 LIBRERIE DI PYTHON UTILIZZATE	43
3.2.2 MODELLI UTILIZZATI	46
CAPITOLO 4: RISULTATI E DISCUSSIONI	49
4.1 RISULTATI	49
4.2 SUDDIVISIONE PER OCCASIONE D'USO	60
4.2.1 LAVORO E UFFICIO	62
4.2.2 VIAGGI E AVVENTURE	63
4.2.3 TEMPO LIBERO E ATTIVITA' QUOTIDIANE	64
4.2.4 EVENTI E OCCASIONI SPECIALI	65
4.2.5 USCITE FUORI PORTA	66
4.3 L'UTILIZZO PRATICO DELLO STUDIO	68
4.4 INIZIATIVE DI MARKETING	69
4.4.1 L'IMPORTANZA DEGLI OBIETTIVI "SMART"	69
4.4.2 CAMPAGNE TEMATICHE STAGIONALI	71
4.4.3 BUNDLE CON PRODOTTI COLLEGATI	78
4.5 CONTROLLO SULL'EFFICACIA DELLE INIZIATIVE MARKETING	81
CAPITOLO 5: CONCLUSIONI E SVILUPPI FUTURI	84
5.1 CONCLUSIONI	84
5.2 LIMITI E RICERCHE FUTURE	86
BIBLIOGRAFIA	88

INTRODUZIONE: IL CAMBIAMENTO RADICALE NEL SETTORE DELLA MODA

1.1 L'IMPORTANZA DELLA MODA AL GIORNI D'OGGI

Al giorno d'oggi la moda, gli accessori e i vestiti che tutti i giorni portiamo, oramai non sono più un mero ornamento ma sono l'espressione del nostro corpo e del nostro io. Ciò non si limita neanche più alla semplice dimensione della bellezza, ma trova la sua vera espressione nelle “pratiche sociali quotidiane che hanno al loro centro il corpo e i suoi modi di apparire e di comunicare”¹. “Il lato estetico della moda, infatti, non può di fatto prescindere da quello valoriale del contatto con i significati etici. Se un vestito echeggia attributi di tipo militare, esso veicolerà inevitabilmente qualcosa di militaresco. Se ad esempio tende troppo al nero, esso potrà rinforzare il suo legame con il lutto, la religione, oppure con la guerra [...]. Un abito color caki a sua volta ci riporterà implicitamente con la memoria allo stile dell'epoca coloniale. Questo aspetto fondamentale della moda, nella sua capacità implicita e inconscia di evocare attributi di tipo storico e dunque etico, è una sua componente centrale che ne rende conto non solo del gusto e della misura, ma anche di suoi significati, caratteri e attributi assai rilevanti”².

Non solo quindi attraverso di essa si riesce ad esprimere il proprio io ma addirittura la propria cultura. Esiste “un collegamento ideale tra mondi e generazioni che i segni del corpo rendono manifesto. Da Piazza Tahrir³ nella “primavera” nordafricana, alle strade della Stoccolma di Lisbeth Salander⁴, c'è una stretta relazione tra lo stile espresso nell'abbigliamento e le forme di ribellione e di anticonformismo”⁵. Questo collegamento intrinseco tra moda e identità culturale emerge ancor di più quando si osservano degli stili che incarnano appieno movimenti sociali. Per esempio, durante le rivoluzioni sociali, come quella della Primavera Araba⁶ il vestiario non era solamente una scelta estetica, ma era simbolo di appartenenza. Allo stesso modo, durante la Rivoluzione francese, gli

¹ Cefalato, P. (2020). La moda oltre la moda. Università degli Studi di Bari Aldo Moro

² Fratini, T. (2015). Sul dominio della moda e i suoi effetti sulla persona nell'epoca odierna: considerazioni pedagogiche. Studi Sulla Formazione/Open Journal of Education, 18(1), 147–159.

³ Piazza situata nel cuore del Cairo, è divenuta un simbolo della Primavera Araba, il movimento di protesta che ha attraversato il Nordafrica e il Medio Oriente all'inizio del 2011. In questa piazza si sono radunati migliaia di cittadini per chiedere la fine del regime. Perciò essa è simbolo di speranza e resistenza per la libertà. Durante la resistenza sono stati iconici i vestiti indossati dai cittadini, i quali molti portavano keffiyeh come simbolo di resistenza e le donne hijab colorati, sottolineando la propria identità culturale e partecipazione alla rivolta.

⁴ Le strade di Stoccolma sono fondamentali nell'universo di Lisbeth Salander, protagonista della serie "Millennium" di Stieg Larsson. Questo luogo riflette il contrasto tra criminalità e vita quotidiana. È iconico l'abbigliamento della protagonista che è punk e gotico, sottolineando lo spirito anticonformista di Lisbeth.

⁵ Cefalato, P. (2020). La moda oltre la moda. Università degli Studi di Bari Aldo Moro

⁶ Serie di rivolte popolari iniziate dal 2010 in diversi paesi arabi come la Tunisia, l'Egitto, la Libia e la Siria. I protestanti chiedevano migliori condizioni socioeconomiche e libertà. Questo movimento ha comportato la caduta di diversi regimi autoritari ed ha avuto un forte impatto sulla geopolitica locale.

indumenti assunsero valore politico. Tra questi, i cosiddetti sans-culottes⁷ che iniziarono a vestire pantaloni lunghi come segno di opposizione alla nobiltà parigina che era solita indossare le cosiddette culotte.

Questi esempi dimostrano che, durante tutta la storia dell'umanità, la moda o più in generale ciò che indossiamo non è solo un modo per coprirsi o adornare il proprio corpo. Essa è una forma di appartenenza, arte, comunicazione che permette agli individui di esprimere la propria identità e i propri valori.

Come nel passato anche oggi scegliere cosa indossare non è questione di tendenza ma rappresenta molto di più, basti pensare come, soprattutto negli ultimi tempi, anche quei "dogmi" e stereotipi da sempre presenti nel mondo della moda stiano cadendo per far spazio alla libera espressione di sé stessi.

Per quanto riguarda la presenza di stereotipi legati agli indumenti, qualche tentativo di abatterli era stato fatto anche in passato, ad esempio nel 1984 Jean Paul Gaultier, celebre stilista francese famoso per il suo stile irriverente e provocatorio, durante una sfilata di Parigi intitolata "Et Dieu Créa l'Homme" fece molto clamore e scalpore dato che sfilarono molti uomini con varie tipologie di gonne. Questo gesto ha sfidato le convenzioni di genere legate all'indumento della gonna, e facendo ciò, ha voluto dimostrare come l'abbigliamento può essere un simbolo di rottura alle norme sociali.

Ancor prima, con la rivoluzione culturale degli anni '60 e '70, l'espressione di sé stessi e soprattutto la libertà, emerse e fu uno degli argomenti centrali portati avanti nella rivoluzione. Questi ideali erano legati a tutta la cultura che si è riflessa inevitabilmente anche sulla moda e sui trend, basti pensare alle iconiche camicie sgargianti e pantaloni a zampa di elefante che furono indossati non solo da donne ma da moltissimi uomini. Questa diffusissima moda in quegli anni simboleggiava un rifiuto alle rigide divisioni di genere e un avvicinamento alla libertà di espressione.

Anche precedentemente c'era stato il tentativo di abbattere gli stereotipi riguardo l'abbigliamento. Durante gli anni '40 divenne celebre la Zoot Suit⁸, con i suoi completi con pantaloni larghi e giacche lunghe, andava in controcorrente rispetto agli abiti tradizionali dell'epoca. Questa tendenza è un

⁷ Erano i radicali militanti della classe operaia durante la Rivoluzione francese. Prendono il nome dalla tipologia di pantaloni lunghi che indossavano, in contrapposizione ai pantaloni corti vestiti dall'aristocrazia francese.

⁸ È uno stile che nacque durante gli anni '30 nei quartieri afroamericani di Harlem a New York. Esso divenne rapidamente molto popolare tra i giovani latini e afroamericani in America. La sua caratteristica principale è che è uno stile particolarmente eccentrico e vistoso.

esempio di come un capo di abbigliamento può sfidare le norme sartoriali tradizionali ed essere simbolo di rifiuto alle norme sociali e razziali.

Passando ai giorni nostri, sono molti e celebri le iniziative per cercare di abbattere i dogmi. Tra queste bisogna citare le sfilate di Alessandro Michele, direttore creativo per la famosa firma Gucci dal 2015 al 2022 e nel 2024 per Valentino, che sono solo alcuni degli esempi tra i più conosciuti. Egli ha creato moltissime collezioni senza un'identificazione di genere ma che permettono all'uomo e alla donna di sentirsi liberi di esprimersi e di far emergere il proprio io interiore anche utilizzando vestiri o accessori comunemente associati ad un genere. Più in particolare la sua collezione più iconica che ha portato al culmine il suo pensiero, è stata la collezione autunno/inverno 2015-2016, nella quale ha presentato un'estetica androgina, con uomini che indossavano dettagli floreali e bluse⁹ trasparenti mentre le donne indossavano capi dal taglio maschile.

Questa abolizione dei dogmi era legata perlopiù al mondo dell'alta moda in delle collezioni che all'occhio della gente potevano e possono essere considerate "provocanti" invece di trasmettere libertà. Questa distanza tra il mondo di tutti i giorni e quello dell'alta moda si è sempre di più assottigliata grazie alla sempre più fitta presenza di influencer, che nei vari social, promuovono l'accettazione di sé stessi e della diversità. Queste figure, grazie al vasto bacino d'utenza presente nelle piattaforme social, sono riuscite concretamente ad abbattere gli stereotipi di genere e a far emergere come la moda, i vestiti e gli accessori siano una reale e concreta espressione di sé stessi. Essi sono riusciti a sottolineare sempre di più il concetto che la moda è un potente mezzo di espressione personale, e non una mera questione di abbigliamento e di trend. Condividendo, sui profili personali, contenuti che celebrano l'autenticità, hanno dimostrato che la moda è accessibile a tutti, senza nessun impedimento di genere, etnia, sesso ed età. Riuscendo così a mettere a proprio agio le persone e facendole sentire libere di poter utilizzare i propri indumenti come strumento di comunicazione.

Attraverso creatività, autenticità, dialogo e narrativa visiva mostrano nelle varie piattaforme social, Instagram, Tik Tok e Facebook¹⁰, che la moda è un importante mezzo espressivo per far emergere la propria identità, i propri valori e le emozioni.

Uno degli esempi più importanti dell'ormai sempre più centrale ruolo degli influencer nel mondo della moda intesa come espressione di sé stessi, è Paloma Elsesser, modella che "rappresenta uno dei nuovi simboli di autenticità, diversità e body positivity, contribuendo a ridefinire i canoni estetici in

⁹ Capi d'abbigliamento femminili simili a camicie, caratterizzate da un taglio morbido e spesso ampio.

¹⁰ In modo da coinvolgere e raggiungere un pubblico giovane ma anche di età adulta

un contesto in continua evoluzione come quello fashion”¹¹. La modella non solo sta portando avanti il suo credo abbattendo i vari dogmi nel mondo della moda, ma soprattutto, sta ispirando centinaia di migliaia di fan a non rinunciare ad esprimere il proprio essere ma a valorizzarlo attraverso i propri indumenti e accessori. Paloma, ha reso concreta la sua azione attraverso la partecipazione alla campagna globale "Dove Self-Esteem Project"¹², che è un progetto avviato con il fine di migliorare l'autostima delle giovani donne.

Perciò negli ultimi decenni l'alta moda, attraverso sfilate e collezioni ha sfidato gli stereotipi legati all'abbigliamento. Tuttavia, questa azione di abbattimento non è solo portata avanti nelle grandi passerelle, ma ognuno di noi sta contribuendo significativamente al processo. Grazie a una maggior accessibilità alla moda e alle figure degli influencer, anche persone comuni stanno giocando un ruolo attivo in questa battaglia per rendere gli indumenti e gli accessori che indossiamo uno strumento di espressione di sé stessi. Permettendo così a tutti di potersi esprimere nel modo più libero, non attraverso le parole in questo caso, ma con ciò che indossano.

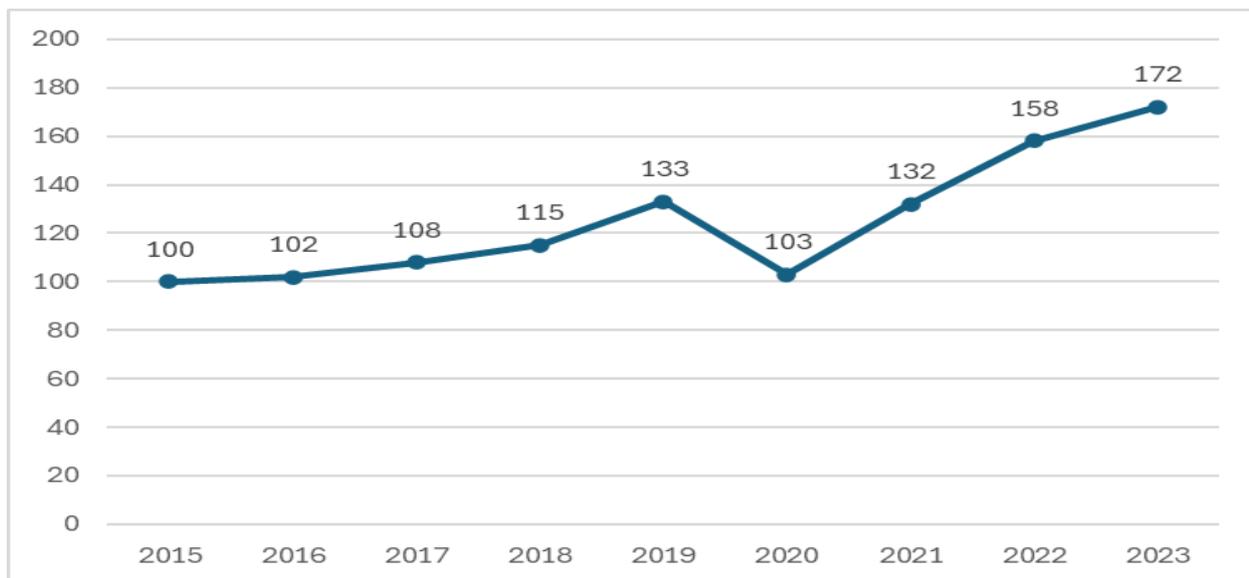
¹¹ Loretelli, V. (2023, December 23). Paloma Elsesser: la bio e la carriera. Retrieved from Harper's BAZAAR website: <https://www.harpersbazaar.com/it/moda/a46187930/paloma-elsesser/>

¹² Progetto lanciato nel 2004 da Dove che ha come obiettivo quello di migliorare l'autostima e la percezione del proprio corpo. Attua questa iniziativa offrendo risorse educative e workshop per affrontare temi come l'immagine di sé.

1.2 CRESCITA ECONOMICA DEL SETTORE

È per questa serie di motivi che la moda sta acquistando sempre più popolarità. Oltretutto quindi, “il fashion sta occupando il centro nella comprensione popolare della cultura moderna. Gode di una copertura senza precedenti nei media occidentali e definisce il tenore della vita urbana come nessun altro mezzo di comunicazione visiva”¹³. Questa sempre più importanza si è inevitabilmente riflessa sull’economia e sul fatturato delle industrie di moda. Ciò è tangibile considerando che nel 2023 il fatturato nel mondo della moda italiana, patria delle più importanti firme, è cresciuto del “4%, per un totale di circa 103 miliardi di euro”¹⁴. Questa tendenza di crescita è stata una costante degli ultimi anni, dal 2015 al 2023 i ricavi sono cresciuti di 7% medio annuo e complessivamente del 72% nell’arco di 8 anni.

Grafico 1: Crescita dei ricavi del settore della moda in Italia (valori annuali indicizzati 2015=100)



Fonte: Banca Ifis¹⁵

Come si può evincere dal grafico sopra illustrato, il settore ha avuto un’impennata nei ricavi a partire dal 2021, dopo i risultati negativi del 2020 che sono il risultato della crisi causata dal Covid.

¹³ Breward, C. (2003). Fashion. In Google Books. OUP Oxford. Retrieved from https://books.google.it/books?hl=it&lr=&id=rzj_AwAAQBAJ&oi=fnd&pg=PP1&dq=importance+of+fashion&ots=RI-lsHpxgf&sig=tQ2aKkqmqmYiR4g_G00H6ZCp30yM&redir_esc=y#v=onepage&q&f=false

¹⁴ Redazione AGI. (2023). Il settore moda chiuderà il 2023 al +4%, a settembre contrazione. AGI. https://www.ansa.it/sito/notizie/economia/2023/12/13/il-settore-moda-chiuderà-il-2023-al-4-a-settembre-contrazione_51c143d3-c9a5-4a62-809c-37a86c8e7fe5.html.

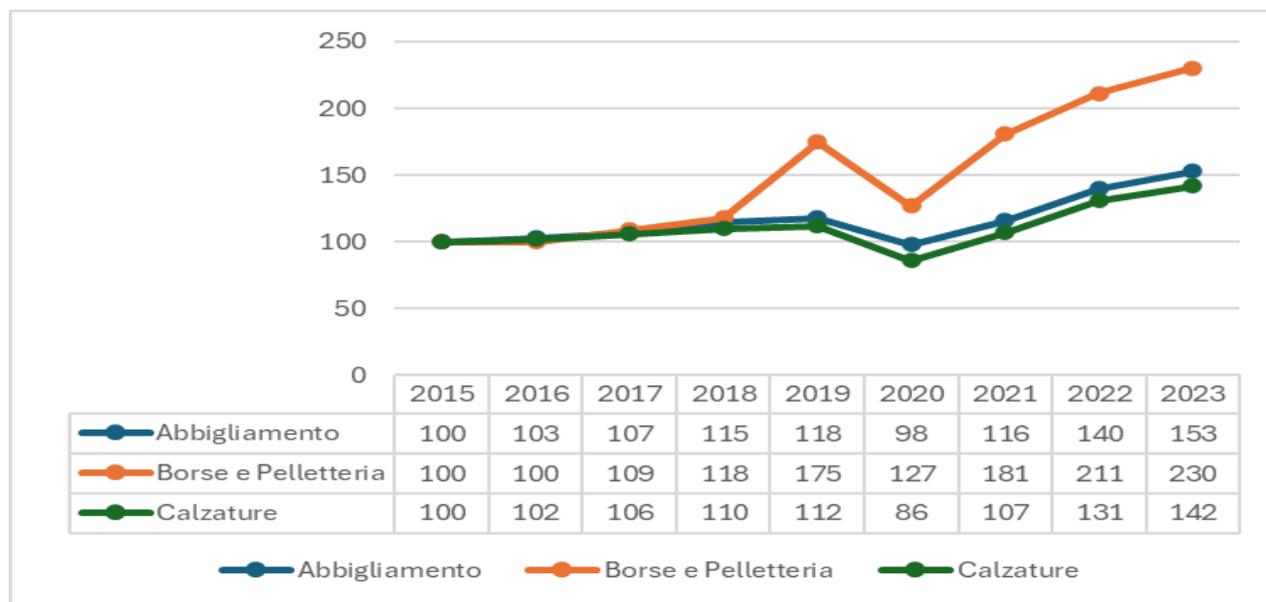
¹⁵ https://www.bancaifis.it/app/uploads/2024/05/Moda-trend-e-consumer-insight_20240229.pdf

Tutti i settori, dall'abbigliamento alla pelletteria, hanno contribuito alla crescita molto ampia dei ricavi delle industrie di questo settore. In generale, secondo un'analisi condotta dall'ufficio studi di Banca Ifis, questa crescita in Italia dei ricavi è dovuta principalmente alle esportazioni che sono cresciute del 17% dal 2019 al 2023, passando da 47 a 53 miliardi di euro.

Inoltre, ciò che ha contribuito molto all'impennata nei ricavi, è la concezione della moda che è leggermente cambiata ed ha acquisito sempre più caratteristiche funzionali a rispondere al desiderio di esprimere sé stessi e alla crescente domanda di benessere.

Andando più nel dettaglio la sezione che ha più contribuito nell'aumentare i ricavi è stato il compartimento delle borse e pelletteria¹⁶.

Grafico 2: Andamento dei ricavi del settore della moda in Italia (valori annuali indicizzati 2015=100) per i comparti abbigliamento, borse, pelletteria e calzature



Fonte: Banca Ifis

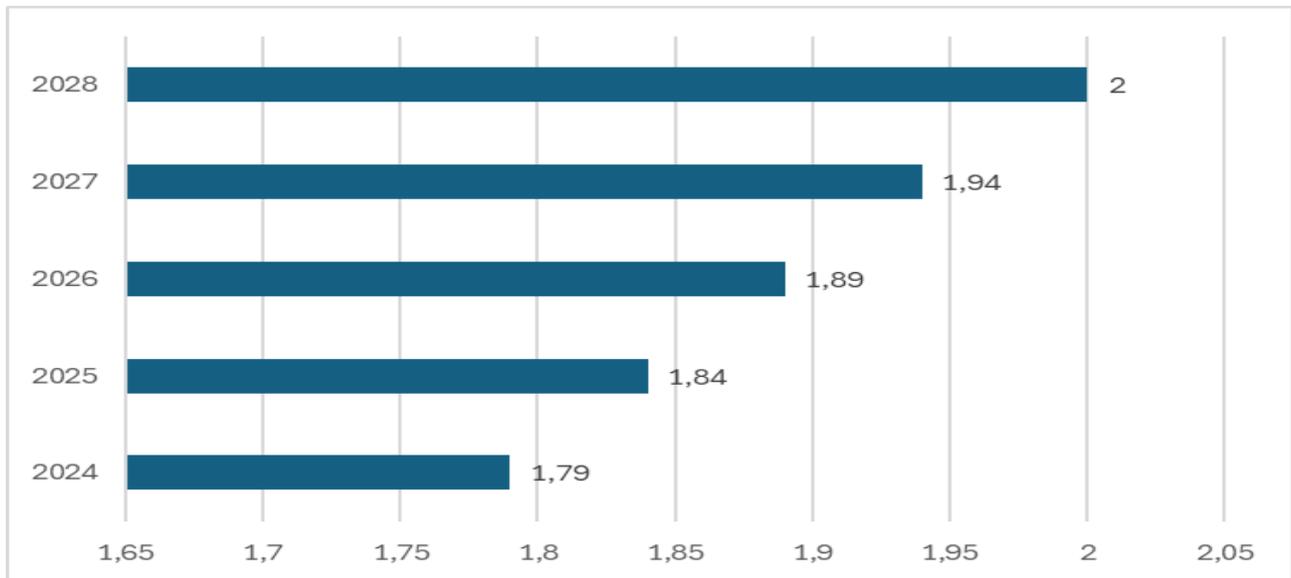
Guardando invece i dati che riguardano tutti i mercati nel 2020 McKinsey afferma che il settore “ha registrato un aumento del 21% dei ricavi nel 2020-21 rispetto all’anno precedente [...] e dovrebbe crescere tra il 5% e il 10% nel 2023, trainato dal forte slancio in Cina (previsto in crescita tra il 9 e il 14%) e negli Stati Uniti (previsto in crescita tra il 5 e il 14%)”¹⁷. Le stime per quanto riguarda i prossimi anni confermano che il settore sta godendo di un momento molto fruttuoso e che la sua ascesa non diminuirà nei prossimi anni avvenire. Queste previsioni sono state confermate dai dati

¹⁶ La quale conta il 37% della Moda italiana

¹⁷ Rapporto McKinsey sulla moda: previsioni sul 2023. Assomac. <https://assomac.it/it/news/news-dal-mondo/rapporto-mckinsey-sulla-moda-previsioni-sul-2023/>

forniti da Statista Consumer Market Insights che conferma che nel 2028 i ricavi arriveranno a toccare i due trilioni di dollari.

Grafico 3: Ricavi del mercato dell'abbigliamento nel mondo da oggi al 2018 (in trilioni di dollari)



Fonte: Statista Consumer Market Insights; Statista¹⁸

A sottolineare l'importanza del settore, addirittura, dati attuali confermano che in totale l'industria è valutata tre bilioni di dollari. La moda britannica vale da sola 26 miliardi di sterline. Si tratta della seconda attività economica più grande al mondo per intensità commerciale, impiega oltre 57 milioni di lavoratori nei paesi sviluppati, di cui l'80% sono donne¹⁹.

Tutti i ricavi o comunque il valore delle aziende del settore non sono imputabili solo agli indumenti di abbigliamento ma giocano un ruolo fondamentale anche gli altri oggetti venduti dalle firme, come mostrato nel Grafico 2.

Infatti, accessori, come borse e scarpe, sono un elemento fondamentale per i guadagni delle industrie di moda. Molte firme basano principalmente i loro guadagni sulla pelletteria, come il noto brand Hermès che deve la maggior parte del suo successo alle sue iconiche borse, cinture e scarpe. Tra queste spiccano i modelli Birkin e Kelly che hanno fatto fare la fortuna all'azienda parigina portandola ai vertici delle aziende del settore.

¹⁸ Smith, P. (2023, September 15). Revenue of the Global Apparel Market 2014-2027. Retrieved from Statista website: <https://www.statista.com/forecasts/821415/value-of-the-global-apparel-market>

¹⁹ Corner, Frances. "Perché La Moda è Importante." Google Arts & Culture, artsandculture.google.com/story/_QKS0J-OeT7HIA?hl=it.

Un'altra grande firma che ha giovato molto delle sue collezioni di borse e di pelletteria è il noto brand parigino Chanel. Il CFO, Philippe Blondiaux, ha spiegato infatti che grazie al continuo aumento dei prezzi di borse ha comportato un grosso aumento dei ricavi per l'azienda nel 2022 e che la metà di essi sono ascrivibili soprattutto alle borse.

Grafico 4: Personal luxury goods market value worldwide from 2009 to 2020, by market segment (in million euros)



Fonte: Bain & Company²⁰

Aziende del settore che non sono nel segmento del luxury hanno sfruttato questo momento di forte crescita del mercato per emergere ed essere considerate come dei veri e propri colossi dell'industria, tra questi bisogna sottolineare l'ascesa del gruppo Inditex, nella quale al suo interno fanno parte Zara, Pull&Bear, Massimo Dutti, Bershka, Stradivarius e Oysho.

Il gruppo ha chiuso il 2023 con un fatturato pari a 35,9 miliardi di euro, registrando un aumento del 10,4% rispetto all'anno prima²¹. A guida di questa impennata del fatturato c'è Zara, noto brand fondato nel 1974 da Amancio Ortega e Rosalía Mera che hanno creato questa azienda basandola sul modello del "fast fashion", che prevede la capacità di reagire rapidamente alle nuove tendenze di moda e di renderle subito disponibili al pubblico attraverso tutti i punti vendita, che ad oggi sono in

²⁰ Sabanoglu, T. (2023). Personal luxury goods market value by market segment worldwide 2020. Retrieved from Statista website: <https://www.statista.com/statistics/883902/personal-luxury-goods-market-value-by-product-category-worldwide/>

²¹ "Inditex: 36 Miliardi Di Ricavi Nel 2023, Dividendo +28%." Wwww.teleborsa.it, 13 Mar. 2024, [finanza.lastampa.it/News/2024/03/13/inditex-36-miliardi-di-ricavi-nel-2023-dividendo-+28percento/MjdfMjAyNC0wMy0xM19UTEI#:~:text=Inditex%2C%20la%20societ%C3%A0%20spagnola%20dell.](https://www.finanza.lastampa.it/News/2024/03/13/inditex-36-miliardi-di-ricavi-nel-2023-dividendo-+28percento/MjdfMjAyNC0wMy0xM19UTEI#:~:text=Inditex%2C%20la%20societ%C3%A0%20spagnola%20dell.) Accessed 21 July 2024.

ogni parte del mondo. Zara, nel gruppo, è quella che ha portato i migliori risultati con un aumento del 21%, seguita da Pull&Bear che ha registrato un +15%²². È importante sottolineare che, nonostante ci sia stato un aumento dei prezzi medi degli items venduti da queste catene del 5% la domanda di abbigliamento è continuata e non ha subito dei cali²³.

Oramai le 10 aziende di vestiti e di moda valutate maggiormente nel 2023 sono diventati questi colossi del fast fashion che sono riuscite ad emergere in un mercato molto competitivo e a distinguersi dai competitors rendendosi riconoscibili pur non essendo dei brand di lusso. Terzo in questa classifica stilata da Kantar²⁴ c'è la già citata Zara che è preceduta solo da Nike e da Shein.

Il primo marchio, Nike, è leader mondiale nel settore dell'abbigliamento sportivo e il suo iconico "swoosh" simboleggia eccellenza sportiva ma è anche icona di stile. Deve molto del suo successo alle collaborazioni con i più grandi atleti di tutti gli sport, da Micheal Jordan per il basket a Serena Williams per il tennis.

Il brand dell'Oregon è un importante esempio, perché è riuscita a farsi largo anche nella cultura pop, ampliando i suoi confini non solo rimanendo nel settore dello sport ma influenzando il mondo della moda. Attraverso collaborazioni con designer conosciuti in tutto il mondo per la produzione di scarpe o collezioni limitate, sono riusciti ad uscire dalla sola sfera sportiva e ad essere riconosciuta come vero e proprio brand di moda. Nonostante i risultati dell'ultimo anno non siano positivi in termini di ricavi, che sono scesi del 2% registrando una dei peggiori risultati degli ultimi anni²⁵, Nike rimane una delle aziende di moda con maggior fatturato.

Grafico 5: Valore del marchio dei principali 10 brand di abbigliamento a livello mondiale nel 2023

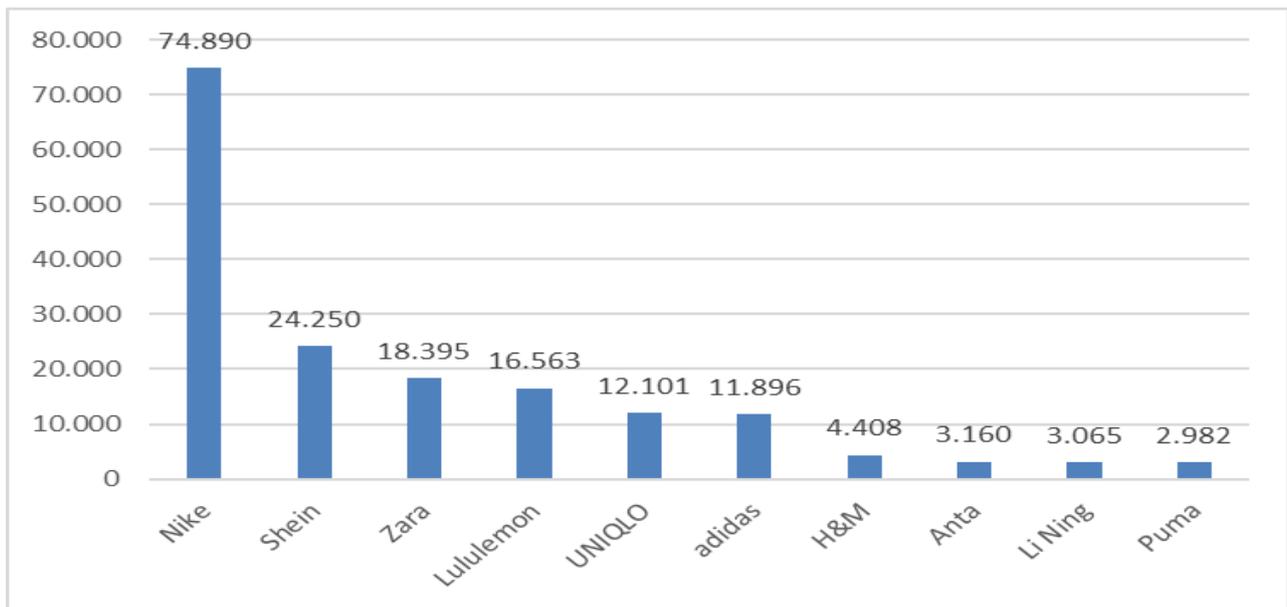
²² Camurati, F. (2024, March 13). Inditex vola nel 2023, ricavi e utili corrono a doppia cifra. Zara a +10% - MilanoFinanza News. Retrieved from MF Milano Finanza website:

<https://www.milanofinanza.it/fashion/inditex-vola-nel-2023-ricavi-e-utili-corrono-a-doppia-cifra-zara-a-10-202403131106285838>

²³ "In Crescita Del 13,5% Anche Febbraio-Marzo 2023: Inditex: Nel 2022 I Ricavi Salgono Del 17,5% a 32,6 Miliardi." www.fashionmagazine.it, 15 Mar. 2023, www.fashionmagazine.it/business/vendite-in-crescita-del-135-tra-febbraio-e-marzo-inditex-nel-2022-i-ricavi-salgono-del-175-a-326-miliardi-111260.

²⁴ Statista. "Brand Value of the Leading 10 Apparel Brands Worldwide 2019 | Statista." Statista, Statista, 2019, www.statista.com/statistics/267931/brand-value-of-the-leading-10-apparel-brands-worldwide/.

²⁵ Finanza, MF Milano. "Nike, Ricavi a 12,6 Miliardi (-2%), Confermando Le Previsioni a Ribasso - MilanoFinanza News." MF Milano Finanza, 28 June 2024, www.milanofinanza.it/fashion/nike-ricavi-a-12-6-miliardi-2-confermando-le-previsioni-a-ribasso-202406281128592617. Accessed 18 July 2024.



Fonte: Kantar²⁶.

Questa continua crescita di domanda per gli items proposti dalle aziende di fast fashion è un trend che è presente da anni e che ha rivoluzionato per sempre il mondo della moda; infatti, marchi come la già citata Zara, H&M e moltissimi altri hanno rivoluzionato il paradigma dell'inaccessibilità della moda alle persone. Infatti, questi colossi riescono ad offrire capi e accessori alla moda a dei prezzi accessibili ad un grande pubblico, permettendo loro di esplorare e adottare nuovi stili, e magari dandogli la possibilità di poter avere accesso ad un vasto mondo di combinazioni e di capi, che riescono di più a incarnare il loro essere o a contribuire ad esprimerlo al meglio. La maggior parte di queste catene che hanno reso accessibile la moda ad un pubblico molto ampio, sono accumulate dal fatto che si trattano di aziende fast fashion.

Il fast fashion si basa su una catena di approvvigionamenti molto efficiente, che permette alle persone solo pochi giorni dopo o comunque poche settimane dopo di indossare gli stili che hanno sfilato nelle passerelle. Inoltre, avendo una rotazione molto veloce, i clienti sono incentivati e curiosi a visitare frequentemente i negozi cercando le ultime novità.

Perciò il fascino di queste aziende fast fashion risiede nella possibilità che danno ai consumatori di poter sperimentare stili diversi, abbinare vari items e di poter avere una vasta gamma di vestiti e accessori nell'armadio ad un bassissimo costo.

In aggiunta una caratteristica che ha contribuito a far aumentare a dismisura l'utilizzo di capi fast fashion, come già detto in precedenza, è l'accessibilità. Ma non solo l'accessibilità di prezzo ma anche

²⁶ Statista. "Brand Value of the Leading 10 Apparel Brands Worldwide 2019 | Statista." Statista, Statista, 2019, www.statista.com/statistics/267931/brand-value-of-the-leading-10-apparel-brands-worldwide/.

la disponibilità quasi immediata ad ottenere i capi di abbigliamento grazie alla sempre di più costante presenza e importanza dell'e-commerce. L'ascesa di quest'ultimo ha contribuito in maniera significativa alla crescita non solo del fast fashion ma della moda in generale, perché è riuscita ad ampliare il bacino d'utenza grazie ad un abbattimento dei confini geografici e portando il consumatore a vivere una vera e propria esperienza d'acquisto totalmente personalizzata.

1.3 E-COMMERCE

Inizialmente i siti e-commerce erano delle semplici piattaforme statiche ma al giorno d'oggi, con l'aiuto della tecnologia sempre più avanzata, sono diventati degli ecosistemi digitali in grado di offrire una customer experience personalizzata e coinvolgente. Le piattaforme e-commerce riescono, attraverso strumenti avanzati come algoritmi di intelligenza artificiale e machine learning, a raccomandare prodotti partendo dalle preferenze individuali del consumatore e ad offrirgli la miglior esperienza basandosi sui suoi comportamenti d'acquisto.

La nascita del commercio elettronico risale agli anni '70 con strumenti tecnologici come la telegrafia e l'Electronic Data Interchange²⁷. Queste tecnologie permettevano alle aziende di scambiarsi i documenti commerciali tra di loro e quindi di concludere delle operazioni tramite l'uso della tecnologia. Ma il primo esempio di e-commerce come lo conosciamo al giorno d'oggi è il "tele-shopping", inventato da Michael Aldrich nel 1979. Egli inventò un sistema che collegava la televisione a un computer sfruttando la rete telefonica e permetteva così agli utenti di avere a disposizione un catalogo di oggetti e fare gli acquisti direttamente dal televisore. Perciò, questa è la prima forma di e-commerce, che però, a differenza dello shopping online di oggi, non sfrutta una connessione internet.

Già negli anni '90 nascevano le prime piattaforme e-commerce simili a quelle che ognuno di noi usa quotidianamente, tra cui le più famose erano Amazon e eBay, ma in generale c'era ancora molto scetticismo e poca fiducia nelle transazioni online. La nascita di questi siti e-commerce fu resa possibile nel 1991 grazie al National Science Foundation²⁸ degli Stati Uniti, che permise l'uso commerciale di internet, consentendo alle aziende e agli individui di utilizzare internet per scopi economici.

Già nei primi del 2000, con la diffusione di internet e la crescita tecnologica, l'e-commerce ha acquistato sempre più importanza nel panorama mondiale con il conseguente maggior investimento delle aziende nelle piattaforme per cercare di sfruttare i nuovi trend di acquisto.

Man mano, con il passare degli anni, si è arrivato ad oggi nel quale strumenti complessi di intelligenza artificiale e machine learning riescono a personalizzare l'esperienza d'acquisto ad ogni singolo consumatore, proponendogli prodotti che potrebbero essere di suo interesse e facendo scoprire ad essi una vasta gamma di items che probabilmente non conoscevano.

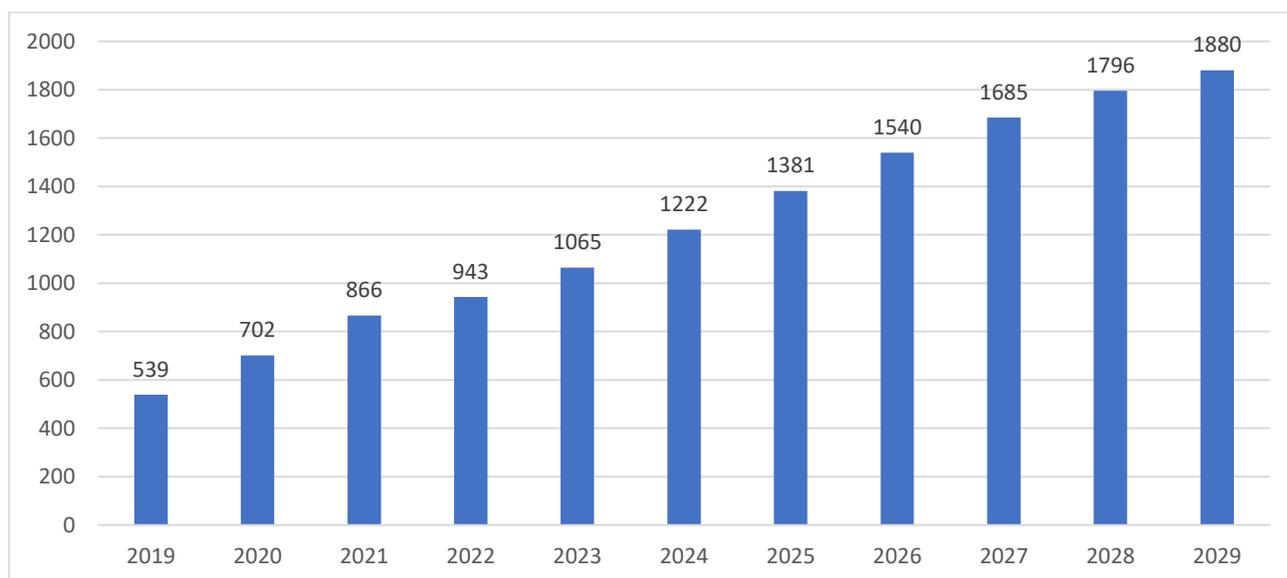
²⁷ È un sistema per lo scambio elettronico di documenti commerciali tra aziende. Automatizza e standardizza il trasferimento di dati, riducendo errori e tempi di elaborazione

²⁸ L'NSF è una agenzia governativa statunitense che finanzia e aiuta la ricerca scientifica e supporta ricerche in vari ambiti tecnologici

Rispetto agli anni precedenti c'è molta più fiducia nelle transazioni online perché sono state apportate delle migliorie sulla sicurezza dei pagamenti, introducendo su tutti l'autenticazione a due fattori²⁹. In aggiunta, molti siti permettono di usufruire di vari metodi di pagamento, dando anche la possibilità di usufruire di servizi come PayPal. Perciò il consumatore si sente sicuro ad acquistare via internet e questa fiducia ha permesso ad un bacino molto ampio di utenti di fare acquisti online.

Oramai ogni settore sta beneficiando delle piattaforme e-commerce e grazie al loro utilizzo stanno aumentando i propri ricavi. L'utilizzo delle piattaforme sta sempre più crescendo e continuerà anche nei prossimi anni come si evince dalla ricerca di Statista Digital Market Insights sui ricavi negli Stati Uniti previsti fino al 2029.

Grafico 6: Ricavi dell'industria dell'e-commerce negli USA dal 2019 al 2029 (in miliardi di dollari statunitensi)



Fonte: Statista Digital Market Insights³⁰

Questi siti permettono al cliente di vedere tutto il catalogo e di confrontare i prezzi, che può così trovare il prezzo migliore, contribuendo in modo incisivo al miglioramento dell'esperienza di acquisto, che non si esplica solo nella possibilità di poter comprare vestiti e accessori comodamente da casa, ma nel permettere tutte queste migliorie sia sulla scelta che sull'esperienza.

Di solito per quanto riguarda la sfera della moda le aziende si affidano all'ambiente del negozio per creare un'esperienza di acquisto. Attraverso l'arredamento, l'illuminazione, la musica e i profumi i

²⁹ Il cosiddetto A2F è un metodo di autenticazione che ti permette di accedere ad un determinato account o sistema solo dopo aver chiesto due diverse forme di accesso

³⁰ Coppola, D. (2023). U.S. e-commerce market size 2016-2022 | Statista. Retrieved from Statista website: <https://www.statista.com/statistics/272391/us-retail-e-commerce-sales-forecast/>

vari brand nei propri negozi riescono a creare un ambiente che riflette la loro identità. Questi particolari negli store non solo trasmettono i valori del brand, ma influenzano le intenzioni di acquisto dei consumatori. Più precisamente le caratteristiche del negozio suscitano delle emozioni ai clienti, portandoli ad acquistare. Perciò, è un potente strumento per invogliare e incentivare le persone all'interno dello store.

Nonostante l'importanza che essi continuano ad avere “negli ultimi anni abbiamo assistito a uno sviluppo sostanziale per quanto riguarda il commercio elettronico di prodotti di moda”³¹. Soprattutto perché l'utilizzo di Internet e delle piattaforme online è visto come un mezzo efficace nel processo di internazionalizzazione, poiché diminuisce significativamente le responsabilità fisiche, le carenze informative e i rischi operativi.

Perciò è importante utilizzare sia i canali fisici che le piattaforme di e-commerce in modo da poter attuare delle strategie differenti ma comunque efficaci. L'integrazione omnicanale permette alle aziende di offrire un'esperienza di acquisto coerente tra i vari canali di vendita, in modo da consentire ai consumatori di spostarsi tra di essi.

Tra i servizi più utilizzati che combinano gli store fisici e quelli online, c'è il “click and collect”. Questa pratica permette ai clienti di ordinare nel sito e di poter ritirare il prodotto presso il negozio, risparmiando i costi di spedizione e allo stesso tempo di poter avere il prodotto nel minor tempo possibile anche se non è presente in negozio.

Per queste ragioni è importante riuscire, oltre che a offrire un servizio efficace nei negozi fisici, a migliorare e rendere efficienti i siti e-commerce in modo da poter avviare iniziative omnicanale.

Come sottolineato in precedenza, internet consente a tutte le aziende, anche quelle di piccole dimensioni, di poter accedere ai mercati globali senza avere uno store in ogni paese. In questo modo, nonostante contribuisca ad aumentare la competizione nel settore, permette a tutte le aziende di poter essere presenti in tutti i paesi del mondo senza costi legati alle aperture di negozi fisici, magazzini e uffici. Le piattaforme, attraverso l'utilizzo di tecnologie avanzate come la blockchain³² e l'Internet of Things³³, possono abbassare sensibilmente il rischio percepito di prodotti falsificati, in particolare nel

³¹ Guercini, S., Bernal, P. M., & Prentice, C. (2018). New marketing in fashion e-commerce. *Journal of Global Fashion Marketing*, 9(1), 1–8. <https://doi.org/10.1080/20932685.2018.1407018>

³² È una tecnologia decentralizzata che permette di creare registri digitali sicuri. Ogni blocco della catena contiene un insieme di transazioni che sono verificate e collegate in ordine cronologico. Essa fonda il suo successo sul fatto che garantisce trasparenza e sicurezza e per questo ha il potenziale di rivoluzionare molti settori, riducendo le frodi e rendendoli più sicuri ed efficienti.

³³ L' IoT fa riferimento a tutti quegli oggetti che comunicano tra loro tramite la rete internet. Essi riescono quindi a scambiarsi informazioni e dati in tempo reale. L'Internet of Things permette quindi di automatizzare i processi e migliorare l'efficienza, rendendo gli ambienti e i sistemi interconnessi tra loro.

settore della moda di lusso che è spesso vittima della contraffazione di prodotti. Grazie a queste due tecnologie i siti e-commerce riescono a garantire una migliore tracciabilità e a garantire l'autenticità dei prodotti evitando così numerose conseguenze negative per le aziende come grosse perdite economiche causate dalla vendita di prodotti non originali e soprattutto riducendo il rischio di danneggiamento della reputazione del brand, che è messa in forte rischio a causa della pessima qualità dei prodotti che erroneamente possono essere associati alla qualità del brand originale. Perciò le aziende, utilizzando piattaforme online sicure e affidabili, possono controllare al meglio la distribuzione dei prodotti, assicurandosi che i loro items originali raggiungano l'acquirente.

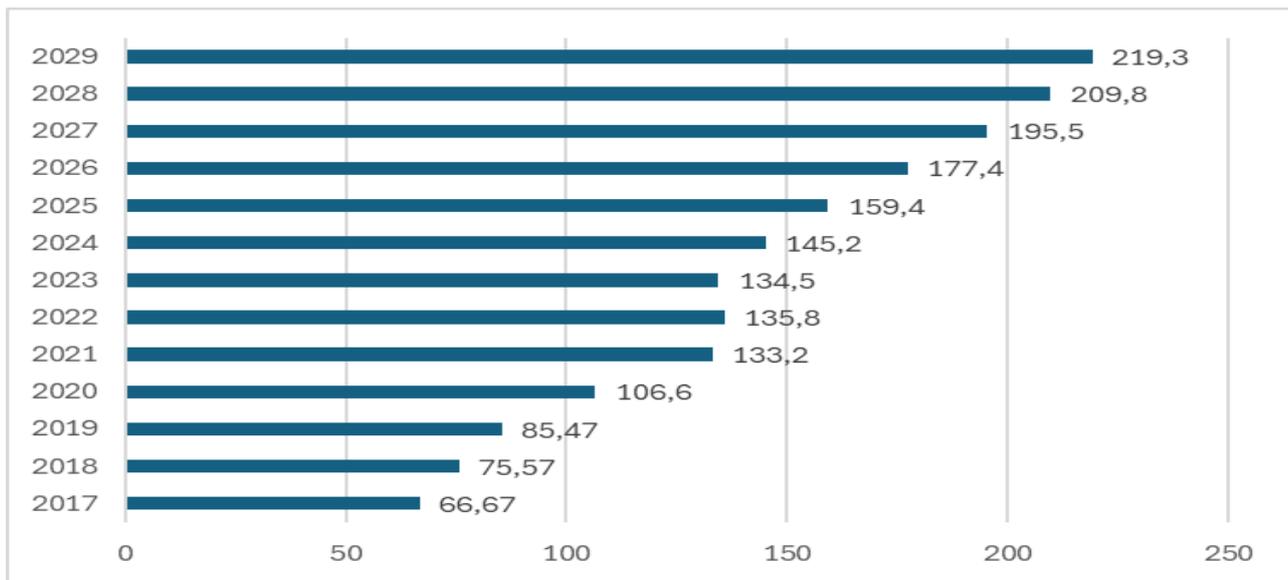
Oltre ad evitare i rischi esposti precedentemente, l'e-commerce è uno strumento che può creare delle opportunità per l'azienda, permettendo al retailer di espandere il proprio business, dato che permette un grosso pubblico di poter conoscere la propria attività e consente agli utenti di comprare con un semplice click. Questa facilità d'uso e tutti i vantaggi che l'e-commerce porta al consumatore ha conseguito ad un forte aumento delle vendite attraverso queste piattaforme, e a sottolineare ciò sono i numeri crescenti sugli acquisti online avvenuti in questi ultimi anni. Più nello specifico le vendite globali e-commerce hanno raggiunto 8000 miliardi di dollari nel 2023, con un incremento del 14,3% rispetto al 2021³⁴. In Italia in particolare nel 2023 si è registrata una crescita del 13% con un valore di 54,2 miliardi di euro³⁵.

È importante sottolineare, come già evidenziato in precedenza, che a inficiare molto su questa crescita delle piattaforme e-commerce è il settore della moda, infatti, questo è il segmento che è cresciuto di più nel 2023 arrivando ad ottenere un +22,3% a livello globale rispetto all'anno precedente. Più nello specifico i dati forniti da Statista Digital Market Insights, confermano che questo trend di acquisti online è cresciuto negli anni precedenti e continuerà a crescere anche nei prossimi anni come si può notare nel grafico seguente.

Grafico 7: Ricavi dell'e-commerce al dettaglio di abbigliamento, calzature e accessori negli USA dal 2017 al 2029 (in miliardi di dollari statunitensi)

³⁴ Alessia Perissinotto. (2023). E-commerce 2024: Previsioni e Strategie d'Investimento | Algoritma. Retrieved July 17, 2024, from Algoritma website: <https://www.algoritma.it/e%0E2%80%91commerce-2024-previsioni-e-strategie-dinvestimento/>

³⁵ Kelly Di Blas. (2024). Rapporto 2023 sull'innovazione digitale in Italia: alcuni progressi inaspettati - BuoneNotizie.it. Retrieved from BuoneNotizie.it website: <https://www.buonenotizie.it/innovazione/2024/02/26/rapporto-2023-sullinnovazione-digitale-in-italia-alcuni-progressi-inaspettati/kelly-di>



Fonte: Statista Digital Market Insights³⁶

Questi dati sottolineano perciò di come sia importante per un'azienda che opera nel settore della moda, di avere una piattaforma e-commerce che permetta l'acquisto online dei suoi prodotti. Oltretutto in un contesto e in un mercato sempre in continua evoluzione e fortemente competitivo come quello del settore della moda, l'offrire questa tipologia di servizio non è più un'opzione, ma un elemento essenziale per avere successo. Perciò, anche le firme più importanti nel panorama mondiale hanno dovuto adattarsi al progressivo rinnovamento e cambiamento del mercato e dunque, per mantenere la posizione nel mercato in un settore pieno di player, hanno dovuto anche loro utilizzare siti e-commerce.

³⁶ Van Gelder, K. (2024). U.S. online apparel and fashion market size 2022. Retrieved from Statista website: <https://www.statista.com/statistics/278890/us-apparel-and-accessories-retail-e-commerce-revenue/>

1.4 AI NEL SETTORE DELLA MODA

Non è la prima volta che l'AI ricopre un ruolo da protagonista nel mondo del fashion, infatti, oltre che nell'e-commerce, è sempre più utilizzata dalle industrie di moda in tutti i suoi processi, dalla progettazione dei capi fino alla fase finale di vendita. Perciò essa non è solo uno strumento ma è motore principale di innovazione durante tutto il processo. Le ragioni di questo suo impiego possono essere spiegate dalla sua capacità di elaborare, gestire e analizzare grandi quantità di dati, identificare modelli complessi e generare previsioni accurate o comunque più precise rispetto ai risultati raggiungibili dalle sole capacità umane. Ciò permette alle aziende del settore di adattarsi rapidamente ai cambiamenti del mercato. È importante però sottolineare che l'AI non può ancora eguagliare l'innata e profonda creatività umana, che ha permesso in passato e che permette tutt'ora di ideare ed esplorare nuovi orizzonti dello stile, combinando varie sfaccettature di stili diversi che riescano a fare breccia sul pubblico. Inoltre, come già evidenziato precedentemente, il design e la moda sono un riflesso delle dinamiche socioculturali e quindi gli stilisti devono essere in grado di interpretare tradizioni, simboli e tendenze che l'intelligenza artificiale può parzialmente intuire.

Proprio nelle primissime fasi, l'intelligenza artificiale ha un ruolo ponderante nell'aiutare lo stilista nella progettazione dei capi e nel prevedere possibili trend e stili che potrebbero ispirare la creazione di nuovi modelli. Sono state fatte molte ricerche accademiche su questa collaborazione tra stilista e macchina e tra le più importanti c'è sicuramente lo studio³⁷ svolto nel 2021 pubblicato sulla rivista "Fashion and Textiles". Questa ricerca mira ad approfondire come, partendo da delle immagini di moda presenti nei social media, l'utilizzo dell'apprendimento automatico riesca a prevedere futuri trend e stili. Partendo da un dataset di immagini recuperate dai social, più precisamente da Instagram e da Pinterest, i ricercatori hanno inizialmente etichettato manualmente le foto indicando stile, brand e stagione e, dopo aver fatto ciò, hanno addestrato il modello di apprendimento automatico, più nel dettaglio uno basato sul CNN³⁸, per estrarre le caratteristiche dalle foto e prevederne il successo futuro. Altri studi invece, a differenza di quest'ultimo appena descritto, che mira a utilizzare AI per analizzare e prevedere le tendenze basandosi su dati esistenti dai social media, sono incentrati sull'innovazione diretta nel design utilizzando l'AI per generare e a perfezionarli secondo specifici criteri estetici e funzionali. Questo è il caso della ricerca "'Unlocking the Potential of Artificial Intelligence in Fashion Design and E-Commerce Applications: The Case of Midjourney" che

³⁷ Tu, Y., Tang, X., & Ling, K. (2021). Using machine learning to forecast fashion trends from social media images. *Fashion and Textiles*, 40(4), 507-522.

³⁸ Le reti neurali convoluzionali sono delle reti neurali profonde spesso utilizzate per l'elaborazione e l'analisi di immagini. Il loro punto di forza è che riescono a ridurre la dimensionalità delle immagini comunque mantenendone le caratteristiche fondamentali, giovando così la ricerca soprattutto per studi sulla classificazione e sul rilevamento facciale

approfondisce l'impatto dell'AI Midjourney³⁹ sulla creazione di capi visivamente accattivanti e sottolinea il potenziale rivoluzionario dell'AI nel settore della moda, migliorando la generazione di concetti e l'ottimizzazione del design.

Queste sono solo due delle moltissime ricerche riguardanti come l'intelligenza artificiale sia utilizzata nel mondo della moda per fare delle previsioni su stili e sulla loro popolarità futura. Ci sono però moltissimi altri utilizzi dell'AI a supporto anche delle aziende non strettamente legate alla moda, questo è il caso del colosso Amazon che sta implementando e sviluppando vari sistemi di apprendimento automatico utili nell'individuare o cercare addirittura di "plasmare" le ultime tendenze. Tra questi sistemi ci sono alcuni molto innovativi, come quello sviluppato da dei suoi ricercatori in Israele che riesce, analizzando le etichette degli indumenti, a classificare un look come elegante o meno in modo da poter dare dei consigli per eventuali modifiche, suggerendo accessori o capi alternativi che possano migliorare l'aspetto complessivo⁴⁰.

Ma l'utilizzo dell'intelligenza artificiale spazia non solo nella previsione di nuovi stili e in classificazioni di determinati modi di vestire, ma può essere di grande importanza anche per aspetti più legati all'economia del business, come la previsione della domanda. Questo aspetto è molto utile vista l'ascesa di questo settore nel panorama globale e il continuo aumento di persone interessate a questo mercato. Nonostante la domanda nell'ambito della moda sia molto complicata da predire perché essa si compone di stagionalità, un'alta propensione agli acquisti spontanei e una vita molto breve dei capi, ci sono vari studi che cercano di fare delle previsioni il più accurate possibili sulla domanda. Ad esempio, in uno studio⁴¹ condotto da Kritika Swaminathan e Rakesh Venkitasubramony, si è analizzato le varie tecniche utilizzate per questo scopo, sia qualitative come il metodo Delfi⁴², che metodi quantitativi fino ad arrivare alle tecniche che utilizzano il machine learning. La ricerca mostra come ci siano varie caratteristiche che impattano la domanda futura, come la presenza di informazioni online, di immagini, singole caratteristiche dei prodotti e altre variabili.

³⁹ è una piattaforma per la creazione di arte generativa con l'intelligenza artificiale. Essa utilizza come input delle frasi o delle parole per poi elaborarle e trasformarle in immagini. Questa tecnologia è molto utilizzata dagli artisti e stilisti che possono sperimentare stili e temi innovativi.

⁴⁰ Knight, W. (2017, August 24). Amazon Has Developed an AI Fashion Designer. Retrieved from MIT Technology Review website: <https://www.technologyreview.com/2017/08/24/149518/amazon-has-developed-an-ai-fashion-designer/>

⁴¹ Swaminathan, K., & Venkitasubramony, R. (2023). Demand forecasting for fashion products: A systematic review. *International Journal of Forecasting*, 40(1). <https://doi.org/10.1016/j.ijforecast.2023.02.005>

⁴² È una tecnica strutturata, sviluppata già negli anni '50, per raccogliere previsioni tramite la somministrazione di questionari. Gli esperti rispondono anonimamente, e dopo ogni ciclo, un facilitatore riassume le risposte e le ridistribuisce al gruppo per ulteriori commenti e revisioni. Questo processo iterativo continua fino a raggiungere un consenso o una comprensione chiara delle opinioni divergenti.

Questo è solo un esempio di come l'intelligenza artificiale possa supportare l'azienda nel trovare tutte quelle caratteristiche che influenzano la domanda futura, e stimarla il più accuratamente possibile in base alla loro presenza o meno.

Ancora, l'utilizzo dell'AI svolge un ruolo cruciale nella produzione e nella logistica. I sistemi di intelligenza artificiale, tramite vari algoritmi di machine learning o deep learning, possono ottimizzare i processi di produzione. Una maggiore efficacia in questi processi permette alle aziende di ridurre gli sprechi e risparmiare molto denaro che può essere investito per nuove iniziative o per migliorare ancor di più la produzione.

Quindi riassumendo, grazie alla sua capacità di elaborare grandi quantità di dati e riconoscere pattern e modelli complessi, l'AI viene spesso utilizzata a supporto delle case di moda durante tutta la filiera produttiva, dall'idea fino alla vendita finale.

Nonostante tutti i benefici che l'intelligenza sta apportando al settore della moda, esiste una significativa diffidenza sul suo utilizzo. Questi dubbi possono essere attribuiti a vari fattori come le preoccupazioni etiche, timori legati all'annullamento della creatività umana, e la percezione di un possibile impatto negativo sull'occupazione.

Il timore principale è quello legato alla possibile diminuzione della creatività umana, dato che si sta affidando sempre più compiti agli algoritmi, il che potrebbe portare ad un omogenizzazione delle tendenze, riducendone così la varietà e l'originalità. Perciò, è importante che l'uomo e l'intelligenza artificiale collaborino e non si sostituiscano in modo da lasciare la creatività agli stilisti ma con il supporto degli algoritmi per la previsione di, possibili nuovi trend.

Un'altra fonte di diffidenza è il timore che con l'introduzione dell'intelligenza artificiale comporti una perdita di posti di lavoro perché c'è la preoccupazione che essa sostituisca il lavoro umano, rendendo alcune mansioni inutili. Questo fenomeno potrebbe essere rilevante nei settori come la produzione e il retail, nei quali l'intelligenza artificiale potrebbe automatizzare i processi che attualmente sono puramente manuali.

Per evitare ciò, è importante che le aziende investano nella riqualificazione dei lavoratori, permettendo loro di sviluppare delle competenze tecnologiche in grado di permettergli di essere comunque necessari per l'azienda.

In conclusione, quindi è importante sottolineare che nonostante ci siano delle preoccupazioni riguardo l'introduzione dell'intelligenza artificiale nel contesto lavorativo, ci sono molte soluzioni che vanno a vantaggio sia dell'azienda che dei lavoratori stessi.

CAPITOLO 2: LITERATURE REVIEW

2.1 L'AI NELL'E-COMMERCE

L'intelligenza artificiale ha trasformato radicalmente il mondo della moda. Come visto nello scorso capitolo gli utilizzi di questi algoritmi spaziano dalla creazione del vestiario o dell'accessorio fino alla vendita finale.

Una finestra di utilizzo dell'AI che ha forte riscontro in termini di ricavi e di fidelizzazione del cliente è nelle piattaforme e-commerce. Essa, infatti, offre una serie di strumenti che migliorano la customer experience, aumentando così le vendite e di conseguenza gli introiti dell'azienda.

Le grandi potenzialità di algoritmi di machine learning e di deep learning combinati con il grande e crescente successo dei siti e-commerce, sono una grande opportunità per le aziende per emergere in questo mercato. Già molti studi e ricerche sottolineano l'efficacia dell'intelligenza artificiale e come già in questo periodo gli utilizzi di essa siano molteplici.

Una delle applicazioni più rivoluzionarie che combina l'AI con i siti e-commerce è la personalizzazione del prodotto che permette così al consumatore di scegliere e modificare le varie componenti e aspetti dello stesso item tra i quali il colore e il tessuto. Ciò si traduce in un vantaggio sia per il consumatore, che riesce così ad avere un prodotto totalmente personalizzato a suo gusto e a vivere un'esperienza d'acquisto più coinvolgente, sia per l'azienda che riesce a ridurre gli sprechi fornendo un item che sicuramente, o con grande probabilità, piacerà al consumatore finale.

Un esempio di successo di personalizzazione del prodotto è sicuramente "Nike by you"⁴³, precedentemente chiamato NikeiD, che permette appunto ai possibili clienti di poter personalizzare la propria scarpa selezionando materiali, colori e altri elementi di design. L'AI in questo caso assicura che il prodotto finale rispecchi le caratteristiche selezionate dal consumatore e che sia esteticamente piacevole. Questa nuova possibilità di personalizzazione del prodotto non è solo un servizio fornito dalle grandi marche, ma anche piccole start-up e designer emergenti stanno sperimentando questa nuova soluzione per emergere in un mercato molto competitivo e con player molto forti e già affermati.

Gli utilizzi dell'intelligenza artificiale sono molteplici e stanno rivoluzionando l'esperienza d'acquisto dei vari consumatori, come l'innovativo servizio offerto da alcuni colossi della moda, tra i quali Asos, Zara e molti altri, chiamato "Virtual Try-On". Questa nuova applicazione dell'AI

⁴³ Originariamente era stato introdotto nel 1999, ma negli ultimi anni questo servizio è stato migliorato permettendo una migliore esperienza ai consumatori. Soprattutto grazie alla visualizzazione 3D che consente ai consumatori di vedere il prodotto in tutti i suoi dettagli

permette ai consumatori di provare virtualmente vestiti o accessori. La tecnologia utilizza avanzate tecniche di realtà aumentata e di modellazione 3D in modo da poter rappresentare al meglio come un determinato abito apparirebbe sulla persona interessata all'acquisto. Questo servizio, perciò, non migliora solo l'esperienza di acquisto del consumatore, ma permette a quest'ultimo di fare scelte più informate, aumentando così la soddisfazione per i prodotti acquistati.

In particolare, questa applicazione dell'AI sta riscontrando molto successo perché riesce a combinare due caratteristiche che sono molto valutate, come la convenienza e la velocità. Essa, infatti, permette a chi naviga semplicemente con un computer o con uno smartphone comodamente da casa propria, di poter esplorare e addirittura indossare degli indumenti senza dover andare nel negozio fisico. Questo livello di personalizzazione riduce significativamente il rischio di insoddisfazione legata agli acquisti online, e aumenta la soddisfazione e la fiducia nel processo di acquisto. Essa riduce drasticamente l'incertezza riguardo la vestibilità e l'aspetto di un determinato capo, e perciò migliora la soddisfazione del cliente che acquistando online solitamente era minata dall'impedimento di provare un vestito. Il "Virtual Try-On" aiuta inoltre le aziende a ridurre drasticamente i resi che rappresentano un costo significativo per le aziende. Infatti i resi, non solo comportano dei costi logistici e di spedizione, ma possono implicare la gestione di capi danneggiati e non più rivendibili.

Attraverso questa applicazione dell'AI quindi possono aumentare l'efficienza e di conseguenza migliorare i margini di profitto. In aggiunta, la riduzione dei resi, riducendo il numero di spedizioni e di smaltimento di prodotti che devono essere riparati o smaltiti, contribuisce ad una catena più sostenibile. Questa caratteristica può contribuire a migliorare l'immagine dell'azienda agli occhi dei consumatori, che la considererebbero attenta alle questioni ambientali.

L'AI per il "Virtual Try-On" elabora i dati di input, partendo dalle misure del corpo del cliente, in modo da far visualizzare come i capi si adatterebbero sul corpo e come il tessuto si adatta alla forma del corpo.

Questi sono solo degli esempi di come l'intelligenza artificiale sta rivoluzionando il rapporto tra e-commerce e i consumatori, trasformando la comunicazione in un'esperienza sempre più personalizzata, in modo da incrementare la soddisfazione del cliente e il legame con il marchio o casa di moda.

Uno dei principali utilizzi dell'AI è proprio nella personalizzazione della comunicazione, e secondo uno studio di McKinsey le aziende che puntano molto su questo servizio e che eccellono nel fornirlo aumentano i ricavi del 40% rispetto alle aziende che non la utilizzano⁴⁴. Perciò è molto importante

⁴⁴ Harreis, H., Koullias, T., Roberts, R., & Te, K. (2023, March 8). Generative AI in Fashion | McKinsey

per le imprese del settore investire per migliorare questo aspetto, in modo da essere competitivi sul mercato e aumentare di conseguenza i propri ricavi.

Un'efficace personalizzazione del sito e-commerce permette, attraverso le interazioni, il comportamento di navigazione, i precedenti acquisti e le varie preferenze, di prevedere gli indumenti o gli oggetti che potrebbero interessare il consumatore, offrendogli oltretutto delle offerte personalizzate e promozioni. Vari studi descrivono il fondamentale ruolo della personalizzazione, sottolineando come questa pratica possa contribuire al successo. Ad esempio, nella ricerca intitolata "The Power of Personalization: Enhancing User Experience in E-commerce" viene evidenziata l'importanza del machine learning per migliorare l'engagement e la soddisfazione del consumatore. L'articolo in questione analizza l'impiego che hanno gli algoritmi di machine learning nel guidare gli utenti attraverso la vastissima offerta di prodotti nei vari siti e-commerce e nel trovare prodotti che meglio incarnano i loro desideri. Sottolinea però l'importanza della salvaguardia dei diritti e della privacy dell'utente, che è un argomento molto discusso e in voga quando si parla di personalizzazione del servizio sulle pagine online.

La privacy è un argomento tanto discusso per la sua grande importanza perché, prima di tutto, essa permette agli individui di mantenere segrete o comunque non rendere pubbliche in modo non voluto le proprie informazioni personali, che sono "il nuovo petrolio"⁴⁵.

Perciò, dato che i siti e-commerce che utilizzano strategie di personalizzazione, per farlo hanno necessità di raccogliere e analizzare una vasta quantità di dati personali, c'è molta diffidenza sull'utilizzo delle informazioni e di conseguenza paura che sia infranto il diritto alla privacy. La trasparenza sull'utilizzo dei dati personali è molto importante e inficia tanto positivamente quanto negativamente sulla soddisfazione del cliente, infatti, più l'azienda o il sito sono chiari e meno sarà la preoccupazione dell'utente, che potrà semplicemente godersi l'esperienza di acquisto senza doversi preoccupare dell'utilizzo dei suoi dati che potrebbero compromettere la sua privacy. Al contrario, una mancanza di trasparenza sull'utilizzo dei dati personali, influenza negativamente il giudizio e l'esperienza del consumatore. Queste preoccupazioni, inoltre si riflettono anche sulla percezione dell'azienda, che sarà vista come poco sicura e quindi di scarsa qualità.

Per affrontare queste ansie, una gran parte delle aziende sta implementando una politica sulla sicurezza dei dati più rigorosa. Questa include l'adozione di misure di sicurezza avanzate e la possibilità per gli utenti di poter controllare le proprie informazioni personali in qualsiasi momento.

⁴⁵ Celebre frase che disse nel 2006 il famoso matematico Clive Humby

Molte aziende, in aggiunta, cercano di ottenere il consenso esplicito del consumatore prima di raccogliere ed utilizzare i loro dati.

Oltre a queste misure, alcune aziende stanno adottando dei sistemi di sicurezza molto avanzati per proteggere i dati dei consumatori, tra i quali crittografia dei dati⁴⁶, firewalls⁴⁷ e monitoraggio continuo per evitare violazioni o accessi non autorizzati ai dati.

Perciò, se c'è trasparenza nell'utilizzo dei dati personali e viene fatta una buona personalizzazione dell'esperienza di acquisto nel sito e-commerce, l'azienda beneficerà di un grosso impatto dal sito stesso. Ma per poter ottenere una buona e soprattutto efficace personalizzazione è necessaria una segmentazione ottimale dei prodotti, in modo da poter classificare o comunque creare dei cluster di consumatori per offrirgli la miglior esperienza di acquisto. Oltre a ciò, una volta individuati i segmenti, è molto importante che l'azienda intraprenda azioni di marketing ben indirizzate ed efficaci.

Segmentare bene i prodotti, che si tratti di vestiti o di accessori come delle borse, non è una sfida facile. Soprattutto la varietà di items presenti nei vari siti può mettere in difficoltà il consumatore a trovare il prodotto specifico che sta cercando o limitare la possibilità che egli scopra prodotti affini a quello di interesse, rovinando così, prima di tutto l'esperienza e, in aggiunta, compromettendo un possibile acquisto di un'alternativa o di un oggetto aggiuntivo. D'altra parte, però una segmentazione troppo dettagliata rischierebbe di confondere e rendere introvabile quel determinato item, peggiorando ancor di più la customer experience.

Probabilmente un utente che è confuso e si fa sormontare dalle centinaia di articoli presenti nel sito, abbandonerà il sito compromettendo non solo questa sessione di navigazione ma anche le future, associando quel determinato e-commerce alla sensazione di sconforto e confusione provata, e preferendo così optare per una alternativa.

Perciò la chiave di una perfetta segmentazione sta nel riuscire a fondere e bilanciare una precisa segmentazione con la semplicità di navigazione.

La facilità di utilizzo è di fondamentale importanza per il successo di una piattaforma e-commerce perché permette al consumatore di poter fare una ricerca intuitiva senza difficoltà. Come sottolineato in aggiunta dall'articolo "Enhanced E-commerce Customer Engagement: A Comprehensive Three-Tiered Recommendation System", questa tipologia di ricerca, appunto definita intuitiva, è un

⁴⁶ Proteggono i dati sia in transito che in archivio, rendendo le informazioni illeggibili per chiunque non abbia l'autorizzazione necessaria.

⁴⁷ sono dispositivi di sicurezza della rete progettati per monitorare e controllare le attività in entrata e in uscita, riuscendo ad intercettare e a prevenire accessi non autorizzati

meccanismo chiave per migliorare l'engagement e la soddisfazione del cliente. Una ricerca intuitiva si raggiunge quando c'è una buona segmentazione unita ad una semplicità di utilizzo e quindi, queste due caratteristiche sono le chiavi per il successo.

In generale un'efficace segmentazione si ottiene con la suddivisione dei prodotti in categorie ben definite. Essa può essere basata su varie caratteristiche, come il prezzo, il tipo di prodotto, la marca, il colore ed altre specifiche. Si può considerare una segmentazione ben riuscita quando essa riduce il tempo e lo sforzo necessario per trovare quel determinato item ricercato.

Questa pratica deve essere di forte interesse e di continuo rinnovamento perché oltre ad attirare nuovi clienti aiuta a mantenere quelli esistenti e soprattutto a renderli fedeli al sito. Poiché un'efficace segmentazione aiuta a prevenire l'overload informatico, che si avvera quando il consumatore si sente asfissiato dalla troppa scelta e dalle quantità di opzioni che, è molto importante migliorare e aggiornare questo aspetto nelle piattaforme. Come viene sottolineato dallo studio "Information Load and Consumer Decision Making", la troppa scelta mal organizzata influisce negativamente sulle decisioni del consumatore portandolo a decisioni d'acquisto ritardate o meno ideali per il consumatore.

In generale, nonostante sia fondamentale la classificazione dei prodotti, come sottolineato precedentemente, la maggior parte delle ricerche scientifiche si basano sulla segmentazione della clientela. Molti di questi studi hanno avuto molto successo e riscontro pratico, come ad esempio lo studio condotto da Yulin Deng e Qianying Gao intitolato "A Study on E-commerce Customer Segmentation Management Based on Improved K-means Algorithm". Questo articolo mira appunto a suggerire e approfondire l'uso di determinati algoritmi per migliorare la clusterizzazione del cliente, in modo da poter attuare delle strategie e delle mosse di marketing indirizzate e coinvolgenti per il consumatore. Più nello specifico, testano l'algoritmo SAPK + K-means⁴⁸ che riesce a superare le limitazioni del solo K-means come, ad esempio, la tendenza a cadere in minimi locali (cioè che si accontenta della soluzione migliore solo in parte rispetto alle possibili soluzioni escludendo così dei possibili risultati migliori). Perciò la ricerca mostra come le tecnologie avanzate e metodi integrati portino ad un miglior risultato.

Un altro esempio di ricerca che punta molto sulla clusterizzazione dei clienti con l'aiuto dell'intelligenza artificiale, è quella eseguita da Rachid Ait Daoud, Abdellah Amine, Bouikhalene

⁴⁸ Esso combina il Self-Adaptive Particle K-means con il tradizionale e molto utilizzato, algoritmo K-means. Il SAPK viene introdotto per migliorare l'accuratezza e l'efficienza della clusterizzazione, limitando l'eccessiva sensibilità agli outliers del solo K-means. Questa combinazione di algoritmi viene utilizzata per analisi di grandi e complicati dataset.

Belaid e Rachid Lbibb. Questi ricercatori hanno utilizzato un approccio che combina il LRFM⁴⁹ con altri metodi di clustering, concentrandosi sul trovare i clienti che hanno maggiore probabilità di passare ad un altro sito. Questo fenomeno può avvenire per svariati motivi come offerte più vantaggiose, prodotti o servizi migliori. Individuando preventivamente quei possibili clienti che sono più propensi a cambiare, l'azienda può intervenire con delle mosse di marketing o varie strategie di retention volte a cercare di impedire che essi diventino dei "customers churn".

Un'altra interessante ricerca, che però si focalizza solo sulla segmentazione dei clienti e non si concentra sulla classificazione degli oggetti presenti nel sito e-commerce, è lo studio condotto da Puspitasari et Al intitolato "Online Consumer Segmentation Study Based on Factors Affecting E-Commerce Selection". Questa ricerca mira, attraverso varie tecniche di clustering, ad analizzare e capire il comportamento degli utenti nel sito e-commerce di moda ponendo al centro della ricerca le recensioni lasciate sotto i prodotti nella sezione commenti. Lo scopo è quello di individuare e segmentare quei consumatori più influenti. Per fare ciò, a differenza di tutti gli studi esaminati in precedenza, utilizza un approccio esplorativo e non quantitativo, impiegando il clustering K-means. Così, in questo modo, sono riusciti a classificare i clienti in cluster per tipologia di linguaggio utilizzato. La novità che porta questa ricerca è che non classifica gli utenti in base a preferenze di prodotti o determinate tipologie di prodotti, ma mira a capire gli atteggiamenti e le aspettative dei consumatori partendo dalle loro espressioni e terminologie.

Perciò molto spesso gli articoli e le ricerche si soffermano molto o prediligono una segmentazione sui clienti tralasciando un tassello molto importante che sta alla base dell'utilizzo dell'e-commerce e quindi della corretta riuscita di tutte le iniziative di marketing legate alla clusterizzazione dei clienti. Cioè la classificazione degli oggetti al suo interno, perché infatti, se il sito e-commerce non è progettato al meglio e combina facilità d'uso e personalizzazione, nonostante ci possano essere degli sforzi e un servizio ad hoc per quella tipologia di cliente, questi ultimi preferiranno un'altra piattaforma più intuitiva che non gli faccia sentire confusi.

Uno studio che è riuscito a combinare la classificazione dei prodotti e la segmentazione della clientela è la ricerca intitolata "Fashion Retail Recommendation System by Multiple Clustering". Gli autori hanno proposto una serie di raccomandazioni su queste due tematiche in modo, prima di tutto, da trovare dei pattern nascosti e poter raggruppare i prodotti in cluster simili. Successivamente, hanno suggerito delle tecniche per predire il comportamento del cliente e personalizzare la customer

⁴⁹ L'acronimo di Length, Recency, Frequency, Monetary, è uno strumento molto utilizzato per segmentare i consumatori. Si basa sulla lunghezza della relazione, il tempo intercorso dall'ultimo acquisto, la frequenza di acquisto e il valore monetario degli acquisti.

experience, al fine di aumentare la possibilità di vendita e di fidelizzazione. Il primo obiettivo, quindi è quello di utilizzare l'intelligenza artificiale per analizzare vari attributi di prodotti come il colore, il materiale e il prezzo. Facendo ciò, si è in grado di suggerire articoli con caratteristiche simili con quelli scelti o osservati dal cliente in modo da potergli offrire delle soluzioni simili.

Fare una classificazione efficace dei propri prodotti in vendita nel sito, permette ai rivenditori anche di poter gestire più efficacemente l'inventario e di avere una base solida su cui partire per implementare promozioni o strategie di marketing per tipologie di prodotti alla moda o destinate a nicchie di consumatori.

Per quanto riguarda invece la segmentazione della clientela lo studio basa la sua ricerca sulla segmentazione per preferenze e comportamenti di acquisto. Inoltre, mira a prevedere ed indentificare, attraverso delle tecniche di mining⁵⁰, le preferenze di chi si interfaccia per la prima volta con il sito e-commerce, evitando così il “cold” start⁵¹ tipico di queste piattaforme.

⁵⁰ Sono delle tecniche che mirano ad estrarre dei pattern dai dati, ad esempio il clustering, classificazione e regressione

⁵¹ Si verifica quando un utente entra per la prima volta su un sito e-commerce. Le raccomandazioni di prodotti risultano molto complicate da fornire, perché mancano le informazioni su cui basarsi.

2.2 CLASSIFICAZIONE DEGLI ITEMS DI MODA

Come detto e analizzato attraverso questi studi la maggior parte delle ricerche si sofferma sulla segmentazione della clientela più che soffermarsi sulla classificazione degli items che renderebbe più fruibile il sito e permetterebbe, a qualsiasi tipologia di utente facente parte di qualunque cluster, di poter utilizzare il sito e-commerce senza sentirsi frustrati ed insoddisfatti.

Nonostante siano di numero molto ridotto, per evitare che i consumatori provino queste sensazioni negative, sono stati realizzati alcuni studi per classificare gli oggetti presenti nel sito e-commerce.

Ad esempio, la ricerca intitolata “Large-scale Item Categorization for e-Commerce” mira a fornire delle soluzioni per la clusterizzazione di prodotti su larga scala, in questo caso contenuti nella piattaforma eBay.

Lo studio è stato molto complicato ed ambizioso perché gli items considerati sono stati di 20.000 categorie diverse e alcune di esse molto eterogenee tra loro. Per fare ciò hanno diviso il lavoro in due parti: la prima si è svolta facendo una classificazione grossolana degli oggetti, mentre nella seconda fase hanno svolto un lavoro più di fine partendo dai gruppi latenti creati nella prima fase e classificando gli oggetti all’interno di essi. Questo studio ha perciò dimostrato che questo metodo è efficace per una grande quantità di dati. Chiaramente per questa ricerca, essendo stata fatta su tutti gli oggetti presenti su una piattaforma come eBay, la classificazione non può essere così approfondita e specifica. Nonostante ciò, essa è stata sicuramente fatta in modo da migliorare l’accuratezza delle classificazioni fino all’ora utilizzate per ingenti quantità di dati.

Una ricerca invece che si concentra solo sulla classificazione degli oggetti di moda, e non su oggetti di categorie molto diverse, è stata fatta da Brian Lao e Karthik Jagadeesh, intitolata “Convolutional Neural Networks for Fashion Classification and Object Detection”. Questo studio si divide in tre fasi, il primo step si concentra sulla classificazione del tipo di abbigliamento che punta a fare una classificazione sulla tipologia di items partendo da un dataset di immagini. La seconda parte invece ha come obiettivo di individuare e attribuire gli attributi come il colore e il materiale. Mentre l’ultima parte si prefigge l’obiettivo di recuperare items simili a quelli preferiti e consigliarli agli utenti col fine di migliorare la loro esperienza di acquisto. I limiti, anche se minori rispetto a quelli della ricerca esposta precedentemente, sono il fatto che si concentra su degli oggetti molto eterogenei di moda, partendo da dei vestiti fino ad arrivare a degli accessori, perdendo così l’accuratezza nella classificazione di items simili tra loro.

Per ovviare a queste limitazioni dovute ad un’ampia gamma di prodotti tra loro molto eterogenei, sono state condotte delle ricerche scientifiche che mirano ad attuare una classificazione, sempre

utilizzando come supplemento modelli di machine learning o deep learning, per delle precise categorie di oggetti nel mondo della moda. Ad esempio, lo studio condotto da Dieinison J. F. Braga et Al, intitolato “Handbags Classification Model via Deep Learning”, si concentra solo sulle borse. Gli autori hanno come obiettivo quello di sviluppare un modello, con l’aiuto di algoritmi di deep learning, capace di indentificare determinati brand e modelli di borse.

Sono partiti da un dataset di 4000 immagini, professionali e non, di circa 200 modelli di 40 note case di moda. Prima di tutto però le immagini sono state ritagliate in modo da eliminare degli elementi, come lo sfondo, che avrebbero potuto compromettere i risultati o comunque peggiorarli. Per raggiungere il loro obiettivo hanno utilizzato l’architettura cosiddetta “Inception”⁵² al fine di addestrare un modello in grado di classificare gli attributi dei diversi modelli di borsa.

I limiti di questo studio però stanno nel fatto che la classificazione parte da un dataset di sole foto senza alcuna descrizione delle caratteristiche che magari dalle semplici foto non sono visibili, come il materiale, le dimensioni o dei piccoli dettagli.

⁵² È una struttura di rete neurale profonda, inizialmente sviluppata da ricercatori che lavoravano per Google, che migliora il riconoscimento delle immagini. Essa ha un impatto molto importante sulle ricerche scientifiche perché consente di catturare meglio i dettagli e quindi di migliorare i risultati.

2.3 GAP DI RICERCA

Come visionato precedentemente, sono stati fatti molti studi con l'intento di classificare gli oggetti contenuti in un sito e-commerce o in generale dei capi di abbigliamento o accessori di moda, ma la maggior parte delle ricerche non si concentra su una determinata categoria ma si focalizza nel classificare tutti gli items presenti sulla piattaforma, non riuscendo così ad attuare una classificazione di oggetti molto simili tra loro. Questa metodologia presenta quindi dei forti limiti perché, a causa dell'eterogeneità dei dati con il quale i vari modelli di machine learning e di deep learning vengono addestrati, l'accuratezza della classificazione è meno precisa e dettagliata. All'interno di una classificazione di una vasta gamma di oggetti, se si va ad analizzare ad esempio solo la categoria delle borse queste ricerche riescono solo a distinguerle in modo molto generalizzato. Elementi come il colore, il materiale, la composizione, la forma e le dimensioni vengono meno in questi approcci molto diffusi dell'intelligenza artificiale.

Le ricerche invece che si concentrano su una determinata tipologia di prodotti, utilizzano principalmente dei dataset di immagini. Questa caratteristica spesso è limitante soprattutto per la difficoltà di analisi quando si tratta di foto prese da internet o dai social network che non sono state appunto concepite o realizzate a fini scientifici o di studio. Infatti, l'illuminazione, la qualità della fotocamera, della foto e l'inquadratura, influenzano notevolmente l'aspetto finale di una immagine. Le foto catturate non in condizioni ottimali possono nascondere dei dettagli o comunque ingannare, creando false impressioni.

Inoltre, le sole foto non riescono a descrivere tutte le sfaccettature e le caratteristiche più tecniche dei prodotti. Soprattutto per quanto riguarda gli accessori di alta moda, la sola immagine non riesce a cogliere tutti i dettagli come le texture utilizzate, il materiale e la composizione.

Perciò, questo studio mira a utilizzare modelli di machine learning di classificazione di items di borsa partendo da tutti i dati tecnici della borsa stessa. Più in particolare verranno utilizzati come strumenti per una corretta classificazione il materiale, le dimensioni, la composizione ed eventuali particolari. Tutto ciò con lo scopo di poter successivamente raggruppare le varie categorie di borse in dei gruppi distinti basati sull'occasione d'uso. Al fine di poter proporre agli utenti che navigheranno sul sito e-commerce, delle iniziative di marketing personalizzate per ogni tipologia considerata e perciò di migliorare l'esperienza d'acquisto e di facilitare la ricerca dell'oggetto desiderato in base all'occasione per la quale la si sta cercando.

Si è scelto di classificare e quindi suddividere le varie tipologie di borse per occasione d'uso, per semplificare ancor di più l'esperienza di acquisto del consumatore. Ad esempio, se un utente sta

cercando una borsa per un evento molto importante e formale, avrà una gamma di prodotti proposti molto mirata, permettendogli di concentrarsi su degli oggetti pensati proprio per quelle occasioni.

Ciò, oltre ad avere un impatto sul tempo di acquisto, aiuterà il consumatore finale a spendere il proprio denaro in modo più saggio. Essi non saranno attratti dal fare acquisti impulsivi delle borse non utili per quella determinata occasione e ciò porta ad evitare di provare quella sensazione di sconforto per le borse acquistate in modo impulsivo che molto probabilmente rimarranno stipate in un angolo della casa. Questa caratteristica farà sentire il consumatore come furbo ed intelligente a non essersi fatto invogliare dalle centinaia di offerte di borse, e ciò farà ricollegare il sito e-commerce a questa sensazione molto positiva, portando così il consumatore a riacquistare nella piattaforma.

Questa metodologia di classificazione non è solo utile per il consumatore finale, ma anche per il venditore stesso. Tale segmentazione permette alle aziende di fare delle strategie e campagne di marketing indirizzate sfruttando molte sinergie con altre compagnie per poter creare delle collaborazioni vincenti e di successo. Ad esempio, per borse sportive, l'azienda potrebbe creare una capsule collection per un determinato evento sportivo, o per quanto riguarda le borse da lavoro potrebbero fare una partnership con delle aziende di oggettistica e cartoleria.

Oltre a collaborazioni, possono essere implementate delle campagne di marketing sfruttando le stagionalità o degli eventi ricorrenti, ad esempio potrebbero essere avviate delle campagne marketing per borse da lavoro nel periodo del rientro dalle vacanze estive o invernali.

Per rendere ulteriormente accattivante il layout della pagina dell'e-commerce, grazie alla classificazione delle borse per occasione d'uso, l'azienda può creare dei moodboard⁵³ inerenti con la determinata categoria. Attraverso colori, immagini e creando una sorta di storytelling la ricerca della borsa desiderata risulterà piacevole e coinvolgente, spronando il consumatore a preferire il sito rispetto agli altri.

Perciò, una corretta classificazione delle borse per occasione d'uso unita a delle strategie di marketing indirizzate, aumenterebbe la soddisfazione del consumatore e il successo della piattaforma e-commerce. Portando così l'utente a preferire il sito, e al venditore di poter emergere e affermarsi in un mercato molto competitivo come quello della moda.

⁵³ È una serie di immagini, colori, testi e materiali che rappresentano un'idea o un concetto. Utilizzato spesso in design, moda e pubblicità, serve a comunicare un determinato messaggio e a rendere più coinvolgente l'esperienza.

CAPITOLO 3: MATERIALI E METODI

3.1 DATASET

Per procedere con l'intento della tesi di ricerca, cioè quello di trovare il miglior metodo di classificazione attraverso il machine learning, utilizzando le caratteristiche delle borse per poi suddividerle in cluster, si è utilizzato un dataset fornito da un sito e-commerce.

Questo dataset raccoglie tutte le borse presenti nei vari siti della stessa azienda in varie parti del mondo, infatti, esso si compone di dati provenienti dal sito e-commerce dell'Italia, del Giappone, della Corea e della Cina.

Più precisamente il dataset si compone di 31.014 borse, divise in: 19.332 dal sito e-commerce in Italia, 7.659 dal sito in Cina, 2.037 da quello in Giappone ed infine 1.986 dalla piattaforma in Corea.

Oltre a questa divisione per paese, nel dataset erano presenti ed indicate le borse che sono nella sezione dedicata all'uomo o alla donna. In generale il dataset si compone di 8142 borse inserite nella categoria uomo e 22.872 in quella donna.

Il dataset ci fornisce varie informazioni sulle caratteristiche e sulle generalità di ogni tipo di borse, ci fornisce un link per l'immagine della borsa stessa, il nome colore dominante espresso anche in rgb⁵⁴, l'url che riconduce direttamente alla pagina web dove viene venduto il prodotto, l'id che è un codice univoco per identificare una determinata borsa, l'id del brand dell'accessorio, il nome del brand, il nome del prodotto, la tipologia di borsa, la valuta, il prezzo, il cluster di prezzo, informazioni sui materiali e informazioni sulle geometrie della borsa.

Prima di tutto per svolgere una buona analisi dei dati si è fatto un data cleaning efficace, perciò sono state decise le variabili che sono state successivamente considerate per lo studio. La variabile Y, cioè la variabile dipendente, è costituita dalla tipologia di borsa. Essa è composta da 15 tipologie, che sono indicate nella tabella sottostante.

Tabella 1: Tipologie di borse presenti nel dataset

⁵⁴ L'RGB (Red, Green, Blue) è un modello di colore utilizzato per rappresentare e visualizzare i colori nei display elettronici. Ogni colore RGB è definito da tre valori che vanno da 0 a 255, rappresentando l'intensità del rosso, del verde e del blu.

TIPOLOGIE DI BORSE	
Backpacks	Laptop & Briefcase
Bag accessories	Luggage and bags
Beach Bag	Messenger Bags
Belt Bag	Mini Bag
Bucket bag	Shoulder bags
Clutch bag	Tote Bag
Crossbody Bags	Travel Accessories
Wallets & Cardholders	

Lo scopo è quindi quello di addestrare un algoritmo di machine learning per classificare, partendo dalle caratteristiche che di seguito verranno spiegate, le tipologie di borse.

Per quanto riguarda le variabili che verranno considerate come indipendenti nel modello della ricerca, sono state utilizzate molte caratteristiche della borsa in modo da rendere lo studio completo.

È stato utilizzato il nome del colore, per ovviare alla moltitudine di sfumature e di tipologie di colore, sono stati raggruppati i vari colori in determinati cluster ideati utilizzando i colori ROYGBIV⁵⁵ o colori conosciuti e nell'immaginario di appartenenza comune.

Tabella 2: Cluster di colori utilizzati per la ricerca

CLUSTER DI COLORI	
Black	Purples
Browns	Reds
Cyans	Whites
Dark Blue	Yellow and Gold
Grays	Pinks
Greens	

Un'altra caratteristica che è stata utilizzata ai fini dell'addestramento del modello è stata il cluster di prezzo. A differenza del lavoro di suddivisione che si è fatto per i vari colori, per il prezzo, il dataset aveva già al suo interno l'informazione inerente al cluster di prezzo. Esso si compone di una scala a 3 valori: 0,1,2.

Nel cluster contrassegnato dal numero "0" sono presenti tutte quelle borse che hanno un prezzo che varia da 0 a 1000€, nel cluster "1" sono inserite gli accessori che hanno un valore tra i 1000 e i 1600€,

⁵⁵ Acronimo inglese utilizzato per descrivere i sette colori che rappresentano una suddivisione semplificata del continuo spettro della luce visibile, cioè il rosso, l'arancione, il giallo, il verde, il blu, l'indaco e il violetto

ed infine nell'ultimo cluster, indicato con il numero "2", sono presenti tutte le borse che hanno un prezzo maggiore di 1600€.

Oltre alle caratteristiche di colore e di prezzo sono state anche considerate le proprietà dimensionali e di composizione delle borse stesse.

In primo luogo, il dataset forniva varie informazioni sulla composizione della borsa e della percentuale di materiali utilizzati per la creazione di essa. Per riuscire a rendere questa caratteristica importante e soprattutto efficace per allenare il nostro modello, si è deciso di dividere in cluster anche questo attributo accorpando insieme tutte quelle borse che erano composte per la maggior parte dallo stesso materiale.

Perciò si è analizzato la percentuale di materiale utilizzato in ogni borsa e si è tenuto solo quello che era maggiormente presente.

Tabella 3: Cluster dei materiali utilizzati per la ricerca

CLUSTER DEI MATERIALI			
Calfskin > 50%	Lambskin >50%	Plastic 100%	Shearling 100%
Canvas >50%	Leather >50%	Plyurethane > 50%	Straw > 50%
Cotton > 50%	Linen > 50%	Polyamide >50%	Weaving
Cowhide 100%	Mix	Polyester > 50%	
Fabric > 50%	Nylon > 50%	PVC 100%	

Per quanto riguarda invece le caratteristiche dimensionali della borsa, sono state divise in 4 attributi: profondità, altezza, larghezza, tracolla.

Anche in questo caso per rendere l'analisi più efficace e dato che ogni borsa aveva delle misure diverse dalle altre, si è deciso di fare dei cluster per ogni attributo, perciò sono stati creati dei gruppi di misure ad intervalli regolari o di simile grandezza per tutte e quattro le caratteristiche.

Tabella 4: I cluster per gli attributi dimensionali delle borse

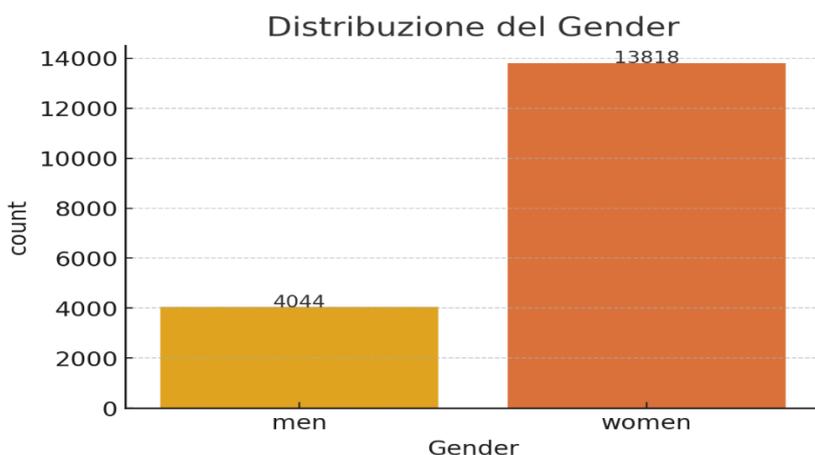
CLUSTER ALTEZZA		CLUSTER PROFONDITA'		CLUSTER LARGHEZZA	
6 -- 10	26 -- 29	1 -- 5	15 -- 19	0 -- 5	31 -- 35
10 -- 14	30 -- 34	6 -- 9	20 -- 25	10 -- 13	36 -- 40
15 -- 18	35 -- 39	10 -- 14	> 25	14 -- 18	41 -- 45
19 -- 22	40 -- 44			19 -- 24	> 45
23 -- 25	45 -- 49			25 -- 30	

CLUSTER TRACOLLA	
10 -- 20	51 -- 55
20 -- 30	56 -- 60
30 -- 36	61 -- 65
37 -- 43	66 -- 70
44 -- 50	> 70

Dopo aver sistemato e reso le variabili idonee ad essere utilizzate ai fini della ricerca, si è svolto un data cleaning. Molte borse non presentavano tutte le caratteristiche di materiale e dimensionali considerate, perciò, sono state eliminate tutte quelle osservazioni che non registravano un valore per tutti gli attributi.

Dopo un attento lavoro di pulizia e sistemazione del dataset, il risultato finale, che è stato utilizzato per l'addestramento e per il test dei vari algoritmi di machine learning, conta 17.862 borse in questo modo distribuite.

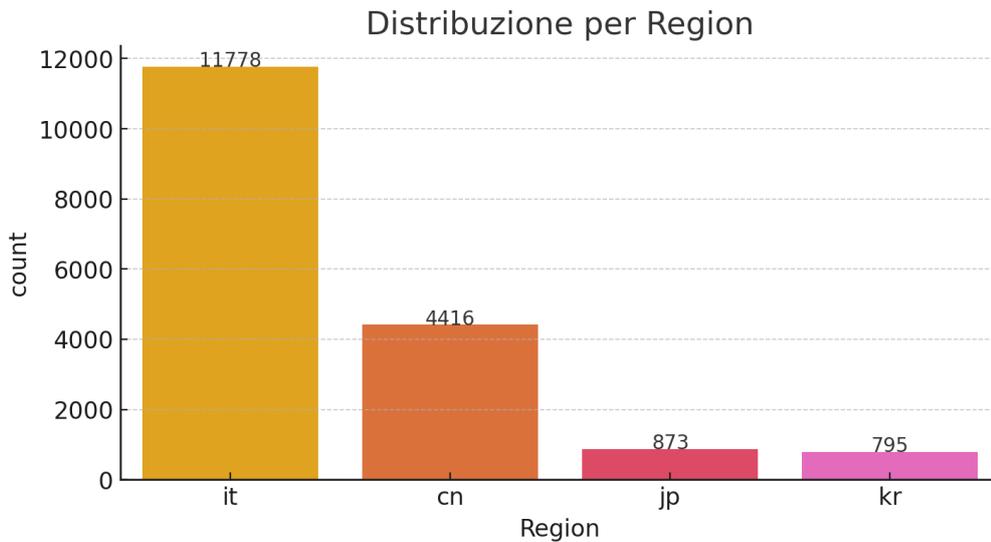
Grafico 8: Distribuzione per genere delle borse



Il grafico ci indica che 4.044 borse presenti sul dataset sono state inserite nella categoria uomo, mentre 13.818 in quella delle donne. Questa composizione non sorprende dato che, nonostante, come evidenziato nel primo capitolo, la moda al giorno d'oggi stia abbandonando i dogmi e le barriere di genere, le borse sono ancora un accessorio che interessa perlopiù il genere femminile piuttosto che quello maschile. Il dato comunque mostra come le borse, un giorno pensate solo per donne, siano comunque di interesse per l'uomo.

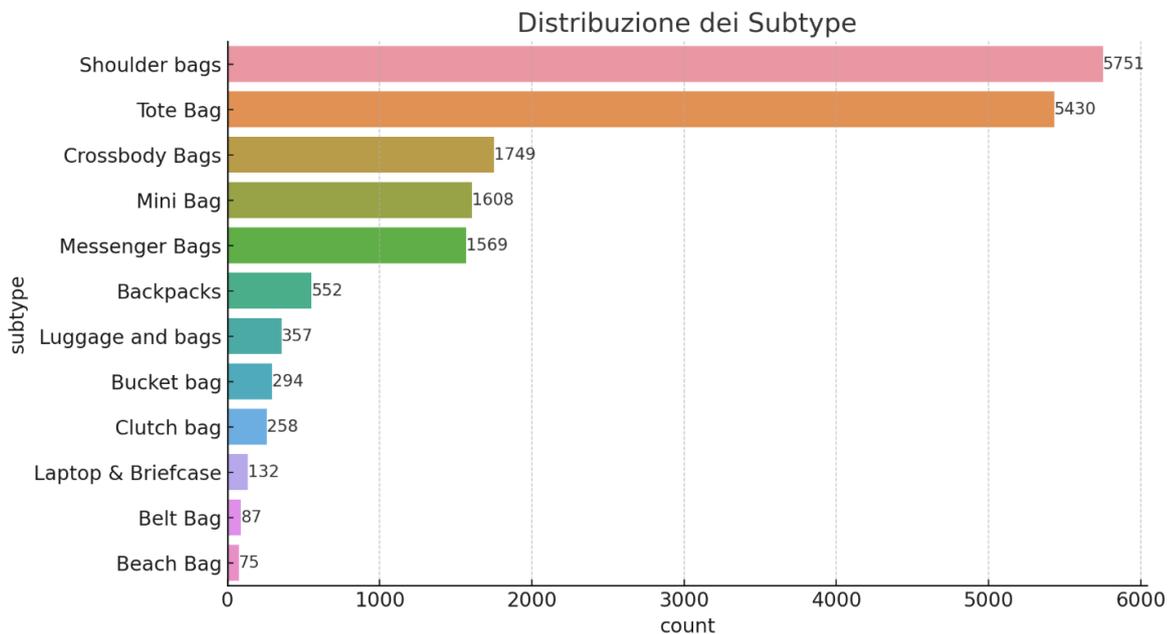
Per quanto riguarda la distribuzione dei dati osservati per regione, c'è una netta prevalenza di osservazioni prelevate dal sito e-commerce in Italia, mentre quelli presenti sul sito in Cina, Giappone e Corea sono molti meno come si evince dal grafico di seguito.

Grafico 9: Distribuzione per regione delle osservazioni



Osservando invece la distribuzione della variabile Y si nota come il dataset non sia perfettamente bilanciato, come si nota di seguito.

Grafico 10: Distribuzione delle tipologie di borse nel dataset



Come si nota dal grafico c'è un grosso sbilanciamento del dataset sulle "Shoulder bags" e le "Tote Bag" che sono molto presenti come osservazioni nel dataset. Il che, da una parte, ci allarma che questo fenomeno può influire le analisi che verranno fatte in seguito, perché potrebbero essere più precise per queste due tipologie data la quantità molto elevata di dati su cui poter allenare il modello. D'altro canto però, ci comunica anche dell'altissima domanda di mercato per queste due tipologie di borse e quindi alla corrispettiva maggiore presenza sul sito.

Il numero molto esiguo di osservazioni per alcune tipologie di borse come la “Beach Bag” e la “Belt Bag” potrebbero influenzare negativamente la riuscita della classificazione per queste due tipologie di accessori, ma, nonostante ciò, è importante sottolineare come questo genere di borse siano molto di nicchia e che quindi non abbiano una presenza ponderante all’interno di un sito e-commerce.

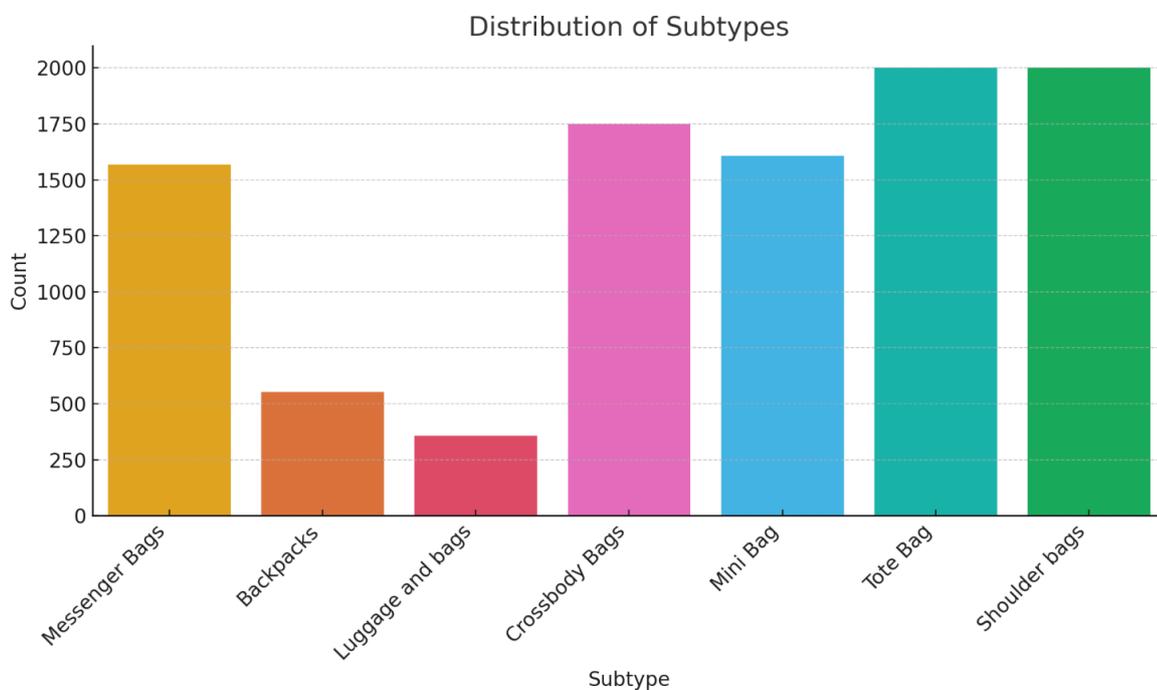
Per evitare che lo sbilanciamento del dataset abbia un impatto negativo sull’analisi, si è deciso di procedere con un metodo per bilanciare il dataset, il *downsampling*⁵⁶.

Per ottenere un set di dati bilanciato, si sono ridotte le due categorie con più osservazioni, “Shoulder bags” e “Tote Bag”, ed inoltre sono state escluse dall’analisi quelle tipologie di borse che hanno troppe poche misure.

Si è deciso quindi, di escludere dall’analisi le seguenti categorie di borse: “Beach Bag”, “Belt Bag”, “Laptop and Briefcase”, “Clutch bag”, “Bucket bag”. Esse non raggiungevano le 300 osservazioni e quindi potevano inficiare negativamente sul risultato finale della classificazione.

Per quanto riguarda le due categorie con più osservazioni sono state ridotte a 2000 misurazioni ciascuna, così da rendere bilanciato il set di dati. Questa correzione dello sbilanciamento dei dati è osservabile dal seguente grafico.

Grafico 11: Distribuzione di frequenza dopo il metodo di downsampling

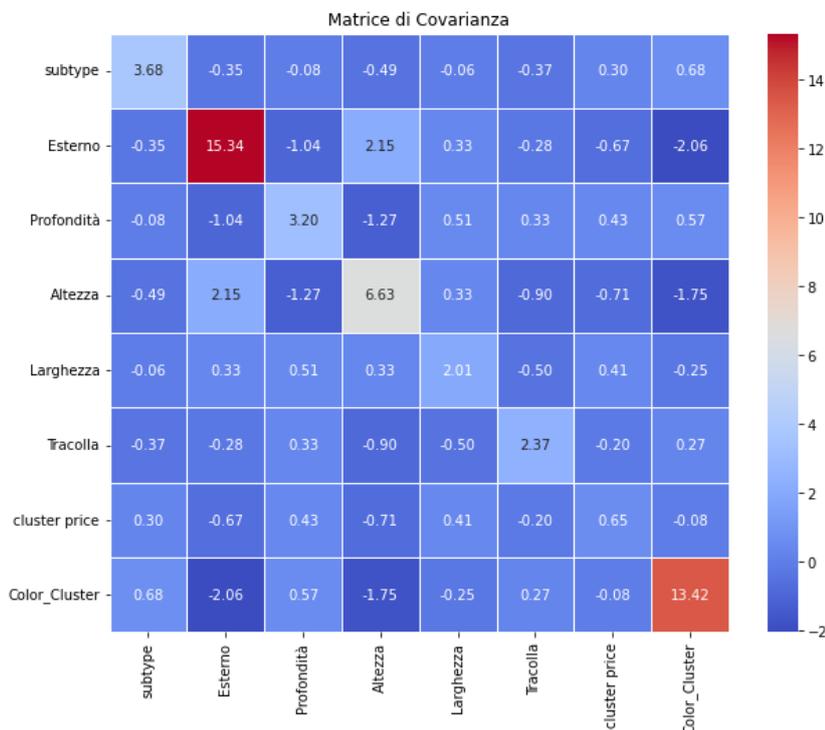


⁵⁶ Il *downsampling* è una tecnica di riduzione del numero di osservazioni in un dataset, utilizzata per semplificare l’analisi o per migliorare l’efficienza computazionale

Considerando tutte le variabili utilizzate per lo studio, sono state analizzate le covarianze⁵⁷ e le correlazioni⁵⁸ tra esse. Per quanto riguarda le covarianze, le variabili tra loro non presentano degli alti valori di covarianza, il che indica che non c'è una forte relazione lineare tra le variabili.

I risultati della covarianza per le variabili di interesse vengono descritti e riassunti dalla tabella di seguito.

Tabella 5: Matrice di covarianza tra le variabili



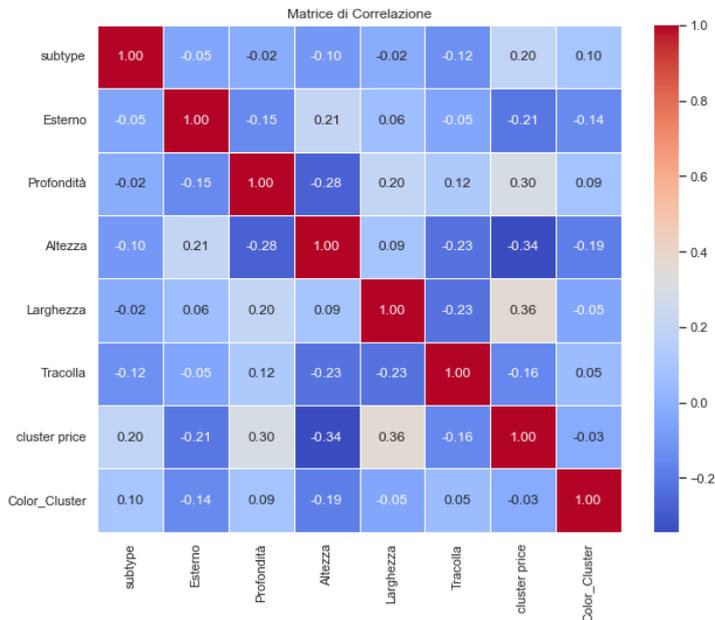
Osservando la tabella sovrastante si notano due valori molto alti, entrambi presenti nella diagonale della tabella stessa. Questi dati spiegano come ci sia un'alta varianza per le variabili "Esterno" e "Color_Cluster", il che dimostra che i due attributi hanno una grande variabilità.

⁵⁷ La covarianza è una misura statistica che misura la direzione, se negativa o positiva, della relazione lineare tra le due variabili. In generale quando la covarianza ha segno positivo, significa che le due variabili tendono a variare nella medesima direzione. Mentre, se la covarianza è negativa, simboleggia che i due attributi tendono a variare in direzione opposta. La covarianza di una variabile con sé stessa è la varianza dell'attributo stesso.

⁵⁸ La correlazione tra variabili è una misura statistica che descrive l'intensità e la direzione della relazione lineare tra due variabili. A differenza della covarianza, poco prima descritta, la correlazione è una misura standardizzata che varia sempre tra -1 e 1, rendendola più facile da interpretare e confrontare.

Per quanto riguarda la correlazione tra le variabili, come illustrato dalla Tabella 6, in valore assoluto non sono di forte intensità. Questa particolarità indica una che non c'è una forte relazione, positiva o negativa, tra le variabili.

Tabella 6: Matrice di correlazione tra le variabili



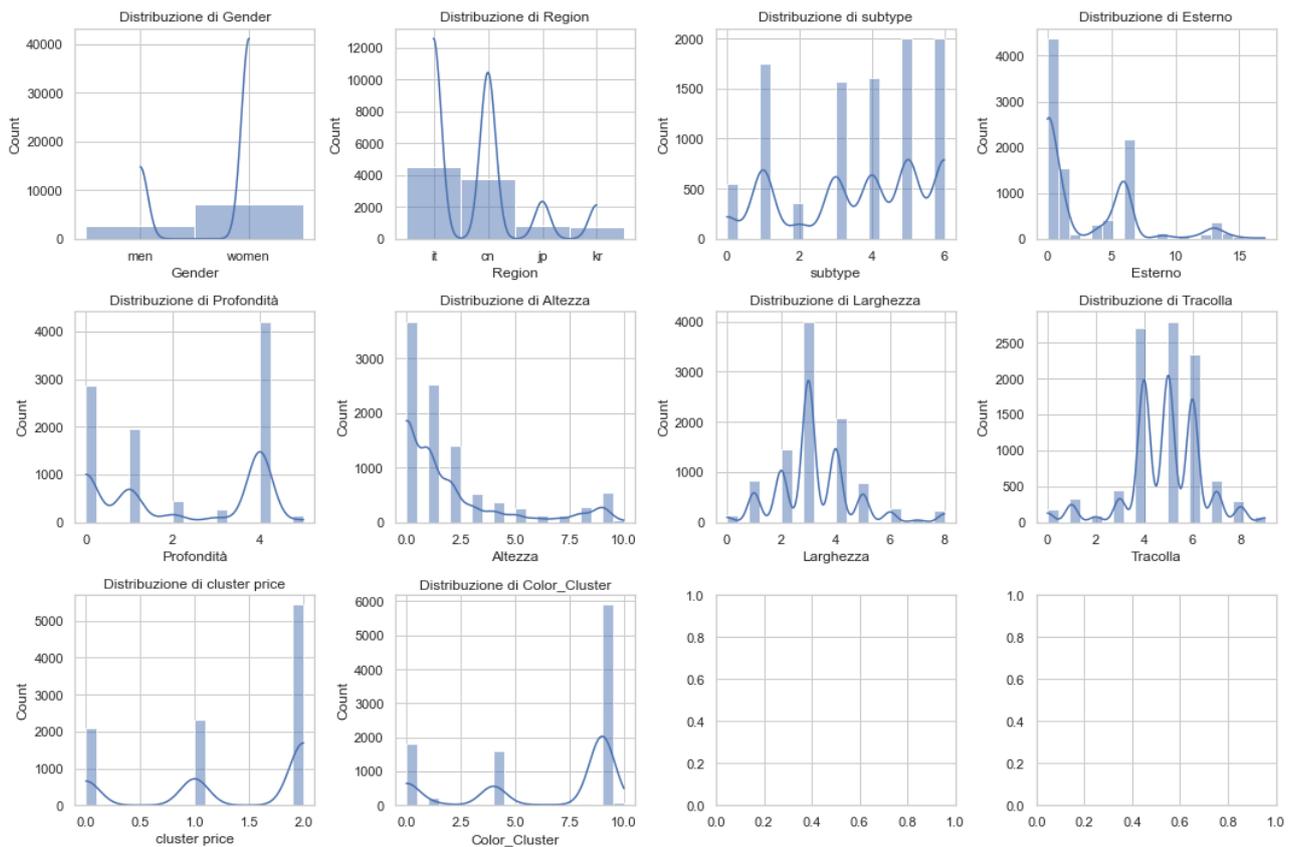
Alcune variabili presentano correlazioni positive con bassa intensità, come ad esempio “Larghezza” e “cluster price”, suggerendo l’esistenza di relazioni tra gli attributi anche se molto deboli. Anche per quanto riguarda correlazioni negative ci sono degli esempi di legami che hanno direzione opposta, anche se molto leggeri, come ad esempio tra “Profondità” e “Altezza”.

Dopo aver analizzato le covarianze e le correlazioni tra variabili, sono state anche calcolate le distribuzioni per tutti gli attributi. Questi dati sono riassunti nella tabella sottostante che, utilizzando istogrammi⁵⁹ con curve di densità⁶⁰, riesce a descrivere al meglio le misure statistiche delle varie variabili.

Grafico 12: Distribuzione di densità delle variabili

⁵⁹ Gli istogrammi sono un tipo di grafico a barre utilizzati per visualizzare la distribuzione di una variabile continua. Le barre mostrano il conteggio di dati in ciascun intervallo di valori.

⁶⁰ Le curve sovrapposte sugli istogrammi rappresentano la densità stimata della distribuzione della variabile, offrendo una visualizzazione più liscia e continua della distribuzione rispetto alle barre dell’istogramma. Queste curve sono spesso chiamate KDE (Kernel Density Estimate).



È interessante osservare, attraverso le curve di densità, la presenza di distribuzioni multimodali⁶¹, che suggeriscono che in una variabile ci siano più picchi. Ad esempio, nelle variabili “Profondità”, “Altezza” e “Larghezza” sono ben visibili picchi distinti, il che potrebbe significare che ci siano misure o taglie comuni nei prodotti.

Oltre ad aver analizzato la densità di frequenza delle variabili è stato oggetto di interesse il boxplot⁶². Questo studio è stato fatto per individuare l’eventuale presenza di outlier⁶³ che potrebbero influenzare negativamente le analisi dei dati.

Come si nota dal grafico sottostante c’è una minima presenza di outlier nel dataset che non compromettono i risultati del nostro studio. Il fatto che ci siano pochi di essi era prevedibile dato che

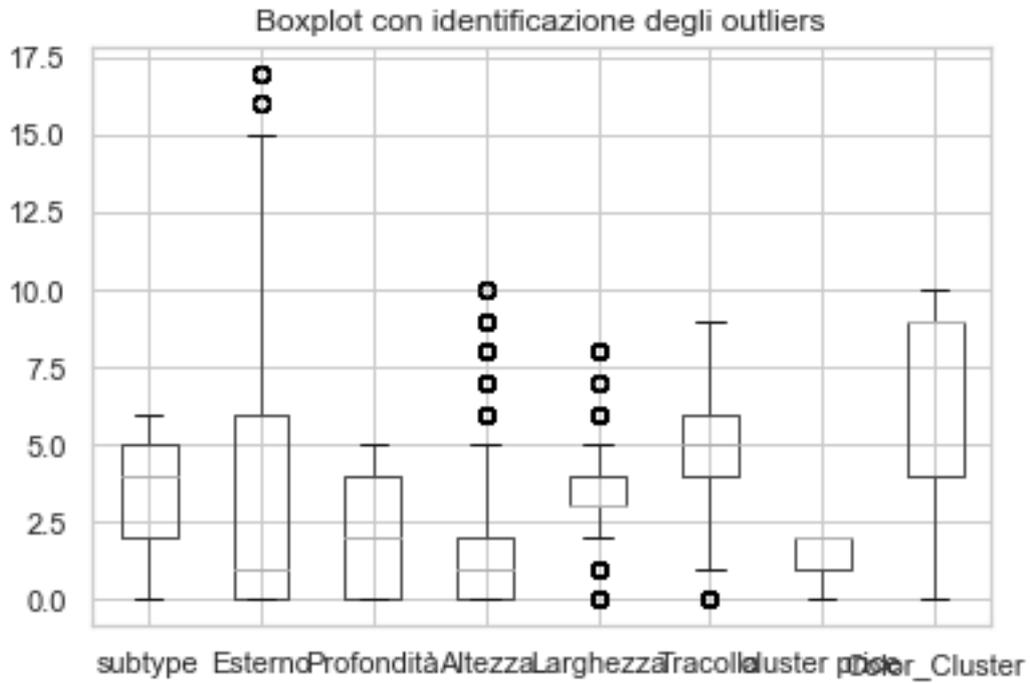
⁶¹ Le distribuzioni multimodali sono distribuzioni di dati caratterizzate dalla presenza di più di un picco, ossia più punti di massima frequenza. A differenza delle distribuzioni unimodali, che hanno un solo picco (come la normale distribuzione a campana), le distribuzioni multimodali indicano che ci sono diverse sottopopolazioni o gruppi nei dati, ciascuno con il proprio centro di frequenza.

⁶² Il boxplot, che è anche conosciuto come diagramma a scatola e baffi, è utilizzato per visualizzare la distribuzione dei dati e se c’è la presenza di outlier. Il boxplot evidenzia la gamma, il mediano, il primo quartile, il terzo quartile, e gli outlier.

⁶³ Gli outlier sono dati che si discostano in modo significativo dal resto delle osservazioni in un dataset. Possono indicare anomalie o eventi rari e, in analisi statistica, influenzano la media e altre misure riassuntive. In generale, gli outlier possono derivare da errori di misurazione o rappresentare informazioni preziose.

è stato fatto precedentemente un importante e accurato lavoro di clustering per quasi tutte le variabili, che hanno così ridotto drasticamente il rischio di presenza degli outlier.

Grafico 13: Boxplot con evidenziati eventuali outlier



3.2 TECNOLOGIE UTILIZZATE

Per trovare l'algoritmo di classificazione che performa meglio per questa tipologia di ricerca, è stato utilizzato come supporto il linguaggio di programmazione chiamato "Python". Esso, nonostante sia stato creato e rilasciato nel 1991⁶⁴, continua ad essere uno dei linguaggi più popolari al mondo.

Python deve la sua notorietà e il suo continuo utilizzo alla sua semplicità e leggibilità, cioè il suo linguaggio è molto chiaro e di facile interpretazione. Questa sua caratteristica lo rende ideale anche per chi non ha molta dimestichezza con la programmazione.

Un'altra caratteristica di Python è che è un linguaggio interpretato, cioè il codice è eseguito riga per riga e ciò permette di testare rapidamente il codice stesso, bisogna aggiungere però che questa caratteristica può influire in modo negativo le prestazioni rispetto a quelle dei linguaggi compilati⁶⁵.

In generale una dei motivi che rende Python uno dei sistemi più noti e utilizzati per la programmazione, da parte di neofiti e di esperti, è quella che ha un vasto ecosistema di librerie e di pacchetti che coprono un'ampia gamma di domini, tra i quali intelligenza artificiale, data science, automazione e molti altri. Nel nostro caso esso è stato utilizzato per allenare vari modelli di machine learning e quindi sono state adottate varie librerie di intelligenza artificiale per riuscire ad ottenere i risultati sull'accuratezza del modello.

Nonostante il linguaggio di programmazione Python sia molto performante, esso ha alcune limitazioni in alcuni suoi aspetti. Ad esempio, come anticipato precedentemente, essendo un linguaggio interpretato, tende ad essere più lento e macchinoso rispetto a linguaggi compilati. Un altro problema da considerare è quello della gestione della memoria, perché rispetto ad altri linguaggi, consuma molta più memoria. Questa caratteristica può essere una limitazione per alcuni studi che analizzano o utilizzano una mole di dati molto elevata.

Infine, nonostante abbia una grande quantità di librerie utilizzabili per svariati fini diversi, Python non è sempre il linguaggio migliore per ogni situazione. Infatti, ci sono dei linguaggi più adatti per

⁶⁴ Wikipedia Contributors. (2021). Python (programming language). Retrieved from Wikipedia website: https://en.wikipedia.org/wiki/Python_%28programming_language%29

⁶⁵ I linguaggi compilati sono dei linguaggi di programmazione il cui codice sorgente viene tradotto in codice macchina eseguibile da un compilatore. Questo processo avviene prima dell'esecuzione, migliorando le prestazioni e l'efficienza del programma stesso. Gli esempi più famosi di questa tipologia sono C, C++, e Java.

un controllo di massima precisione delle risorse hardware⁶⁶. Per queste tipologie di lavoro sono più indicati linguaggi come “C”, “C++” e “Rust”.

Per precisione, nella ricerca è stato utilizzato “Spyder” che è un IDE⁶⁷ potente e versatile, specificamente progettato per il calcolo scientifico e l'analisi dei dati in Python. Sebbene abbia alcuni limiti in termini di prestazioni su progetti molto grandi e flessibilità per applicazioni diverse dalla scienza dei dati, rimane uno strumento molto utilizzato e molto utile per chiunque lavori con Python nel contesto della ricerca scientifica e dell'analisi dei dati.

⁶⁶ Le risorse hardware sono le componenti fisiche di un dispositivo elettronico che consentono il funzionamento del sistema. Includono elementi come CPU, RAM, disco rigido, scheda grafica, schede di rete, e periferiche come tastiere e monitor.

⁶⁷ Un IDE, acronimo di Integrated Development Environment, è un software che fornisce agli sviluppatori un set completo di strumenti per scrivere, testare, debug e mantenere il codice software.

3.2.1 LIBRERIE DI PYTHON UTILIZZATE

Come anticipato nel paragrafo precedente, è stato utilizzato Python come linguaggio di programmazione anche per la possibilità di utilizzare una vasta gamma di librerie che possono essere adottate per la ricerca. Più in particolare, sono state utilizzate maggiormente le seguenti librerie che verranno descritte e verranno analizzate i loro contributi allo studio.

PANDAS: Pandas, abbreviazione di "Panel Data System", è una libreria essenziale per la gestione e la manipolazione dei dati in Python, sviluppata sulla base di NumPy. NumPy fornisce array multidimensionali e funzioni matematiche avanzate, che costituiscono il fondamento su cui Pandas sviluppa le sue strutture dati principali: le Series e i DataFrame.

Le Series sono una struttura dati unidimensionale con etichette per ciascun elemento, assimilabile ad una colonna in un foglio di calcolo o una tabella SQL. Ogni elemento in una Series ha un'etichetta di indice, che consente di accedere ai dati in modo intuitivo e diretto.

Il DataFrame è una struttura dati bidimensionale simile a una tabella, composta da una collezione di Series allineate su un indice comune. I DataFrame hanno etichette sia per le righe che per le colonne, facilitando l'accesso e la manipolazione dei dati. Questa struttura è particolarmente utile per gestire dati eterogenei, permettendo di immagazzinare diversi tipi di dati (numerici, stringhe, date, ecc.) in colonne diverse.

Pandas offre un'ampia gamma di funzionalità per il data wrangling⁶⁸, che includono l'importazione e l'esportazione di dati da e verso diversi formati come CSV, Excel, JSON, SQL e molti altri. Le operazioni di filtraggio, ordinamento, raggruppamento e aggregazione possono essere eseguite in modo efficiente e con poche righe di codice. Pandas semplifica anche la manipolazione di date e tempi, la gestione di dati mancanti, e l'operazione su dati con diverse frequenze temporali, il che è essenziale per l'analisi delle serie temporali.

Grazie alla sua integrazione con altre librerie scientifiche Python come Matplotlib e Seaborn, Pandas permette anche di creare visualizzazioni di dati direttamente dai DataFrame, facilitando l'analisi esplorativa dei dati. La stretta integrazione con NumPy consente di eseguire calcoli vettorializzati⁶⁹ e operazioni matematiche avanzate su grandi dataset in modo efficiente.

⁶⁸ Il data wrangling è il processo di trasformazione e pulizia dei dati grezzi in un formato più adatto per l'analisi.

⁶⁹ Le operazioni vettorializzate sono dei calcoli applicati simultaneamente ad intere array invece che a singoli elementi.

Pandas è altamente utilizzato nelle fasi di preelaborazione dei dati, dove la pulizia, la trasformazione e la preparazione dei dati per l'analisi successiva sono cruciali. Le sue funzioni intuitive e le API potenti permettono di gestire dataset complessi e di grande dimensione, rendendo il flusso di lavoro più efficiente e riducendo significativamente il tempo necessario per l'analisi dei dati.

Inoltre, Pandas supporta l'elaborazione di dati da fonti diverse come fogli Excel, database SQL, file di testo e persino dati ottenuti via API. Questa flessibilità rende Pandas uno strumento di scelta non solo per la scienza dei dati, ma anche per l'analisi finanziaria, la ricerca scientifica, l'analisi di mercato e molte altre applicazioni che richiedono l'elaborazione di dati strutturati⁷⁰.

NUMPY: abbreviazione di "Numerical Python", è una libreria fondamentale per l'elaborazione numerica in Python. Essa ha un ruolo centrale nella computazione scientifica in Python, fornendo un potente supporto per la manipolazione e l'elaborazione di dati numerici. Grazie alla sua compatibilità con altre librerie come Pandas e Matplotlib, NumPy è essenziale per operazioni di data science, per il calcolo scientifico, per l'elaborazione di immagini e per lo sviluppo di algoritmi.

NumPy offre strutture dati efficienti, come array multidimensionali e una vasta gamma di funzioni matematiche per operazioni complesse su grandi volumi di dati.

In conclusione, riassumendo, NumPy è essenziale per il trattamento di grandi volumi di dati numerici, perché riesce ad ottimizzare le operazioni computazionali e garantire prestazioni molto elevate anche su dataset di notevoli dimensioni⁷¹.

MATPLOTLIB: essa è una libreria altamente versatile per la generazione di grafici bidimensionali e tridimensionali. L'utilizzo di grafici e quindi la conseguente possibilità di visualizzare i dati, è un aspetto cruciale dell'analisi, poiché consente di comprendere meglio le caratteristiche e le relazioni presenti nei dataset, facilitando la comunicazione dei risultati.

Questa capacità di rappresentare visivamente le informazioni rende i dati immediatamente comprensibili e interpretabili, migliorando così la qualità delle decisioni basate sui dati. Inoltre, l'integrazione di Matplotlib con altre librerie Python, come Pandas e Seaborn, estende ulteriormente le sue capacità, permettendo analisi avanzate e visualizzazioni interattive che ampliano la comprensione dei risultati della ricerca⁷².

⁷⁰ McKinney, W. (2011). pandas: a foundational Python library for data analysis and statistics. Python for high performance and scientific computing, 14(9), 1-9.

⁷¹ Bressert, E. (2012). SciPy and NumPy: an overview for developers.

⁷² Ari, N., & Ustazhanov, M. (2014, September 1). Matplotlib in python. <https://doi.org/10.1109/ICECCO.2014.6997585>

Questa libreria è stata utilizzata sia per la rappresentazione del dataset, che è visibile nel sottoparagrafo precedente dove sono state fatte le analisi dei dati, ma anche per illustrare graficamente i risultati dei vari algoritmi utilizzati per la ricerca.

SCIKIT-LEARN: rappresenta lo strumento chiave per l'implementazione di algoritmi di machine learning e tecniche di data mining in Python. Questa libreria open-source⁷³, costruita sulla base di NumPy, SciPy e Matplotlib, include una vasta gamma di algoritmi per la classificazione, il clustering, la regressione, la riduzione della dimensionalità, e la selezione delle caratteristiche, tutti facilmente integrabili nel workflow analitico. Inoltre, Scikit-learn si integra perfettamente con altre librerie e strumenti di data science in Python, come Pandas per la gestione dei dati, NumPy per l'elaborazione numerica, e Matplotlib o Seaborn per la visualizzazione dei dati⁷⁴.

Per la ricerca sono state utilizzati i seguenti modelli presenti nella libreria di Scikit-learn:

- Gradient Boosting
- Naive Bayes
- Random Forest
- SVM

⁷³ Una libreria open source è una raccolta di codice disponibile gratuitamente, che gli sviluppatori possono utilizzare, modificare e distribuire. Una caratteristica che le contraddistingue è che essendo open source, il codice è aperto al contributo e alla revisione della comunità globale, migliorandone sicurezza e qualità.

⁷⁴ Hao, J., & Ho, T. K. (2019). Machine Learning Made Easy: A Review of Scikit-learn Package in Python Programming Language. *Journal of Educational and Behavioral Statistics*, 44(3), 348-361. <https://doi.org/10.3102/1076998619832248>

3.2.2 MODELLI UTILIZZATI

Come ribadito più volte nei precedenti capitoli e paragrafi, la ricerca si concentra sul trovare il miglior modello di machine learning per classificare le tipologie di borse (variabile chiamata “subtype” nel dataset) utilizzando delle variabili riguardanti le caratteristiche delle borse stesse, come la dimensione, il colore o il prezzo.

Per sancire se la classificazione, utilizzando queste variabili sia efficiente e accurata, sono stati addestrati vari modelli di machine learning per poi, se i risultati fossero stati positivi, decretare quale modello sia il più adatto per questo tipo di problema.

Il primo ad essere stato testato è stato il Random Forest. Questo modello combina la potenza di molti alberi decisionali per migliorare la precisione e ridurre l'overfitting, un problema comune con gli alberi decisionali singoli. Più nel dettaglio, il Random Forest costruisce una moltitudine di alberi decisionali durante il processo di addestramento. Ogni albero viene addestrato con campioni casuali del dataset originali⁷⁵, e una volta fatto ciò, il risultato finale è dato dalla maggioranza delle predizioni di tutti gli alberi di cui è composto⁷⁶.

Questo modello viene utilizzato per la sua elevata accuratezza e precisione anche con dataset molto complessi ed inoltre, attuando il “bagging”, riduce drasticamente il rischio di overfitting rispetto all'utilizzo di un singolo albero decisionale. Pur non essendoci presenti numerosi outlier nel dataset utilizzato per la ricerca, questo modello è comunque molto robusto contro gli outlier.

Inoltre, è stato scelto di utilizzare il Random Forest, perché esso consente di fornire una stima sull'importanza che ciascuna variabile ha avuto per la buona riuscita dell'addestramento. Ciò permette di identificare quali attributi sono stati più rilevanti e d'impatto per le predizioni⁷⁷.

Nonostante i tanti punti di forza del Random Forest, questo modello ha anche dei punti di debolezza, ad esempio il fatto che può richiedere l'utilizzo di molta memoria, soprattutto per analisi con dataset complessi, o anche il tempo richiesto per completare l'addestramento che, a causa della presenza di molti alberi, può essere molto prolungato. Il punto di debolezza più importante, osservando la nostra

⁷⁵ Questo metodo è chiamato “bagging” o “bootstrap aggregating”

⁷⁶ Parmar, A., Katariya, R., Patel, V. (2019). A Review on Random Forest: An Ensemble Classifier. In: Hemanth, J., Fernando, X., Lafata, P., Baig, Z. (eds) International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018. ICICI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 26. Springer, Cham. https://doi.org/10.1007/978-3-030-03146-6_86

⁷⁷ Editorial. (2022). Pros and cons of Random Forest Algorithm. Retrieved from RoboticsBiz website: <https://roboticsbiz.com/pros-and-cons-of-random-forest-algorithm/>

ricerca, è quello che questo modello può essere influenzato negativamente da dati sbilanciati, che è una caratteristica presente per alcune variabili del nostro studio.

Successivamente è stato utilizzato il Gradient Boosting. Questo modello si basa sulla creazione sequenziali di modelli, nel quale ciascun modello cerca di correggere gli errori commessi da quello precedente. Esso viene fatto attraverso il calcolo del residuo, cioè la differenza tra predizione e valore reale. Una volta calcolato il residuo il modello successivo tenterà di correggerlo. Questo procedimento viene ripetuto n volte fino a che non viene raggiunto un limite che è il risultato finale del modello⁷⁸.

Il Gradient Boosting è noto per la sua elevata accuratezza e precisione anche in casi molto complessi. Per quanto riguarda i punti di debolezza sono che, nonostante sia molto robusto, è suscettibile all'overfitting ed inoltre talvolta è di difficile interpretazione⁷⁹.

Successivamente è stato utilizzato il Naive Bayes. Esso è un algoritmo di classificazione multiclasse⁸⁰ che assume prima di tutto che tutte le variabili siano indipendenti una dall'altra. Questa caratteristica viene confermata dalla matrice di correlazione, precedentemente calcolata e illustrata, che sottolinea l'indipendenza tra le variabili.

Questo modello calcola a priori le probabilità di tutte le categorie (c) della variabile dipendente P(C), e successivamente calcola le probabilità condizionali per ogni variabile indipendente xi dato che una borsa sia un determinato "subtype" P(xi|C).

Infine, prende una serie di dati e calcola la probabilità per ogni classe della variabile dipendente usando il teorema di Bayes⁸¹ e si sceglie la classe con probabilità a posteriori più alta.

$$P(C|X) \propto P(C) \cdot \prod_{i=1}^n P(x_i|C)$$

$$\operatorname{argmax}_C P(C|X)$$

⁷⁸ Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*, 7, 21.

⁷⁹ GeeksForGeeks. (2020). ML - Gradient Boosting. Retrieved from GeeksforGeeks website: <https://www.geeksforgeeks.org/ml-gradient-boosting/>

⁸⁰ Un algoritmo di classificazione multiclasse può gestire problemi e analisi con più di due classi, quindi la variabile dipendente non deve essere per forza binaria ma può avere più categorie

⁸¹ Il teorema di Bayes è una formula fondamentale in statistica e calcolo della probabilità che descrive la probabilità condizionale di un evento, data l'evidenza di un altro evento.

Questo algoritmo viene molto utilizzato per la sua rapidità e leggerezza. Inoltre, è anche molto performante su dataset di grandi dimensioni.

Nonostante queste caratteristiche che lo rendono molto affidabile e utile, esso è molto sensibile ai dati non bilanciati, come nel caso della nostra variabile dipendente. Poiché si basa sulle probabilità condizionali, le classi più rare possono essere sottorappresentate, portando a modelli che hanno un bias verso le classi più frequenti⁸².

Poi è stato utilizzato l'XGBoosting che è un algoritmo di machine learning basato su alberi di decisione, progettato per essere altamente efficiente, flessibile e portabile. Esso utilizza la stessa metodologia del Gradient Boosting, ma a differenza di questo utilizza altre tecniche avanzate come l'apprendimento parallelo degli alberi, l'eliminazione delle features ridondanti e il pruning⁸³.

Le caratteristiche per il quale il seguente modello è molto utilizzato sono il fatto che è noto per essere uno degli algoritmi più veloci e performanti. In aggiunta, grazie alle tecniche avanzate che utilizza, riesce a prevenire l'overfitting. I punti negativi del suo utilizzo sono la sua complessità e che può richiedere l'utilizzo di molta memoria se viene utilizzato per dataset complessi e di grande dimensione⁸⁴.

Infine, è stato testato il modello SVM⁸⁵ che cerca di trovare un iperpiano ottimale che separa i dati nel modo più netto possibile. Cercando, quindi, di massimizzare il margine tra le classi di dati⁸⁶.

Esso è utilizzato sia per classificazioni lineari che non lineari, come nel caso della ricerca, ed è molto performante quando ci sono un numero di variabili molto alti rispetto al numero delle osservazioni.

Inoltre, l'SVM è poco sensibile agli outlier ma non performa in modo eccelso con la presenza di dataset molto grandi e complicati. Perciò, nonostante sia un modello molto utilizzato per studi di classificazione, nel caso specifico di questa ricerca potrebbe non performare al meglio a causa della grandezza del dataset.

⁸² IBM. (2023). What is Naïve Bayes | IBM. website: <https://www.ibm.com/topics/naive-bayes>

⁸³ Il pruning è una tecnica utilizzata negli alberi di decisione per ridurre la complessità del modello e prevenire l'overfitting. Consiste nel rimuovere rami dell'albero che hanno poca o nessuna importanza nella predizione finale. Ciò migliora la generalizzazione del modello, rendendolo più semplice e robusto su nuovi dati.

⁸⁴ Hachcham, A. (2023). XGBoost: Everything You Need to Know. Retrieved from Neptune.ai website: <https://neptune.ai/blog/xgboost-everything-you-need-to-know>

⁸⁵ Support Vector Machine

⁸⁶ Yue, S., Li, P., & Hao, P. (2003). SVM classification: Its contents and challenges. *Applied Mathematics-A Journal of Chinese Universities*, 18(3), 332–342. <https://doi.org/10.1007/s11766-003-0059-5>

CAPITOLO 4: RISULTATI E DISCUSSIONI

4.1 RISULTATI

Dopo aver lavorato sul dataset in modo da renderlo pronto per le analisi e dopo aver scelto accuratamente quali modelli testare per lo studio, si è proceduto testando i vari algoritmi di machine learning per capire quale tra quelli provati sia migliore per questo tipo di task.

Dopo aver importato il dataset e trasformato le variabili da categoriche a numeriche, utilizzando un Label Encoder⁸⁷, e dopo aver definito la variabile dipendente e le variabili indipendenti, si è iniziato a testare i vari modelli.

Il primo ad essere utilizzato è stato il Gradient Boosting, e per far sì che il modello possa generalizzare su dati mai visti prima, si sono divisi i dati nel seguente modo: l'80% è stato utilizzato per addestrare il modello, mentre il restante 20% è stato utilizzato per testare il modello e per vederne l'accuratezza e i risultati.

Per valutare il seguente modello e quelli successivi si è calcolata l'accuratezza, cioè la percentuale di previsioni corrette sui campioni testati, la matrice di confusione, che mostra attraverso una tabella come il modello ha classificato ciascun campione, suddividendo le previsioni in numero di campioni che appartengono ad una determinata classe e che il modello ha classificato in modo corretto. Essi sono tutti quei valori presenti sulla diagonale della matrice. Mentre gli elementi fuori dalla diagonale rappresentano le classificazioni errate. Cioè, per ogni riga, essi indicano quante volte una classe è stata erroneamente predetta come una differente.

Infine, per valutare i risultati, è stata utilizzato il "Classification report" che è una tabella nel quale per ogni categoria vengono indicati i risultati in modo dettagliato, soffermandosi sulla "precision", "recall", "f1-score".

La "precision" per ogni classe che compone la variabile dipendente viene calcolata come il rapporto tra i campioni di quella classe correttamente classificati, chiamati anche veri positivi, e la somma tra i veri positivi e i campioni di altre categorie ma erroneamente classificati come appartenenti a questa, detti anche falsi positivi.

La "recall", invece, viene definita anche come sensibilità del modello e viene calcolata come il rapporto tra i veri positivi e la somma tra i veri positivi e i campioni di quella classe erroneamente classificati come di altre, detti anche falsi negativi.

⁸⁷ Il Label Encoder è uno strumento utilizzato per trasformare variabili categoriche in valori numerici. Funziona assegnando a ciascuna categoria un numero intero, rendendo i dati comprensibili agli algoritmi di machine learning che richiedono input numerici.

Infine, "f1-score" è una metrica per valutare il modello che combina la "precision" e la "recall". Essa è la media armonica tra le due operazioni sopra descritte, cioè è una media che riesce a bilanciare i due valori qualora siano di grandezze molto differenti. Con la media armonica si riesce a penalizzare un valore molto basso di una delle due metriche per far sì che un'elevata "precision" o "recall" non sbilanci e comprometta l'analisi del modello⁸⁸.

Per quanto riguarda il Gradient Boosting, il modello ha registrato un'accuratezza dell'87%. Per verificare che non ci sia un problema di overfitting, si è confrontata l'accuratezza sia sul set di training sia sul set di test, ed è risultato che non c'è overfitting dato che i due risultati sono molto simili. Infatti, l'accuratezza sul set di training è del 87,3%, segno che il modello generalizza bene ai dati non utilizzati per l'addestramento.

Come mostrato nella tabella di seguito, le metriche di "precision", "recall" e "f1-score" sono molto positive, le classi che hanno delle prestazioni leggermente inferiori sono "Crossbody Bags" e "Shoulder Bags". Ma, ciò nonostante, il modello ha degli ottimi risultati.

Tabella 7: Classification report del Gradient Boosting

Classe	Precision	Recall	F1-Score	Support
Backpacks	1.00	0.99	1.00	124
Crossbody Bags	0.75	0.84	0.79	328
Luggage and bags	1.00	1.00	1.00	77
Messenger Bags	0.91	0.97	0.94	323
Mini Bag	0.89	0.96	0.93	327
Shoulder Bags	0.84	0.70	0.77	386
Tote Bag	0.88	0.84	0.86	402

Per quanto riguarda la matrice di confusione, essa ci mostra come il modello funzioni molto bene per alcune classi come "Luggage and bags" e "Backpacks" che hanno un numero elevato di previsioni corrette sulla diagonale. Mentre per altre classi come "Shoulder Bags" segnala dei problemi, perché proprio per questa categoria 64 campioni sono stati classificati erroneamente come "Crossbody Bags". Anche le "Tote Bag" vengono spesso classificate in modo errato in altre categorie.

⁸⁸ Goutte, C., Gaussier, E. (2005). A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. In: Losada, D.E., Fernández-Luna, J.M. (eds) Advances in Information Retrieval. ECIR 2005. Lecture Notes in Computer Science, vol 3408. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-31865-1_25

Tabella 8: Matrice di confusione del Gradient Boosting

Classe	Backpacks	Crossbody Bags	Luggage and bags	Messenger Bags	Mini Bag	Shoulder Bags	Tote Bag
Backpacks	123	0	0	0	0	0	1
Crossbody Bags	0	276	0	10	4	13	25
Luggage and bags	0	0	77	0	0	0	0
Messenger Bags	0	0	0	313	2	8	0
Mini Bag	0	5	0	2	314	6	0
Shoulder Bags	0	64	0	6	26	272	18
Tote Bag	0	22	0	12	5	26	337

Infine, per concludere l'analisi dei risultati del modello, si è calcolata quale tra le variabili ha maggiormente condizionato la buona riuscita della classificazione.

Tabella 9: Importanza delle features del Gradient Boosting

Caratteristica	Importanza
Altezza	0.272904
Profondità	0.216709
Larghezza	0.139465
Tracolla	0.131274
Esterno	0.093421
Color_Cluster	0.073619
cluster price	0.072608

Come si evince dalla tabella le caratteristiche dimensionali della borsa, in particolare l'altezza e la profondità, hanno molta influenza nei risultati ottenuti. Invece le caratteristiche come il colore e il prezzo hanno un impatto minore, ciò era prevedibile dato che esse non sono elementi distintivi per le tipologie di borse ma anzi possono essere condivise da più categorie.

Successivamente è stato utilizzato il modello Random Forest ed anche per esso, il dataset è stato diviso per il 20% destinato per il test, mentre l'80% per il training. È stato impostato che il modello dovesse utilizzare cento alberi. Si è deciso questo numero perché è abbastanza alto da ottenere dei

buoni risultati ma non eccessivo da aumentare in modo esponenziale il costo computazionale. Inoltre è stato impostato che essi possano avere una profondità⁸⁹ massima di [2, 3, 4, 5, 10, 15, 20, 25, 30].

Dopo aver impostato il modello come descritto, è risultato che l'accuratezza è dell'95,5%. Questo numero così elevato potrebbe indurre a pensare che si sia verificato l'overfitting, e per controllare ciò, sono state confrontate l'accuratezza del test con l'accuratezza dell'addestramento ed è risultato che sono molto simili, dato che quella del training è dell'95,3%. Ciò conferma che il modello generalizza molto bene ai dati di test.

Analizzando più approfonditamente i risultati per ogni singola categoria di borsa, come si evince dalla tabella seguente, il modello ha dei risultati molto positivi, addirittura con "precision", "recall" e "f1-score" generalmente superiori a 0.90. La classe per il quale il Random Forest ha delle prestazioni leggermente peggiori sono "Shoulder Bags", ma comunque sono molto positive.

Tabella 10: Classification report del Random Forest

Class	Precision	Recall	F1-Score	Support
Backpacks	1.00	1.00	1.00	124
Crossbody Bags	0.89	0.93	0.91	328
Luggage and bags	1.00	1.00	1.00	77
Messenger Bags	0.95	0.99	0.97	323
Mini Bag	0.98	0.99	0.98	327
Shoulder bags	0.98	0.87	0.92	386
Tote Bag	0.95	0.97	0.96	402

Successivamente è stata calcolata la matrice di confusione, che riconferma come il modello ha degli ottimi risultati per tutte le categorie di borse. Solo le "Crossbody Bags" e le "Shoulder Bags" hanno un tasso di errore relativamente più alto. Le "Crossbody Bags" vengono principalmente confuse con le "Messenger Bags", le "Shoulder Bags" e le "Tote Bags", mentre le "Shoulder Bags" vengono frequentemente confuse con le "Crossbody Bags".

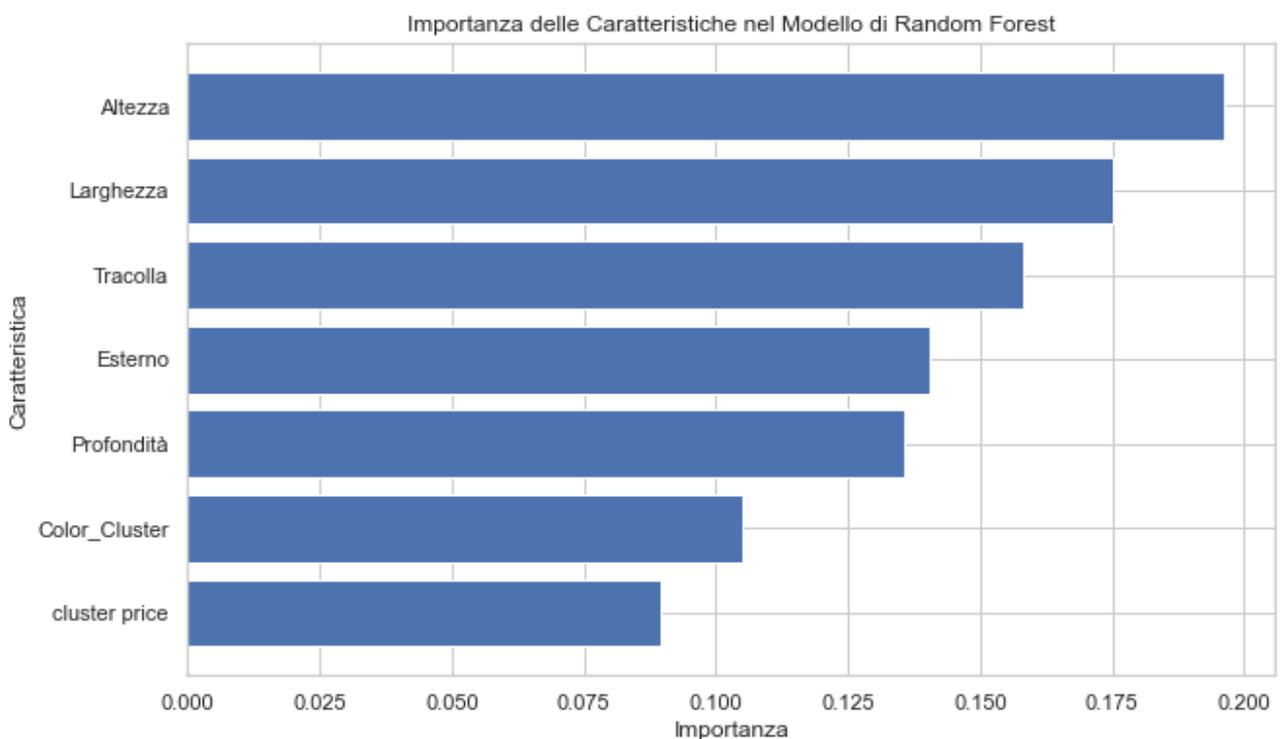
Tabella 11: Matrice di confusione del Random Forest

⁸⁹ Per profondità di un albero si intende il numero massimo di livelli (nodi) che l'albero può avere

	Backpacks	Crossbody Bags	Luggage and bags	Messenger Bags	Mini Bag	Shoulder bags	Tote Bag
Backpacks	124	0	0	0	0	0	0
Crossbody Bags	0	306	0	8	0	5	9
Luggage and bags	0	0	77	0	0	0	0
Messenger Bags	0	0	0	321	2	0	0
Mini Bag	0	1	0	0	323	1	2
Shoulder bags	0	36	0	3	2	337	8
Tote Bag	0	2	0	6	3	0	391

Infine, si sono volute evidenziare quali tra le variabili hanno avuto un impatto maggiore sul modello.

Grafico 14: Importanza delle features del Random Forest



Anche per questo modello, le caratteristiche dimensionali della borsa hanno influito in modo preponderante nel risultato della classificazione, portando a dei risultati eccellenti.

Mentre, viene confermato che il colore e il prezzo sono gli elementi che hanno un impatto minore sulle prestazioni del Random Forest.

Dopo aver testato il Random Forest e il Gradient Boosting, è stato utilizzato il modello Naive Bayes.

Anche in questo caso il dataset è stato diviso in 80% per il training e il 20% per il test. Nonostante il modello sia molto utilizzato e preciso, a causa soprattutto del dataset non del tutto bilanciato e della non perfetta indipendenza tra le variabili, l'accuratezza è risultata molto bassa, più precisamente del 45%.

Analizzando nel dettaglio i risultati per ogni singola tipologia di borsa, per le "Luggage and Bags" la classe è gestita molto bene dal modello, ma lo stesso commento non si può ampliare alle "Tote Bags" e le "Crossbody Bags" che hanno valori di "f1-score" molto bassi. Bisogna sottolineare come la categoria "Mini Bag" ha una "recall" estremamente bassa (0.21), ciò suggerisce che il modello non è in grado di riconoscerla correttamente nella maggior parte dei casi, nonostante una precisione più alta.

Tabella 12: Classification report del Naive Bayes

Classe	Precision	Recall	F1-Score	Support
Backpacks	0.62	0.83	0.71	124
Crossbody Bags	0.29	0.67	0.41	328
Luggage and Bags	0.85	1.00	0.92	77
Messenger Bags	0.45	0.51	0.48	323
Mini Bag	0.67	0.21	0.32	327
Shoulder Bags	0.51	0.49	0.50	386
Tote Bag	0.58	0.18	0.27	402

A sottolineare come questo modello non performi bene per questo studio, è la matrice di confusione. Come si evince dalla tabella seguente, infatti, le "Mini Bags" vengono spesso confuse con le "Crossbody Bags" e le "Messenger Bags", mentre le "Tote Bags" vengono frequentemente classificate come le "Crossbody Bags" e le "Shoulder Bags". In più questo modello spesso scambia le "Shoulder Bags" con le "Crossbody Bags".

Tabella 13: Matrice di confusione del Naive Bayes

Classe	Backpacks	Crossbody Bags	Luggage and Bags	Messenger Bags	Mini Bag	Shoulder Bags	Tote Bag
Backpacks	103	0	2	12	5	0	2
Crossbody Bags	0	221	0	50	0	48	9
Luggage and Bags	0	0	77	0	0	0	0
Messenger Bags	5	114	0	166	10	20	8
Mini Bag	41	148	0	65	68	5	0
Shoulder Bags	3	139	1	21	1	188	33
Tote Bag	15	129	11	51	17	107	72

Per quanto riguarda l'importanza delle variabili per la classificazione, a differenza dei precedenti modelli anche il materiale esterno della borsa ha una notevole importanza.

Tabella 14: Importanza delle features del Naive Bayes

Caratteristica	Importanza
Altezza	1.990134
Esterno	1.098476
Color_Cluster	0.862616
Larghezza	0.752201
Tracolla	0.370673
cluster price	0.150623
Profondità	0.135923

Successivamente è stato testato anche il modello XGBoosting, e anche in questo caso, come era stato fatto per il Random Forest, sono stati creati 100 alberi decisionali, con profondità massima di 6 nodi. Anche per questo modello si è diviso il dataset, destinando il 20% al test mentre l'80% al training. Una peculiarità del XGBoosting è il "learning rate", infatti, come spiegato nel paragrafo precedente, nel seguente modello ogni albero tenta di correggere l'errore residuo dei precedenti. Il "learning rate" è un fattore di ridimensionamento che limita la quantità di correzioni che ogni albero può apportare, e in questo caso si è scelto il valore 0.3⁹⁰.

⁹⁰ Questo valore è considerato alto, cioè ogni albero corregge in modo aggressivo gli errori dei precedenti

In generale il modello ha performato molto bene, registrando un'accuratezza del 95,5%. A causa del valore molto alto si è voluto confrontare l'accuratezza del test con quella dell'addestramento in modo da capire se si trattasse di un caso di overfitting. Dato che l'accuratezza del training è di 95,3%, non ci troviamo di fronte all'overfitting ma il modello riesce a generalizzare bene ai nuovi dati.

Analizzando più nel dettaglio i risultati per ogni singola classe di borsa, si è riscontrato un'elevata performance su tutte. C'è solo da sottolineare che "Crossbody Bags" e "Shoulder bags" hanno dei risultati leggermente peggiori nella "recall", ma comunque molto positivi come si evince dalla tabella di seguito.

Tabella 15: Classification report del XGBoosting

Classe	Precision	Recall	F1-Score	Support
Backpacks	1.00	1.00	1.00	124
Crossbody Bags	0.89	0.93	0.91	328
Luggage and Bags	1.00	1.00	1.00	77
Messenger Bags	0.95	0.99	0.97	323
Mini Bag	0.98	0.99	0.98	327
Shoulder Bags	0.98	0.88	0.92	386
Tote Bag	0.95	0.97	0.96	402

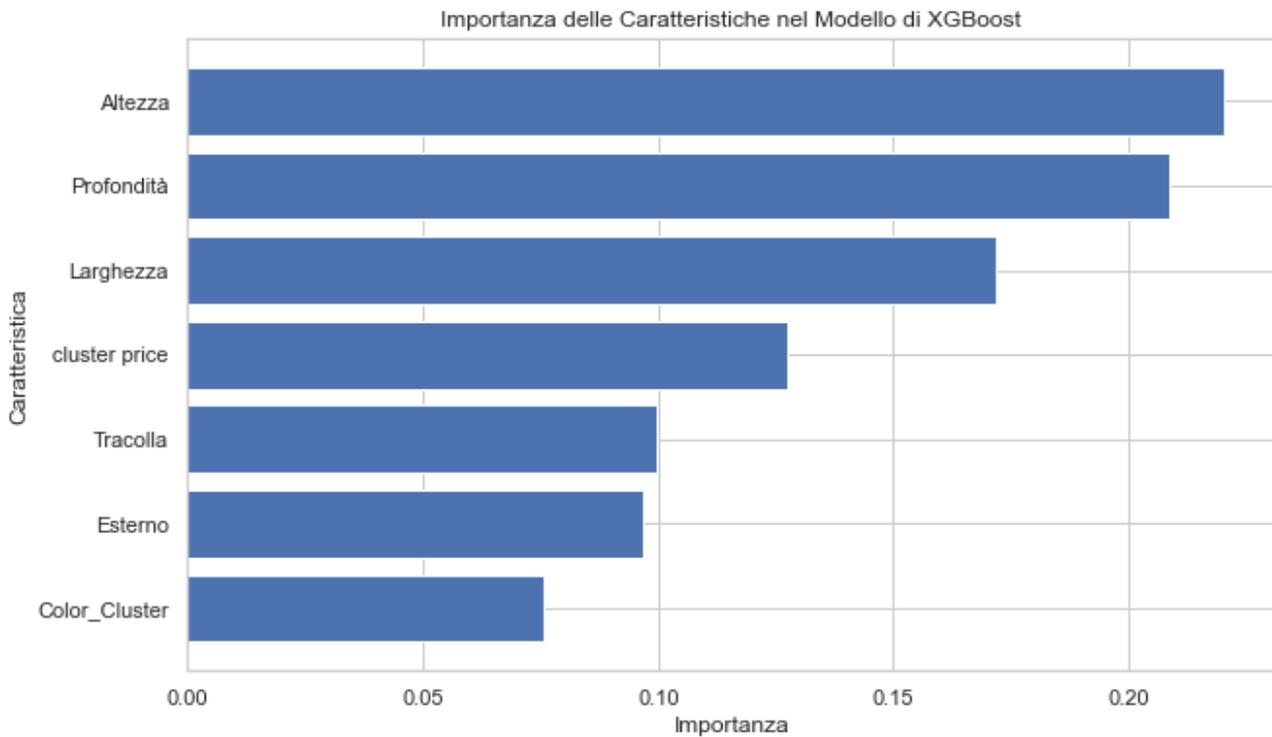
Questi ottimi risultati di classificazione sono anche confermati dalla matrice di confusione.

Tabella 16: Matrice di confusione del XGBoosting

Classe	Backpacks	Crossbody Bags	Luggage and Bags	Messenger Bags	Mini Bag	Shoulder Bags	Tote Bag
Backpacks	124	0	0	0	0	0	0
Crossbody Bags	0	304	0	8	0	7	9
Luggage and Bags	0	0	77	0	0	0	0
Messenger Bags	0	0	0	321	2	0	0
Mini Bag	0	1	0	0	323	1	2
Shoulder Bags	0	34	0	3	2	339	8
Tote Bag	0	2	0	6	3	0	391

Infine, si è voluto calcolare quali tra tutte le caratteristiche hanno avuto un impatto maggiore per il raggiungimento di questo risultato. Dal grafico sottostante si nota che ad inficiare molto sulle prestazioni sono le caratteristiche dimensionali della borsa. Invece, come avveniva per gli altri modelli, il colore e il prezzo non sono molto impattanti.

Grafico 15: Importanza delle features nel XGBoosting



Infine, come ultimo modello si è testato il modello SVM, dove anche in questo caso l'80% del dataset è stato utilizzato per l'addestramento del modello mentre il restante 20% per il test. Per una migliore prestazione del modello le caratteristiche sono state normalizzate usando il codice "StandardScaler", e ciò è stato molto importante perché l'SVM è molto sensibile alla scala dei dati.

Il modello, nonostante le accortezze utilizzate, non ha performato in modo ottimale, ottenendo un'accuratezza del 53,5%.

Tabella 17: Classification report del SVM

Classe	Precision	Recall	F1-Score	Support
Backpacks	0.77	0.79	0.78	428
Crossbody Bags	0.39	0.51	0.44	1421
Luggage and Bags	0.93	1.00	0.97	280
Messenger Bags	0.51	0.55	0.53	1246
Mini Bag	0.50	0.70	0.58	1281
Shoulder Bags	0.53	0.41	0.46	1614
Tote Bag	0.58	0.32	0.41	1598

La tabella “Classification report del SVM” mostra come le classi che sono meglio gestite dal modello sono “Luggage and Bags” e “Backpacks”, ma le altre tipologie di borse hanno dei risultati molto negativi in termini di classificazione.

Questi risultati non positivi vengono sottolineati anche dalla matrice di confusione, che mostra come spesso le classi vengano classificate in modo errato. Tra queste quella che più spesso non viene riconosciuta è la “Crossbody Bags” che spesso viene considerata erroneamente come “Tote Bags”, “Mini Bag” e “Shoulder Bags”. Il modello riesce a gestire bene solo le “Backpaks” e le “Luggage and Bags”, come viene mostrato dalla tabella sottostante.

Tabella 18: Matrice di Confusione del SVM

	Backpacks	Crossbody Bags	Luggage and Bags	Messenger Bags	Mini Bag	Shoulder Bags	Tote Bag
Backpacks	102	0	0	6	10	3	3
Crossbody Bags	0	178	0	34	46	46	24
Luggage and Bags	0	0	77	0	0	0	0
Messenger Bags	11	78	0	169	26	20	19
Mini Bag	6	33	0	51	232	5	0
Shoulder Bags	4	99	1	4	80	164	34
Tote Bag	3	71	3	58	60	76	131

Per quanto riguarda le caratteristiche che più hanno avuto un impatto, ad inficiare maggiormente sono state le dimensioni e, diversamente dagli altri modelli, il prezzo.

Tabella 19: Importanza delle features nel SVM

Caratteristica	Importanza
Larghezza	3.386443
Altezza	2.752441
Tracolla	2.267204
cluster price	2.161796
Profondità	0.571568
Esterno	0.539695
Color_Cluster	0.200847

In conclusione, l'analisi ha mostrato come diversi modelli di machine learning possano avere performance variabili a seconda dello scopo e del dataset. Il Random Forest e il XGBoosting si sono dimostrati i modelli più efficaci, con accuratezze superiori al 95% e senza avere problemi come l'overfitting, confermando la loro capacità di generalizzare bene ai nuovi dati. Entrambe, come in generale tutti i modelli utilizzati, hanno beneficiato delle caratteristiche dimensionali delle borse come altezza e profondità, risultando cruciali per una corretta classificazione.

Questa osservazione conferma che le caratteristiche delle borse che non sono visibili o comunque non possono essere rilevate in modo accurato da un'immagine, hanno un forte impatto positivo nelle task di classificazione delle borse, e che quindi possono essere utilizzate dai vari siti e-commerce per fini manageriali.

Tornando ai modelli, al contrario, il Naive Bayes e l'SVM hanno mostrato performance inferiori, in particolare per alcune categorie di borse, evidenziando la necessità e l'importanza di scegliere accuratamente il modello in base alla natura dei dati. Hanno confermato che questi modelli non sono indicati per ricerche che prevedono l'utilizzo di un dataset molto grande e complesso e non perfettamente bilanciato, come nel caso di questo studio.

4.2 SUDDIVISIONE PER OCCASIONE D'USO

Come indicato precedentemente quando è stato affrontato il gap di ricerca, lo studio condotto non si limita solo a voler trovare quale algoritmo di machine learning è più adatto e quale dia migliori risultati per la classificazione delle borse. Lo scopo ultimo della ricerca è quello di, dopo aver suddiviso le tipologie di borse precedentemente classificate per occasione d'uso, proporre delle iniziative di marketing e manageriali in grado di migliorare le prestazioni e l'incidenza del sito e-commerce sul fatturato dell'azienda e di conseguenza, di dare dei consigli o una nuova metodologia sulla classificazione delle borse nella piattaforma.

Si è deciso di dividerle per occasioni d'uso per impostare il sito e-commerce su un approccio euristico, basandolo così su regole pratiche e intuizioni capaci di migliorare l'esperienza dell'utente e facilitarne la navigazione. Comprendendo le diverse occasioni in cui i consumatori sono più propensi a prendere decisioni d'acquisto, le aziende possono adattare i loro sforzi di marketing per raggiungere e coinvolgere efficacemente i potenziali clienti. Questo approccio di segmentazione riconosce che il comportamento dei consumatori può variare significativamente a seconda dell'occasione, che si tratti di una festività, di un evento speciale o anche di un giorno della settimana⁹¹.

Nel contesto di questa ricerca e quindi di un catalogo e-commerce di borse, suddividendo le borse in base all'occasione d'uso, le aziende possono allineare le loro offerte di prodotti alle esigenze e preferenze specifiche dei clienti, oltre che a personalizzare la comunicazione e le promozioni. Così facendo aumenterebbero la rilevanza e l'efficacia dei messaggi di marketing, portando così a una maggiore soddisfazione del cliente e a un potenziale incremento delle vendite.

Lo studio, quindi propone una clusterizzazione di borse secondo un modello euristico che è l'occasione di utilizzo. In un processo decisionale la parte euristica ha una grande importanza come viene sottolineato dalla ricerca svolta da K. Krabuanrat e R. Phelps svolta nel 1998⁹². Un'altro studio afferma che "l'euristica è importante nelle prime fasi del processo decisionale per formare un insieme di considerazioni, che consiste nell'eliminare la maggior parte dei prodotti da ulteriori considerazioni"⁹³. Perciò la suddivisione di borse in cluster creati secondo l'occasione di utilizzo permette una ricerca più immediata e intuitiva del prodotto che stiamo cercando.

⁹¹ Indeemo. (2024). The Importance of Occasions in Consumer Research. website: <https://indeemo.com/blog/consumer-research-occasions>

⁹² Krabuanrat, K., & Phelps, R. (1998). Heuristics and rationality in strategic decision making: An exploratory study. *Journal of Business Research*, 41(1), 83–93. [https://doi.org/10.1016/s0148-2963\(97\)00014-3](https://doi.org/10.1016/s0148-2963(97)00014-3)

⁹³ Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, 62(1), 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>

Perciò, prima di tutto le aziende quando inseriranno la borsa nel loro sito e-commerce potranno utilizzare i modelli più performanti per classificarla e una volta fatto ciò verrà inserita nel cluster della tipologia d'uso per rendere più facile la navigazione all'utente e per poter avviare delle iniziative di marketing mirate.

Come sottolineato nella descrizione del dataset e delle sue varie variabili i tipi di borsa presenti in questo studio sono molteplici: backpacks, beach bag, belt bag, bucket bag, clutch bag, crossybody bag, laptop & briefcase, luggage and bag, messenger bag, mini bag, shoulder bag, tote bag, travel accessories, wallet & cardholder.

Analizzando le caratteristiche e le dimensioni di queste varie tipologie di borse esse possono essere divise in 5 categorie come elencate di seguito.

Tabella 20: La divisione delle borse per occasione d'uso

OCCASIONE D'USO	TIPOLOGIE DI BORSE
Lavoro e Ufficio	Laptop & Briefcase, Messenger Bag
Viaggi e Avventure	Luggage and Bag, Travel Accessories
Tempo libero e Attività Quotidiane	Crossybody Bag, Shoulder Bag
Eventi e Occasioni Speciali	Clutch Bag, Mini Bag
Uscite fuori porta	Backpack, Beach Bag, Bucket Bag, Tote Bag

Bisogna precisare che all'interno delle categorie sono stati esclusi i portafogli e i porta carte perché non sono considerate delle tipologie di borse ma degli accessori a sé stanti.

4.2.1 LAVORO E UFFICIO

Analizzando la segmentazione delle tipologie di borse per occasione d'uso, partendo dalla categoria "Lavoro e Ufficio", sono state inserite all'interno di questa occasione d'uso le seguenti borse" e le "Messenger Bag".

In generale le borse per l'ufficio o più in generale per il lavoro non sono solo accessori di moda ma riflettono anche la professionalità e l'organizzazione di chi le indossa. Inoltre, la borsa per essere considerate da lavoro deve avere determinate caratteristiche tecniche. Esse, ad esempio, devono avere degli appositi scompartimenti per contenere il computer portatile, i documenti ed altri oggetti utili. Una buona organizzazione della borsa riduce lo stress legato alla confusione e al tempo impiegato per cercare gli oggetti che, inevitabilmente è maggiore nelle borse dotate di un unico scompartimento. Un'altra caratteristica di questa tipologia di borse sono i materiali resistenti e robusti utilizzati per far fronte all'uso quotidiano. In collegamento a questa peculiarità le borse cosiddette "da lavoro" hanno spesso un materiale di alta qualità, come la pelle, che, oltre ad essere resistenti, conferiscono eleganza e serietà a chi la utilizza.

La scelta delle borse "Laptop & Briefcase" e "Messenger Bag" è frutto di un'osservazione delle caratteristiche di esse. Entrambe offrono una combinazione di eleganza, funzionalità e organizzazione che sono fondamentali per rispondere al bisogno quotidiano di un lavoratore.

Più nello specifico le borse "Laptop & Briefcase" sono state appositamente progettate per trasportare laptop e tablet, ed, oltre a ciò, spesso hanno degli scompartimenti aggiuntivi per mantenere il contenuto della borsa in modo distinto e organizzato. Oltre a ciò, come è osservabile dal dataset, per la maggior parte dei casi sono composte da pelle o comunque da materiali di alta qualità, che gli conferiscono professionalità, che è una caratteristica richiesta nell'ambito lavorativo.

Per quanto riguarda invece le "Messenger Bag" riescono a combinare l'utilità, data dalla grandezza e dalla presenza di vari scompartimenti, alla praticità, permessa dalla presenza della tracolla che consente al lavoratore di avere le mani libere durante gli spostamenti tra uffici e riunioni.

Le "Messenger Bag" pur avendo un aspetto professionale, riescono a conferire a chi le sta utilizzando un aspetto casual, che è ideale per alcuni ambienti lavorativi meno formali.

Per questa serie di ragioni, queste due tipologie di borse sono ideali per l'ambiente lavorativo.

4.2.2 VIAGGI E AVVENTURE

Le borse da viaggio non solo vengono scelte per la loro estetica ma devono garantire comfort e sicurezza. Infatti, questa tipologia di borse ha il compito di rendere i viaggi rilassati, cioè devono far sì che il possessore della borsa non si preoccupi dell'integrità del contenuto e della resistenza a spostamenti e urti del bagaglio.

Le borse da viaggio e avventura sono spesso realizzate con materiali molto resistenti e di qualità, come il nylon e varie tipologie di pelle. Essi sono stati scelti apposta per la loro capacità a resistere alle più disparate condizioni atmosferiche, agli strappi durante i vari trasporti e all'usura, garantendo la sicurezza del contenuto della borsa.

Queste tipologie di borse si distinguono dalle altre per le dimensioni, infatti, per quanto riguarda i "Luggage and Bag" sono dotati di misure elevate, che conferiscono alla borsa molto spazio. Mentre, in modo diametralmente opposto, i "Travel Accesories" hanno dimensioni molto ridotte per essere inserite comodamente nelle valigie o nelle stesse "Luggage and Bag".

La scelta di inserire queste due tipologie di borse all'interno di questo gruppo deriva dalle caratteristiche uniche e specifiche per rendere il viaggio confortevole e sicuro. Infatti, esse offrono una buona capacità di stoccaggio, organizzazione e resistenza, ed inoltre, con l'utilizzo combinato di entrambe, i "Travel Accesories" conferiscono maggiore convenienza e organizzazione dello spazio.

Quindi che si tratti di viaggi di lavoro, in famiglia o all'avventura, è molto importante avere il giusto equipaggiamento anche in termini di bagagli e borse per godersi al meglio il viaggio e per renderlo meno stressante.

4.2.3 TEMPO LIBERO E ATTIVITA' QUOTIDIANE

Oggi giorno le vite di ognuno di noi sono diventate frenetiche e di conseguenza tutti gli oggetti si sono adattati a questo cambiamento di stile di vita. Anche le borse sono diventate degli strumenti che aiutano le persone a gestire al meglio la quotidianità e a godersi il tempo libero da impegni.

Queste determinate tipologie di borse devono avere delle caratteristiche che le permettano di essere effettivamente utili per rispondere alle necessità di chi le indossa.

Le “Shoulder Bag” sono delle borse molto in voga per il tempo libero e per le attività quotidiane. Queste borse sono progettate per essere portate facilmente su una spalla, consentendo a chi le indossa un accesso facile e veloce agli oggetti contenuti. In generale esse sono disponibili in vari formati, il che le rende perfette per ogni evenienza durante le attività giornaliere e di svago.

Sono composte da una varietà molto alta di materiali, infatti, possono avere una composizione in pelle o di tessuti, garantendo un'ampia versatilità di modelli che permettono di rispecchiare i gusti personali sul design della borsa.

Parlando delle “Crossbody Bag”⁹⁴, esse sono ideali per tutte le necessità legate alla quotidianità. Questa tipologia di borsa distribuisce il peso in modo uniforme su tutto il corpo, permettendo a chi le indossa di non sovraccaricare la spalla e di poterla indossare comodamente anche per un periodo di tempo prolungato. In aggiunta consentono di avere le mani libere e quindi di non avere delle limitazioni.

Le caratteristiche che distinguono le “Crossbody Bag” sono la loro sicurezza e la comodità, che la rendono una tipologia di borsa adatta per quelle attività quotidiane che richiedono anche un tempo prolungato. Inoltre, la maggior parte di esse dispone di piccoli scompartimenti interni che permettono di poter organizzare il contenuto della borsa in separate sezioni, così da facilitare la ricerca di ciò che serve senza dover svuotare la borsa intera.

Per tutta questa serie di ragioni è stato deciso di inserire queste due tipologie di borse nella categoria “Tempo libero e Attività quotidiane”.

⁹⁴ Note anche come borse a tracolla lunga

4.2.4 EVENTI E OCCASIONI SPECIALI

Come sottolineato nelle categorie precedenti, le borse oltre che ad arricchire il look, hanno anche delle funzioni pratiche che permettono di essere molto utili durante le attività quotidiane o durante il lavoro. Tuttavia, per eventi speciali, la scelta della borsa è molto importante ed ha un impatto molto forte sull'aspetto finale della persona.

Tra le tipologie di borse quelle più indicate per queste tipologie di occasioni sono la “Clutch Bag” e la “Mini Bag”.

Le prime sono delle borse di dimensioni ridotte senza manici, spesso di forma rettangolare, che sono progettate per essere portate in mano e che conferiscono molta eleganza a chi le indossa. Esse sono sufficientemente piccole per essere portate comodamente in mano, ma comunque hanno una grandezza tale da poter contenere gli oggetti necessari come i documenti e il telefono.

Invece le “Mini Bag” sono delle borse, che, come indica il nome, sono di dimensioni ridotte ma comunque abbastanza grandi da contenere gli oggetti indispensabili. Esse, a differenza delle “Clutch Bag” possono avere la tracolla e quindi dare anche un'alternativa di stile per chi le indossa.

Entrambe offrono la comodità e la praticità di contenere solo l'essenziale, riducendo drasticamente il peso e l'ingombro. Questa caratteristica le rende ideali per eventi e serate speciali, dove non si vuole essere appesantiti da borse voluminose. Inoltre, grazie alla loro forma ed essendo composte da materiali ricercati e pregiati, riescono ad aggiungere un tocco di classe e di eleganza all'outfit.

Per questi motivi sono ideali per gli eventi speciali e sono un must – have nel guardaroba di ogni donna o uomo.

4.2.5 USCITE FUORI PORTA

Le gite fuori porta sono l'occasione per staccare dalla propria routine e dal lavoro di tutti i giorni. Perciò la parola che più riassume queste situazioni deve essere relax, e il compito di rendere la giornata rilassante lo ha anche la borsa, che deve essere adatta al tipo di giornata.

Che sia al mare, in montagna o un semplice picnic, la borsa giusta può fare la differenza in termini di comodità e comfort. A causa appunto dell'eterogeneità di esperienze che si possono fare durante le gite fuori porta, questo cluster di borse contiene molte tipologie di esse per rispondere a tutte le varie necessità. Ma nonostante apparentemente esse possano sembrare molto differenti tra loro hanno delle caratteristiche come le dimensioni, la durabilità e la comodità.

Entrando più nel dettaglio, esse devono avere la grandezza necessaria a poter contenere tutto il necessario, che sia cibo, bevande o eventuali accessori come i libri, giochi e occhiali da sole.

Tutte le borse selezionate sono composte da materiali molto resistenti come il nylon e varie tipologie di tessuti studiati appositamente per essere resistenti a varie condizioni ed elementi come l'acqua, la sabbia e la polvere. Inoltre, queste tipologie di materiali sono anche molto facili da pulire, come il nylon e il poliestere, che hanno questa caratteristica.

Per questa determinata occasione d'uso sono state incluse le seguenti tipologie di borse: "Backpack", "Beach Bag", "Bucket Bag", "Tote Bag".

I "Backpack", o più comunemente conosciuti come zaini, sono noti per la loro ampia grandezza e divisione in più scompartimenti. Ciò permette a chi li indossa di poter avere con sé una grande quantità di oggetti, che possono essere organizzati nelle varie sezioni e tasche dello zaino, in modo molto comodo. Infatti, gli zaini distribuiscono il peso uniformemente sulle spalle e sulla schiena, riducendo drasticamente la fatica e l'ingombro durante la giornata.

Per quanto riguarda le borse da spiaggia, le "Beach Bag", la loro caratteristica principale è lo spazio e la capienza, perché sono pensate per poter portare gli asciugamani, creme solari, libri e oggetti per la spiaggia. Anche parlando di materiali, essa è perfetta per queste occasioni perché è perlopiù realizzata da componenti molto resistenti e leggeri come la tela, il nylon e altri materiali che sono resistenti all'acqua. Nonostante siano delle borse concepite per essere funzionali, sono disponibili in vari colori e design che conferiscono all'accessorio anche un ruolo estetico.

In questa categoria sono state inserite anche le "Bucket Bag", che, come suggerisce il nome, hanno una forma a secchiello. Anch'esse si contraddistinguono dall'elevata capacità di stoccaggi al loro

interno, ma avendo una dimensione ridotta rispetto alle “Beach Bag”, consentono un facile accesso agli oggetti all’interno della borsa.

L’ultima tipologia di borsa inserita in questo cluster sono le “Tote Bag”. La caratteristica principale è la grande spaziosità al suo interno ed anche un facile accesso all’oggetto che si sta cercando.

Esse sono delle borse che pur essendo molto semplici hanno un grande impatto sullo stile della persona oltre che ad essere molto funzionali soprattutto per picnic e gite di una giornata.

La classificazione attraverso algoritmi di machine learning, e quindi con un assetto data-driven, consente non solo di migliorare l’efficienza del sito, ma anche di offrire un servizio migliore in termini di esperienza e di soddisfazione per il cliente e, di conseguenza, delle vendite e dei ricavi.

La suddivisione in cluster di prodotti omogenei per occasione d’uso, per avere la massima efficacia, deve essere accompagnata da delle iniziative di marketing precise e personalizzate per il cliente.

Queste iniziative potrebbero migliorare l’esperienza e la fidelizzazione del cliente, che si tradurrebbero in un maggiore introito per l’azienda, oltre che ad una maggiore visibilità e originalità della strutturazione del sito.

4.3 L'UTILIZZO PRATICO DELLO STUDIO

Come analizzato e descritto nel capitolo precedente, i vari algoritmi di machine learning utilizzati per la classificazione, hanno dei risultati molto positivi e accuratezza significativa nel raggruppamento delle borse in categorie.

Questi risultati, oltre a poter essere di spunto per ricerche future o di approfondimenti, hanno un reale impatto e utilità nel gestire i siti e-commerce. Implementando e utilizzando questi algoritmi, è possibile automatizzare il processo di classificazione delle borse presenti nel sito e addirittura quelle che verranno inserite successivamente, per poi poter creare delle determinate categorie di borse, descritte nel capitolo precedente, in base all'occasione d'utilizzo.

Quindi il fine ultimo di questo studio è quello di permettere di migliorare l'organizzazione del sito e-commerce partendo da un'automatizzazione della classificazione delle borse eseguita con le sole caratteristiche, per poi poterle inserire in cluster ben distinti per occasione d'uso. Perciò, esso non migliora solo il layout delle pagine e-commerce, ma facilita e migliora l'esperienza e la navigazione degli utenti nel sito.

Grazie alla classificazione accurata e al successivo inserimento delle borse nei cluster descritti, i clienti possono esplorare il sito in modo più intuitivo e trovare l'item perfetto per l'occasione in modo molto veloce. La riduzione drastica del tempo necessario per la ricerca minimizza la frustrazione e l'insoddisfazione, portando così il cliente a preferire quel determinato sito rispetto a quello dei competitors.

Inoltre, questa divisione in cluster permette all'azienda di avviare delle iniziative di marketing mirate e personalizzate, che hanno molto successo perché riescono a creare delle “esperienze utente uniche e memorabili”⁹⁵. Il marketing personalizzato può anche aumentare l'engagement degli utenti che si interfacciano con la piattaforma e-commerce, la conversione e migliorare il rapporto tra il cliente e la piattaforma creando una forte fidelizzazione.

⁹⁵ Marketing personalizzato – un approccio alla personalizzazione della comunicazione con i clienti - Copymate. (2024). Retrieved from Copymate website: <https://copymate.app/it/blog/multi/marketing-personalizzato-un-approccio-alla-personalizzazione-della-comunicazione-con-i-clienti/#:~:text=Il%20marketing%20personalizzato%20pu%C3%B2%20aumentare>

4.4 INIZIATIVE DI MARKETING

4.4.1 L'IMPORTANZA DEGLI OBIETTIVI "SMART"

Come più volte sottolineato la ricerca del migliore algoritmo di classificazione e la successiva classificazione delle borse per occasione d'uso, devono essere accompagnate da delle iniziative di marketing volte a migliorare l'engagement e l'esperienza del consumatore.

Le iniziative di marketing hanno il compito di sfruttare la classificazione precedentemente eseguita, per proporre delle azioni mirate a quel determinato cliente e soprattutto a quella determinata occasione di utilizzo della borsa.

Il primo passo per creare una campagna marketing è quella di definire bene gli obiettivi e il come calcolarli. Sul come monitorarli nei prossimi sottocapitolo saranno enunciati e descritti alcuni KPI utili a questo scopo, mentre per quanto riguarda gli obiettivi è importante che essi siano "SMART"⁹⁶. Esso è un acronimo di specific, measurable, achievable, relevant, time-bound.

Specific: l'obiettivo deve essere ben definito, perché più un obiettivo è specifico più è facile capire la direzione che si deve prendere per raggiungere quel determinato risultato.

Measurable: l'obiettivo deve essere misurabile attraverso degli indici o specifici KPI che possano valutare i progressi e che siano utili per poter aggiustare in corso d'opera le azioni intraprese.

Achievable: l'obiettivo deve essere realistico, cioè deve poter essere raggiunto tenendo conto delle risorse a disposizione. Un risultato atteso troppo ambizioso potrebbe demoralizzare il team e compromettere o addirittura danneggiare l'operato futuro, mentre un obiettivo troppo facile da raggiungere potrebbe non incentivare le varie persone a dare il massimo nel proprio lavoro.

Time-Bound: l'obiettivo deve avere una scadenza definita e decisa, in modo da poter definire un arco temporale di monitoraggio. Questa temporalità aiuta a mantenere il focus e a motivare i vari team a dare il massimo entro una scadenza predefinita.

È perciò importante che gli obiettivi siano "SMART", soprattutto perché definiscono con chiarezza la direzione da dover prendere ma anche permettono di misurare i progressi, motivare il team e ottimizzare le risorse a propria disposizione.

⁹⁶ Questa proposta riguardo le caratteristiche degli obiettivi manageriali è stata introdotta nel 1981 da George T. Doran nella sua pubblicazione intitolata "There's a S.M.A.R.T. way to write management's goals and objectives". Nel quale sottolinea che gli obiettivi devono essere: specifici, misurabili, raggiungibili (achievable), realistici, temporizzati

In ultima, bisogna sottolineare che questi obiettivi devono essere in linea con la strategia aziendale complessiva, in modo che essi contribuiscano significativamente al successo dell'azienda.

4.4.2 CAMPAGNE TEMATICHE STAGIONALI

Una delle strategie di marketing che le aziende che utilizzano i siti e-commerce e la classificazione di borse proposta da questa ricerca potrebbero attivare è sicuramente il lancio di campagne a tema per ogni stagione.

Questo permetterebbe di sfruttare le esigenze stagionali, mantenendo l'interesse dei clienti alto in tutti i periodi dell'anno, offrendogli anche delle promozioni ad hoc per invogliarli ancor di più ad acquistare la borsa.

Durante il periodo di primavera ed estivo l'azienda nel proprio sito e-commerce potrebbe promuovere maggiormente le borse pensate per le uscite fuori porta e per viaggi ed avventure. Questo perché, in questi periodi, le persone iniziano a trascorrere più tempo all'aperto e a spostarsi nei fine settimana per piccole gite oppure programmano e partono per lunghi viaggi durante l'estate. Per far fronte al bisogno di comodità e di organizzazione per picnic, escursioni, giornate in spiaggia o in montagna, le borse facente parti del cluster delle gite fuori porta potrebbero essere di molto interesse per i consumatori. Invece per i viaggi di più giorni o che comportano dei grossi spostamenti, sono perfette quelle borse rientranti nel gruppo dei viaggi e avventura.

La strategia che potrebbe essere adottata è quella di creare una landing page dedicata. Secondo la definizione di Korniiichuk, R., & Boryczka, M., essa “nel marketing digitale [...] è una pagina web indipendente creata ai fini di una campagna di marketing o pubblicitaria”⁹⁷.

Essa potrebbe mostrare le borse in ambientazioni primaverili come foreste di montagna, ruscelli, parchi, spiagge per evocare e aumentare il desiderio di partire. Mentre per invogliare l'acquisto di borse specifiche per viaggi lunghi e dall'altra parte del mondo, le landing page potrebbero avere come ambientazione un aeroporto o un qualcosa che ricordi dei viaggi molto lunghi e che faccia venire il desiderio di comodità e di voglia di iniziare un viaggio.

Di seguito saranno presentati alcuni esempi.

⁹⁷ Korniiichuk, R., & Boryczka, M. (2021). Conversion Rate Prediction Based on Text Readability Analysis of Landing Pages. *Entropy*, 23(11), 1388. <https://doi.org/10.3390/e23111388>

Tutti i filtri ☰

Borse to

> Ordina per ▾



Nuova stagione
Ferragamo
Borsa a tracolla Fiamma media



stagione
spalla Joie

Tutti i filtri ☰

Borse

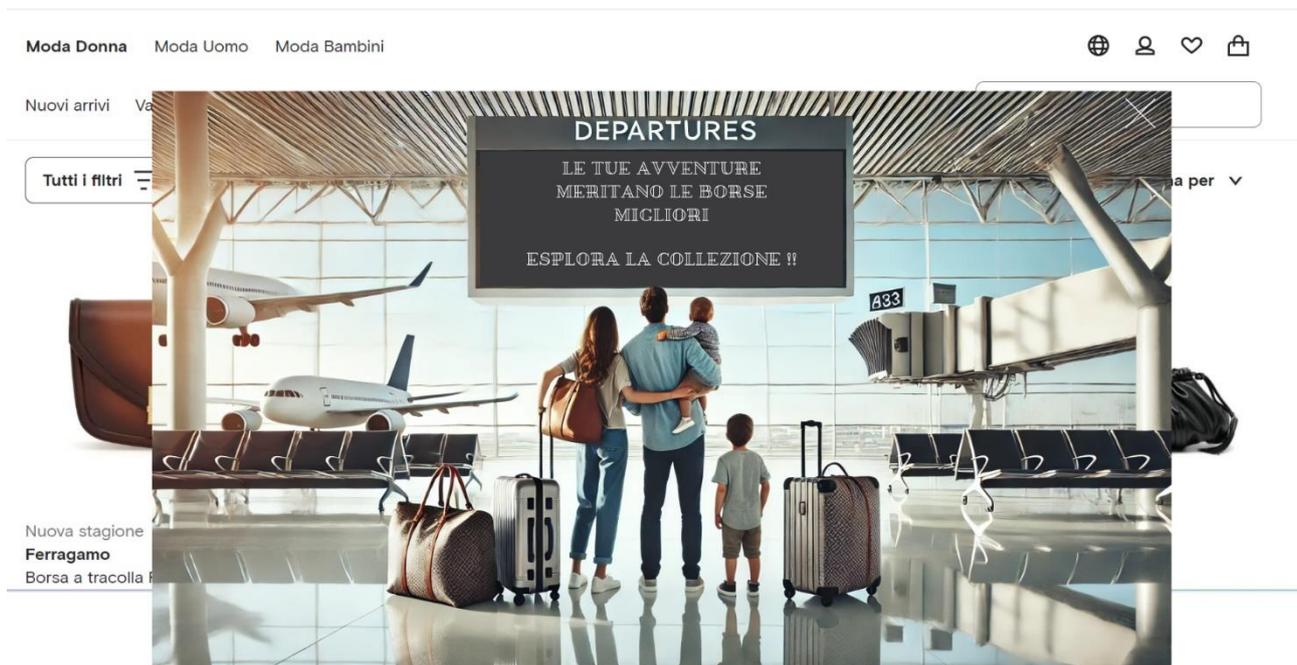
> Ordina per ▾



Nuova stagione
Ferragamo
Borsa a tracolla Fiamma me



zione
sulla Joie



Oltre alla parte grafica della landing page, si possono inserire delle mappe con dei percorsi suggeriti per una gita fuori porta, con l'opzione di scaricarle dopo aver inserito la propria e-mail. In questo modo gli utenti potrebbero ricevere un'offerta personalizzata in base alle loro preferenze.

La prima mail potrebbe contenere la mappa completa con la possibilità di impostare dei filtri come la distanza che si vuole percorrere, il livello di difficoltà, la presenza o meno di bambini ed altre informazioni utili per consigliare il miglior programma per un giorno. Le mail successive potrebbero includere delle promozioni in modo da spingere il consumatore a comprare quella determinata borsa.

Mentre per le stagioni come l'autunno e l'inverno, il sito e-commerce dovrebbe promuovere altre categorie di borse. Infatti, nel periodo di settembre, con il rientro dalle ferie estive e con il rientro in ufficio e a scuola, molte persone cercano di rinnovare il proprio stile, gli accessori da lavoro e da ufficio, comprese le borse e porta laptop.

In questo periodo le aziende dovrebbero anche promuovere quelle borse che sono utilizzate per le attività quotidiane e per il tempo libero. Con il rientro dai vari viaggi estivi, ricomincia la routine di tutti i giorni e le persone tendono a ripensare al loro stile e ai loro accessori per aggiungere un tocco di stile ai loro outfit e per stravolgere un minimo la routine, dandogli un senso di nuovo inizio.

Infine, il periodo invernale e autunnale è un periodo di molte festività, dal Natale a Capodanno e con molte cerimonie nei mesi che precedono l'inverno. Per questo motivo le aziende dovrebbero mettere in risalto e promuovere le borse per eventi e occasioni speciali.

Quindi, durante questo periodo le categorie di borse che vengono utilizzate per l'ufficio, per eventi speciali e per la quotidianità sono molto rilevanti e ricercate. Perciò, promuovere questi articoli con campagne mirate e offerte speciali può aiutare, prima di tutto, ad aumentare e catturare l'interesse degli utenti e in seguito anche ad incrementare le vendite e di conseguenza il fatturato dell'azienda.

Anche per queste tipologie di borse una strategia di marketing è quella di utilizzare in questi periodi delle landing page personalizzate in modo da spingere il consumatore a cliccare sulla Call To Action.

In generale la landing page deve essere composta, oltre che da un'immagine suggestiva che invogli il cliente a cliccare, da un'apposita CTA⁹⁸. Quest'ultima è uno degli elementi più importanti della landing page perché invoglia e invita l'utente a cliccare per scoprire di cosa si tratta quella determinata offerta o proposta.

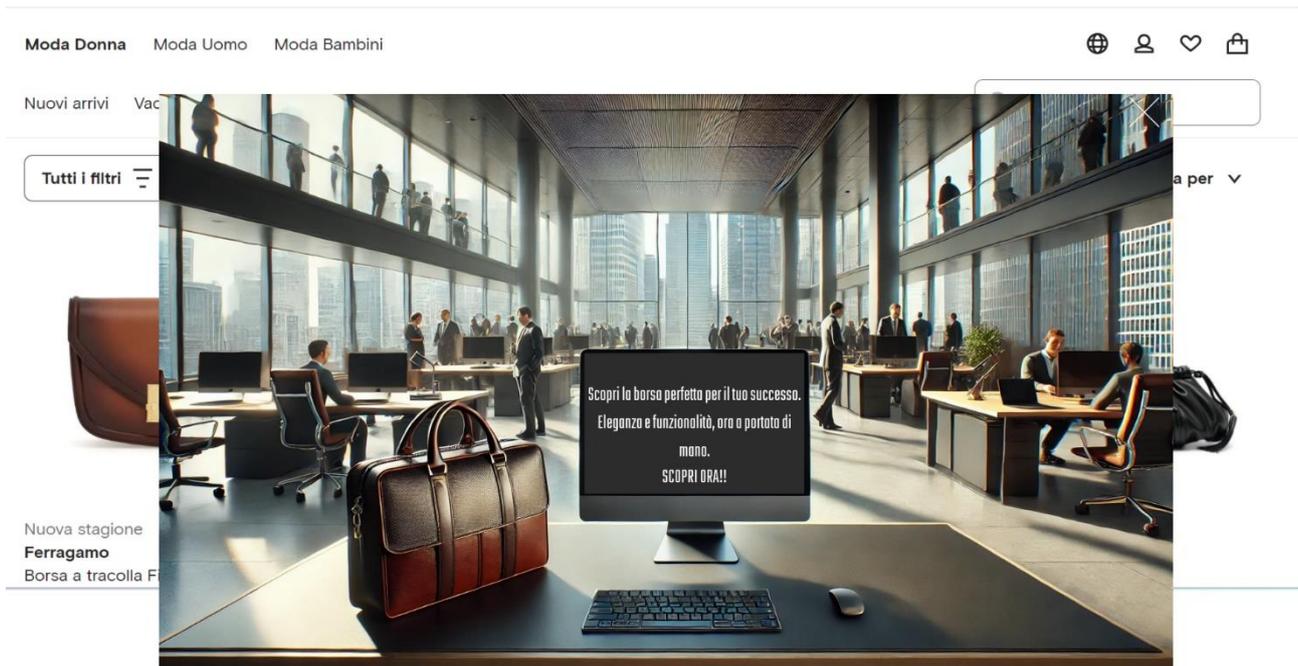
Una CTA deve essere chiara e ben strutturata perché fornisce un suggerimento preciso sul cosa fare al consumatore. In un mondo digitale dove sono presenti molte possibili distrazioni e la soglia dell'attenzione sta sempre di più diminuendo, è molto importante che i consumatori siano guidati nei vari passaggi successivi che lo portano ad acquistare.

Per quanto riguarda nello specifico le landing page, di seguito saranno mostrati degli esempi di come le aziende dovrebbero impostarle e proporle.

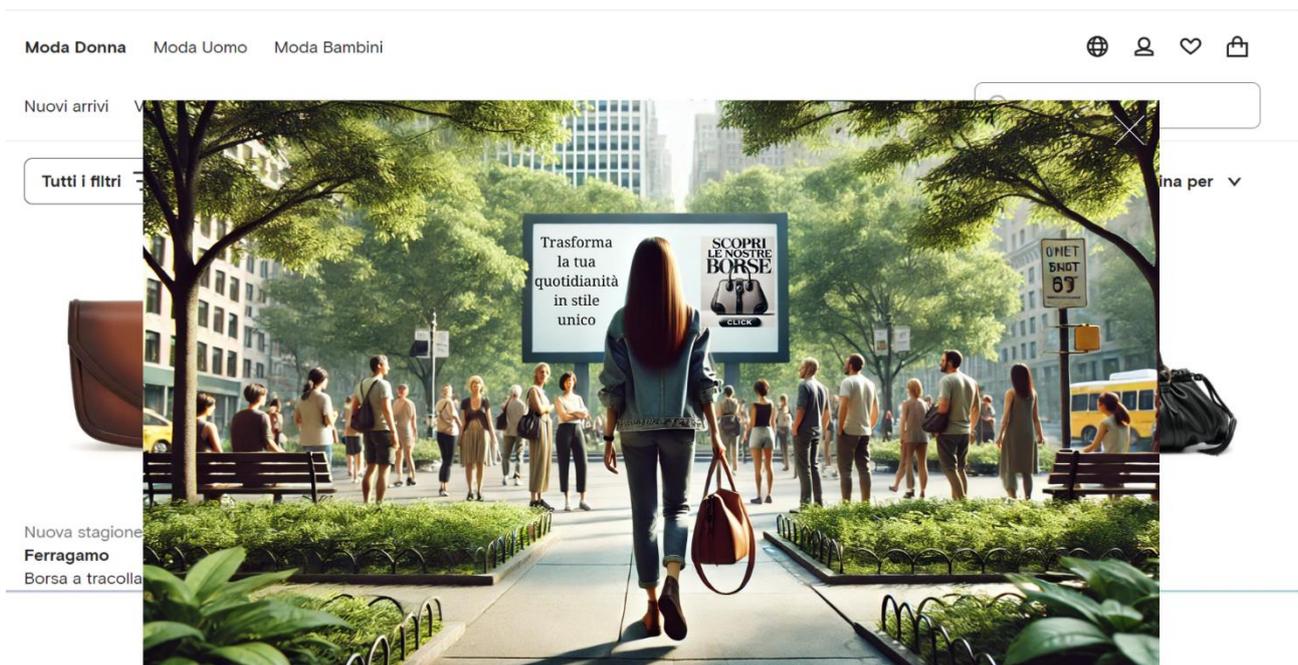
Esse dovrebbero mostrare un ambiente lavorativo o comunque un contesto che richiami il lavoro, in modo da invogliare il cliente a rinnovare il proprio outfit giornaliero comprando delle borse che uniscano lo stile con la praticità.

Di seguito un esempio di landing page per questa occasione d'uso.

⁹⁸ È un acronimo di Call To Action, cioè un invito all'utente a compiere una determinata azione

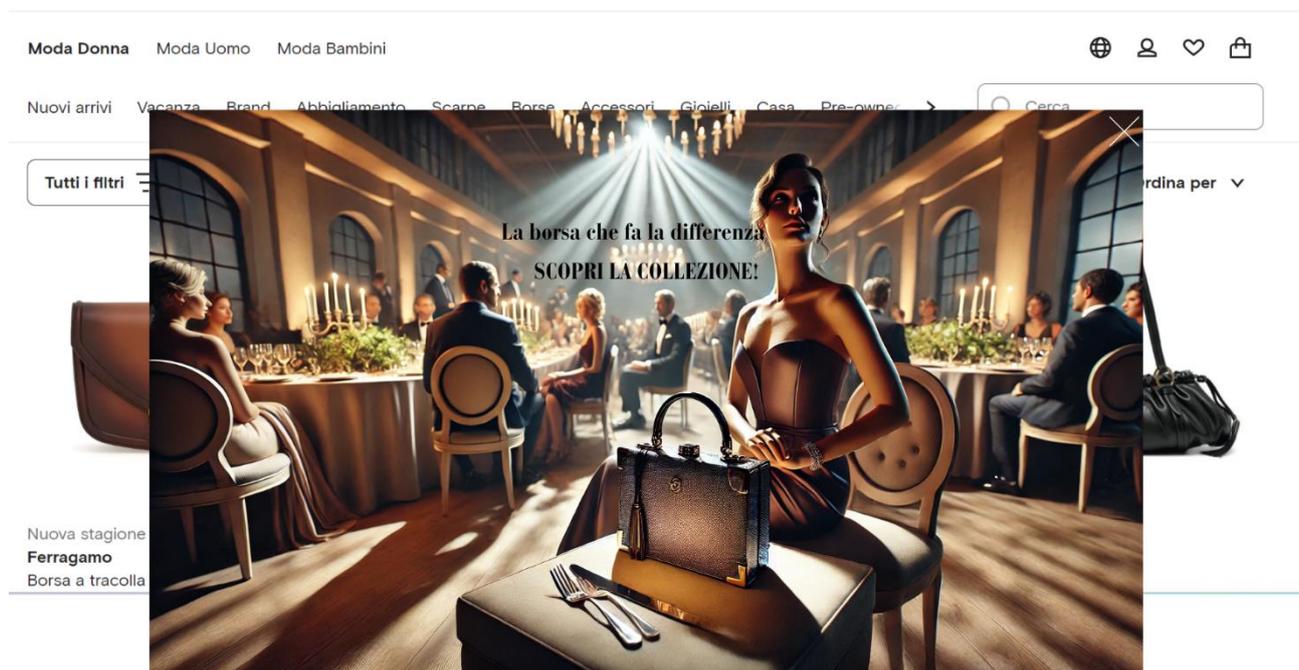


Per quanto riguarda le landing page per invogliare l'acquisto delle borse per le attività quotidiane e per il tempo libero, dovrebbero essere mostrate delle scene di vita quotidiana, come una passeggiata al parco, un aperitivo in città e altre attività che si svolgono nel tempo libero. Il messaggio che deve emergere dall'immagine è quello di unione tra quotidianità e stile, e per fare ciò la Call To Action deve sottolineare questa caratteristica come nell'esempio di seguito.



Infine, per le borse usate per eventi serali e grandi occasioni, dovrebbero essere fatte delle landing page raffiguranti una cena o un evento particolarmente elegante nel quale viene risaltata l'importanza della borsa che emerge sul resto della scena. Il tutto deve essere accompagnato da messaggi che facciano risaltare ancor di più l'importanza della borsa in queste occasioni, invitando così l'utente ad acquistare una di queste tipologie di borse.

Di seguito un esempio di landing page per questa occasione d'uso.



Oltre alla personalizzazione delle landing page per ogni tipologia di occasione d'uso, potrebbero essere offerte delle guide stagionali su dei consigli su come utilizzare al meglio le borse durante l'inverno e l'autunno. Focalizzandosi su come abbinare quella determinata tipologia di borsa o come organizzare al meglio il contenuto di essa.

Addirittura, le aziende potrebbero pensare ad un calendario editoriale che verrebbe inviato tramite mail ai clienti che copra tutte le stagioni, pianificando in anticipo i temi che verranno toccati. In generale il calendario dovrebbe contenere delle informazioni inerenti a quella determinata stagione e che possano essere di interesse per il cliente.

Ad esempio, a settembre dovrebbero esserci dei contenuti focalizzati sul ritorno a lavoro e alla quotidianità, come ad esempio "Le migliori borse da ufficio", "Le borse che tutte le tue colleghe invidieranno" e "Come organizzare la tua borsa per una routine senza stress".

Durante le feste natalizie invece i contenuti dovrebbero focalizzarsi su possibili idee regalo o sulle nuove tendenze che verranno lanciati nell'anno successivo e che saranno in voga per tutto l'anno. Esempi di contenuti potrebbero essere “Migliori borse con cui iniziare l'anno”, “Guida ai regali di Natale: Borse per tutti i gusti” oppure ancora “La borsa perfetta per Capodanno: Accessori per brillare”.

Per la primavera e l'estate i numeri dovrebbero riguardare le migliori borse da avere per le mete estive e per le gite fuori porta o suggerire delle borse perfette come regalo per la Festa della Mamma o per altre celebrazioni come feste di laurea e cerimonie. I titoli potrebbero essere “Le migliori borse per l'estate”, “Guida ai regali per la Festa della Mamma” e “Pronti a partire? Scopri le borse che ti accompagneranno nella tua avventura”.

4.4.3 BUNDLE CON PRODOTTI COLLEGATI

Un'altra strategia di marketing che può essere avviata per aumentare le vendite del sito e-commerce sfruttando la classificazione per occasione d'uso, è un'iniziativa incentrata sui bundle. In generale un bundle è un'offerta che combina due o più prodotti in un unico pacchetto, che è venduto ad un prezzo minore rispetto alla somma dei due prezzi singoli degli oggetti.

Questa tipologia di promozione è molto utilizzata nei siti e-commerce per incentivare i clienti a comprare più prodotti, aumentando così la percezione di convenienza da parte dell'utente e incrementando la spesa media per singolo cliente.

I bundle possono essere di varie tipologie, tra le più utilizzate sono quelli stagionali, tematici o personalizzati, ma sfruttando la classificazione di borse per occasione d'uso, le aziende che utilizzano questa tipologia di classificazione, dovrebbero proporre dei bundle per occasione d'uso.

È importante sottolineare che questa strategia incentrata sui bundle, ha molto effetto e può portare a dei concreti miglioramenti in termini di vendite e di spesa media per consumatore, se l'azienda nel sito e-commerce ha una varietà di prodotti e non vende solo borse. Nel caso, infatti, che essa vendesse anche altri prodotti come accessori, vestiti, occhiali da sole, ed altro, la strategia potrebbe portare ad un forte incremento delle vendite perché riuscirebbe a convincere il consumatore a comprare più prodotti, anche se di categorie diverse, e quindi di venire a conoscenza di molti items venduti nel sito.

Per costruire un bundle che abbia successo e che sia attrattivo per i clienti, secondo la ricerca condotta da Yadav M.S. intitolata "Bundle Evaluation in Different Market Segments: The Effects of Discount Framing and Buyers' Preference Heterogeneity"⁹⁹, bisogna considerare prima di tutto il fatto che ci sono livelli di interesse diversi per ogni singolo oggetto componente il bundle. Se viene offerto un bundle con una borsa e un altro accessorio per alcuni consumatori sarà più interessante acquistare la borsa mentre per altri l'altro accessorio.

Collegato a questa disparità di interesse per gli items che compongono il bundle, il modo in cui vengono applicati gli sconti influenza in modo significativo l'attrattività del bundle stesso. Infatti, i risultati di questa ricerca, dimostrano che i consumatori tendono a valutare in maniera più positiva e più attrattiva i bundle che presentano lo sconto sul prodotto che preferiscono maggiormente.

A livello pratico, dato che il bundle verrà offerto quando l'utente si interessa ad un'occasione d'uso per la ricerca di una borsa, lo sconto maggiore dovrà essere applicato sulla borsa piuttosto che

⁹⁹ Yadav, M. S. (1995). Bundle Evaluation in Different Market Segments: The Effects of Discount Framing and Buyers' Preference Heterogeneity. *Journal of the Academy of Marketing Science*, 23(3), 206–215. <https://doi.org/10.1177/0092070395233005>

sull'altro accessorio. In questo modo, l'utente troverà molto attrattiva l'offerta e sarà più invogliato a completare l'acquisto del bundle.

Per le borse utilizzate per il lavoro e ufficio, l'azienda potrebbe proporre dei bundle che comprendano degli accessori o degli oggetti complementari. Come, ad esempio, con una custodia per laptop o un portadocumenti in modo da aggiungere valore all'acquisto. Addirittura, per le aziende con una grossa notorietà e grandezza, si potrebbero proporre dei bundle con altre aziende che si occupano di oggettistica come penne di grande valore, agende molto eleganti e con aziende tecnologiche.

In questo modo, oltre ad aumentare l'appetibilità del bundle, l'azienda che ha il sito e-commerce migliorerebbe ancor di più la sua immagine per le prestigiose collaborazioni oltre che il suo rapporto con le aziende con il quale collabora.

Per quanto riguarda le borse per viaggi e avventure, l'azienda potrebbe offrire dei bundle non solo per far percepire la convenienza ma anche per rafforzare la percezione del marchio come fornitore di tutto il necessario per un viaggio che combina la comodità allo stile.

Potrebbero essere offerti in un pacchetto le borse da viaggio con gli accessori, in modo da rispondere alle esigenze degli utenti. Oppure potrebbero offrire nel pacchetto degli oggetti utili durante il viaggio come un cuscino da viaggio e una maschera per dormire.

Questa strategia potrebbe attrarre molte altre aziende interessate a collaborare con il sito e-commerce e quindi con l'azienda, che aumenterebbe l'appetibilità dei bundle e il proprio fatturato.

Per le borse destinate al tempo libero e alle attività quotidiane, l'azienda potrebbe creare bundle che includano accessori utili per migliorare la praticità e lo stile nella vita di tutti i giorni. Questi bundle potrebbero essere progettati proprio per adattarsi ad una varietà di attività quotidiane come lo shopping, la palestra, le uscite con amici e altre situazioni di svago.

I bundle potrebbero essere composti dalla borsa con l'aggiunta di un ombrello, con una scontistica sui libri pattuita con delle case editrici, oppure con degli auricolari. Questi sono solo dei possibili esempi, ma è importante che il pacchetto sia composto, oltre che dalla borsa chiaramente, da oggetti che sono molto utili durante le giornate e che ci accompagnano per tutte le ore del giorno fuori casa.

Parlando di borse per eventi ed occasioni speciali, i bundle potrebbero essere fatti grazie a collaborazioni con aziende di lusso o con degli eventi o catene di ristoranti sparse nel territorio dal quale si sta cercando. È importante che emerga un senso di eleganza e di esclusività dalla combinazione della borsa con l'altro oggetto, che sia materiale o che sia un buono o un biglietto.

Infine, per borse per gite fuori porta, l'azienda potrebbe proporre dei pacchetti con accessori ed oggetti pratici e funzionali per chi ama fare dei viaggi di breve durata e stare all'aria aperta.

I bundle potrebbero essere composti, grazie a delle collaborazioni con aziende esterne, di borracce termiche, molto utili per qualsiasi tipo di gita fuori porta e anche riutilizzabile in altri contesti. Potrebbe anche comprendere crema solare, o una coperta da picnic. La questione che accomuna questi oggetti e che deve emergere è la forte utilità di essi in qualsiasi occasione vengano usate le borse per gite fuori porta, in modo da invogliare ancor di più il consumatore a procedere con l'acquisto.

4.5 CONTROLLO SULL'EFFICACIA DELLE INIZIATIVE MARKETING

Per calcolare l'impatto effettivo delle campagne marketing personalizzate possono essere utilizzati vari KPI¹⁰⁰, utili anche per misurare l'effettiva efficacia delle iniziative e per capire se stanno avendo l'effetto desiderato o meno.

Per calcolare i risultati delle campagne di marketing possono essere utilizzati i seguenti indicatori:

Engagement Rate: viene utilizzata per misurare il livello di coinvolgimento medio degli utenti su un sito web in rapporto al numero di visitatori unici (ogni visitatore viene conteggiato una sola volta, indipendentemente da quante volte abbia visitato la piattaforma)¹⁰¹.

$$\text{Engagement Rate} = \frac{\text{Numero totale di interazioni (click, tempo speso, ecc.)}}{\text{Numero totale di visitatori unici}} \times 100$$

Un risultato alto di questo KPI indica che in media gli utenti interagiscono molto con il sito, il che suggerisce che il contenuto è molto interessante e coinvolgente. Mentre un risultato scarso sottolinea che, anche se il sito attira molti visitatori, pochi di loro interagiscono con il contenuto e quindi è sintomo di poco interesse e poco coinvolgimento.

Conversion Rate: misura la percentuale di utente presenti nel sito che completano l'azione di acquisto. Questo KPI è fondamentale per valutare l'effettiva efficacia dell'iniziativa di marketing, poiché riflette la capacità di quest'ultima nel trasformare i visitatori in clienti attivi¹⁰².

$$\text{Conversion Rate} = \frac{\text{Numero totale di conversioni}}{\text{Numero totale di visitatori o lead}} \times 100$$

Un alto tasso di conversione dimostra che il sito è ottimizzato per spingere l'utente a comprare. Mentre un basso tasso di conversione può essere sintomo di vari problemi come l'usabilità del sito,

¹⁰⁰ Key Performance Indicators sono delle metriche utilizzate per misurare l'efficacia di una determinata iniziativa implementata nell'azienda. Essa può riguardare l'area vendite, marketing, soddisfazione del cliente.

¹⁰¹ Niciporuc, T. (2014, December 1). Comparative analysis of the engagement rate on Facebook and Google Plus social networks. Retrieved from ideas.repec.org website: <https://ideas.repec.org/p/sek/iacpro/0902287.html>

¹⁰² Di Fatta, D., Patton, D., & Viglia, G. (2018). The determinants of conversion rates in SME e-commerce websites. *Journal of Retailing and Consumer Services*, 41(1), 161–168. <https://doi.org/10.1016/j.jretconser.2017.12.008>

cioè se il processo di registrazione, pagamento, o checkout è complicato o poco intuitivo tanto da spingere l'utente ad abbandonare il sito prima di acquistare l'articolo.

Potrebbe essere anche un problema sulle campagne di marketing che non sono ben indirizzate o abbastanza coinvolgenti da indurre il consumatore ad acquistare gli articoli presenti nel sito.

È molto utile tenere sempre monitorato questo KPI per capire l'efficacia della campagna e della qualità della user experience. Ma soprattutto fornisce degli insight molto utili per migliorare la strategia e le iniziative di marketing.

Customer Retention Rate: è uno dei KPI più utilizzati per misurare la fidelizzazione del cliente. Essa, infatti, riesce a controllare e a dare un valore alla capacità dell'azienda di mantenere i propri clienti nel tempo. Questo indice è molto rilevante per l'azienda perché il costo per acquistare nuovi clienti è maggiore rispetto al costo di trattenere i clienti già esistenti¹⁰³.

$$\mathbf{[(clienti\ fine\ periodo - clienti\ acquisiti\ nel\ periodo) / clienti\ inizio\ periodo] \times 100}$$

Un alto tasso di retention sottolinea che i clienti sono soddisfatti dei prodotti e dell'azienda stessa e che vogliono continuare ad acquistare nel sito. Mentre un basso risultato è sintomo di insoddisfazione del cliente o di una concorrenza molto forte nel quale l'azienda non riesce a differenziarsi e ad emergere.

Churn Rate: collegato al KPI precedente bisogna considerare anche il tasso di abbandono, il cosiddetto "Churn Rate". Esso misura la percentuale di clienti che hanno abbandonato o comunque non acquistano più dall'azienda in un determinato periodo¹⁰⁴.

$$\text{Churn Rate} = \frac{\text{Numero di clienti persi durante il periodo}}{\text{Numero totale di clienti all'inizio del periodo}} \times 100$$

Un churn rate alto comporta per l'azienda una perdita di ricavi, dato che ogni utente rappresenta dei possibili ricavi futuri ed inoltre, un alto tasso di abbandono significa che l'azienda dovrà spendere di più per acquisire nuovi clienti piuttosto che mantenere quelli già presenti.

¹⁰³ Berkeley Electronic Press. (2019). Association for Information Systems - Log in. Retrieved from Aisnet.org website: <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1197&context=jais>

¹⁰⁴ Mutanen, T. (2006). Customer churn analysis -a case study Author Teemu Mutanen. Retrieved from https://publications.vtt.fi/julkaisut/muut/2006/customer_churn_case_study.pdf

Questi sono solo alcuni dei KPI utili per monitorare l'operato e l'efficacia delle iniziative di marketing personalizzate proposte dalle aziende nei propri siti e-commerce.

CAPITOLO 5: CONCLUSIONI E SVILUPPI FUTURI

5.1 CONCLUSIONI

Il settore della moda sta vivendo un periodo di forte trasformazione e stravolgimento di idee e di dogmi sul quale poggia da moltissimi anni. Questa trasformazione si intreccia con le potenti innovazioni tecnologiche, tra le quali l'utilizzo dell'intelligenza artificiale e del crescente ricorso dell'e-commerce.

Questa tesi ha esplorato come tali tecnologie, più in particolare modelli di machine learning, possano essere utilizzate per migliorare la classificazione di prodotti di moda, con focus sulle borse, utilizzando le caratteristiche dimensionali, di materiale, di colore e il prezzo delle stesse. Lo studio è partito da un dataset costituito da una grossa quantità e informazioni riguardanti tutte queste caratteristiche di borse prese da un sito e-commerce nel mercato, e tramite vari modelli di machine learning si è dimostrato come, soprattutto con modelli come il Random Forest e l'XGBoosting, l'accuratezza è molto elevata e precisa.

Lo scopo della tesi non era solo quello di dimostrare come attraverso questi modelli la classificazione di borse utilizzando solo le caratteristiche sopra descritte, sia molto efficace. L'intento ultimo era quello di sfruttare la classificazione per poi segmentare le borse in cluster per occasione d'uso. Lo studio, quindi propone una clusterizzazione di borse secondo un modello euristico che è l'occasione di utilizzo. In un processo decisionale la parte euristica ha una grande importanza nel facilitare la ricerca come viene sottolineato da varie ricerche e soprattutto dalla ricerca svolta da K. Krabuanrat e R. Phelps¹⁰⁵. Questa suddivisione, che include borse da lavoro, da tempo libero, per eventi speciali e altre occasioni, consente una segmentazione più dettagliata del mercato, facilitando al contempo la ricerca del consumatore e il processo decisionale.

Partendo da questa segmentazione si possono creare strategie di marketing più mirate ed efficaci. Attraverso l'analisi delle preferenze dei consumatori e la suddivisione dei prodotti in base a determinati contesti d'uso, le aziende possono sviluppare campagne promozionali più personalizzate, migliorando così il coinvolgimento e la fidelizzazione del cliente. Ad esempio, durante i periodi di transizione stagionale, come il rientro dalle vacanze estive o l'inizio del nuovo anno scolastico, la promozione di borse da lavoro o per ufficio può essere intensificata, sapendo che vi sarà una maggiore domanda per questi prodotti. Allo stesso modo, la classificazione delle borse per eventi speciali permette di creare campagne specifiche per le festività, i matrimoni o altre occasioni formali,

¹⁰⁵ Krabuanrat, K., & Phelps, R. (1998). Heuristics and rationality in strategic decision making: An exploratory study. *Journal of Business Research*, 41(1), 83–93. [https://doi.org/10.1016/s0148-2963\(97\)00014-3](https://doi.org/10.1016/s0148-2963(97)00014-3)

facilitando al consumatore la scelta del prodotto più adatto alla circostanza. La capacità di offrire suggerimenti personalizzati riduce il tempo di ricerca e aumenta le probabilità di conversione, migliorando al contempo la soddisfazione del cliente.

Al contempo sono stati anche proposti dei KPI per misurare l'efficacia e l'efficienza delle campagne di marketing in modo da poterle analizzare e migliorarle nel caso non stiano dando i risultati attesi.

5.2 LIMITI E RICERCHE FUTURE

La ricerca attuale presenta alcuni limiti che offrono l'opportunità per ulteriori ricerche future. Innanzitutto, uno dei principali limiti dello studio è il fatto che il campione considerato è stato preso solo da un catalogo e-commerce di un'unica azienda, anche se sono stati considerati cataloghi di prodotti offerti in nazioni diverse. Perciò, nonostante il numero di dati utilizzati per la ricerca sia molto alto, potrebbe non essere sufficiente per generalizzare i risultati ottenuti dall'attuale studio.

Oltretutto le caratteristiche del dataset e il suo inevitabile sbilanciamento verso le tipologie di borse più richieste sul mercato possono aver compromesso o comunque peggiorato il risultato dei vari addestramenti dei modelli considerati in fase di analisi. Pertanto, ricerche future potrebbero replicare l'analisi su un dataset più bilanciato e che sia composto da modelli di borse presi da vari siti e-commerce. Sarebbe infatti interessante replicare l'esperimento, usando i dati da cataloghi completamente diversi, che presentano stili e clientela molto eterogenea, in modo da avere un risultato che sia il più completo e utile possibile.

In secondo luogo, un ulteriore limite è quello di aver utilizzato solo alcune delle caratteristiche delle borse, infatti oltre a quelle considerate ci sono molti altri attributi che possono essere considerati come la presenza o meno della cerniera, la presenza di tasche interne e altre caratteristiche che potrebbero essere utili a migliorare i risultati della classificazione. Tuttavia, per mitigare questo limite, si è cercato di utilizzare tutti gli attributi disponibili e presenti nella descrizione della borsa, in modo da considerare per lo studio più variabili possibili.

In terzo luogo, si sono valutate solo le caratteristiche delle borse e non si è fatto uso delle immagini di esse presenti nelle pagine e-commerce. L'immagine del prodotto, come sostenuto nei capitoli precedenti, può non essere sempre ideale per una classificazione perché può non rendere chiare alcune caratteristiche, come le dimensioni, o addirittura nascondere degli attributi che per la classificazione potrebbero essere molto importanti. Perciò le ricerche future potrebbero quindi utilizzare sia le caratteristiche delle borse sia l'immagine di esse, in modo da combinare la precisione delle variabili sulle dimensioni e sui materiali con l'efficacia delle fotografie. È necessario precisare però, che le immagini che devono essere utilizzate dovrebbero essere prive di movimento e tutte con la stessa illuminazione, in modo da rendere ancor più precisa la classificazione ed ottenere dei risultati più accurati.

È importante studiare la metodologia più adatta per combinare questi due tipologie di classificazione perché sono molto eterogenee tra loro, il che rischierebbe di, oltre che far diventare il lavoro molto lungo e dispendioso, di non dare i frutti attesi. Si consiglia quindi di cercare delle ricerche che utilizzano la stessa metodologia ma che classificano diversi oggetti, dato che ricerche di

classificazione di borse tramite il machine learning combinando le immagini e le caratteristiche descrittive delle borse non sono ancora state fatte.

Infine, sono stati utilizzati i modelli più famosi di machine learning come il Random Forest, il Gradient Boosting e il Naive Bayes. Perciò, nonostante questi metodi citati e gli altri modelli utilizzati siano molto validi e abbiano dato, come analizzato precedentemente, dei buoni risultati, sono stati testati solo degli algoritmi di classificazione molto conosciuti e utilizzati nelle altre ricerche scientifiche. Pertanto, ricerche future potrebbero utilizzare dei modelli più complessi e più adatti alla tipologia di dataset utilizzata, in modo da allenare più algoritmi di classificazione possibili per trovare il modello che meglio performa per questa tipologia di dati.

BIBLIOGRAFIA

Alessia Perissinotto. (2023). E-commerce 2024: Previsioni e Strategie d'Investimento | Algoritma. Retrieved July 17, 2024, from Algoritma website:

<https://www.algoritma.it/e%E2%80%91commerce-2024-previsioni-e-strategie-dinvestimento/>

Ari, N., & Ustazhanov, M. (2014). Matplotlib in python.

<https://doi.org/10.1109/ICECCO.2014.6997585>

Bellini, P., Nesi, P., Palesi, A. L. I., & Pantaleo, G. (2021). Fashion retail recommendation system by multiple clustering. In Proceedings of the 27th International DMS Conference on Visualization and Visual Languages (DMSVIVA 2021), Pittsburgh, PA, USA (pp. 29-30).

Berkeley Electronic Press. (2019). Association for Information Systems - Log in. Retrieved from Aisnet.org website: <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1197&context=jais>

Blázquez, M. (2014). Fashion shopping in multichannel retail: The role of technology in enhancing the customer experience. *International Journal of Electronic Commerce*, 18(4), 97–116

Braga, D. J., SS, L. B., de Souza, C. P., & da Silva, T. L. C. (2019). Handbags classification model via deep learning. In International Conference on Machine Learning and Data Mining (pp. 1-11).

Bressert, E. (2012). SciPy and NumPy: an overview for developers.

Breward, C. (2014). *Fashion*. Oxford University Press, Oxford History of Art.

[Link to the book](#)

Camurati, F. (2024). Inditex vola nel 2023, ricavi e utili corrono a doppia cifra. Zara a +10% - MilanoFinanza News. Retrieved from MF Milano Finanza website: <https://www.milanofinanza.it/fashion/inditex-vola-nel-2023-ricavi-e-utili-corrono-a-doppia-cifra-zara-a-10-202403131106285838>

Changjo Yoo, Jonghee Park, Deborah J. MacInnis, Effects of Store Characteristics and In-Store Emotional Experiences on Store Attitude, *Journal of Business Research*, Volume 42, Issue 3, 1998, Pages 253-263

- Chen, S. (Joseph), Tran, K. T., Xia, Z. (Raymond), Waseem, D., Zhang, J. A., & Potdar, B. (2023). The double-edged effects of data privacy practices on customer responses. *International Journal of Information Management*, 69(Volume 69), 102600. <https://doi.org/10.1016/j.ijinfomgt.2022.102600>
- Chen, T.-Y., Wu, H.-L., Yun, J.-H., Wu, H.-L., & Hua, J. (2024). Electronic Word-of-mouth. *Asia Marketing Journal*, 26(2). <https://doi.org/10.53728/2765-6500.1629>
- Coppola, D. (2023). U.S. e-commerce market size 2016-2022 | Statista. Statista; Statista. <https://www.statista.com/statistics/272391/us-retail-e-commerce-sales-forecast/>
- Corner, F. (n.d.). Perché la moda è importante. Google Arts & Culture. https://artsandculture.google.com/story/_QKS0J-OeT7HIA?hl=it.
- Corner, Frances. “Perché La Moda è Importante.” Google Arts & Culture, artsandculture.google.com/story/_QKS0J-OeT7HIA?hl=it. Accessed 17 July 2024.
- Deng, Y., Gao, Q. (2020). RETRACTED ARTICLE: A study on e-commerce customer segmentation management based on improved K-means algorithm. *Inf Syst E-Bus Manage* 18, 497–510. <https://doi.org/10.1007/s10257-018-0381-3>
- Di Fatta, D., Patton, D., & Viglia, G. (2018). The determinants of conversion rates in SME e-commerce websites. *Journal of Retailing and Consumer Services*, 41(1), 161–168. <https://doi.org/10.1016/j.jretconser.2017.12.008>
- Doran, George T. (1981). There’s a S.M.A.R.T. way to write management’s goals and objectives. *Management review* 70.11 1981: 35. Business Source Premier
- Editorial. (2022). Pros and cons of Random Forest Algorithm. Retrieved from RoboticsBiz website: <https://roboticsbiz.com/pros-and-cons-of-random-forest-algorithm/>
- Finanza, MF Milano. (2024). “Nike, Ricavi a 12,6 Miliardi (-2%), Confermando Le Previsioni a Ribasso - MilanoFinanza News.” MF Milano Finanza, 28 June 2024, www.milanofinanza.it/fashion/nike-ricavi-a-12-6-miliardi-2-confermando-le-previsioni-a-ribasso-202406281128592617.
- Fratini, T. (2015). Sul dominio della moda e i suoi effetti sulla persona nell’epoca odierna: considerazioni pedagogiche. *Studi Sulla Formazione/Open Journal of Education*, 18(1), 147–159.
- GeeksForGeeks. (2020). ML - Gradient Boosting. Retrieved from GeeksforGeeks website: <https://www.geeksforgeeks.org/ml-gradient-boosting/>

Ghansawant, S., Kumbhar, G., Kawase, N., Deepak, K., & Sharma. (2023). The Power of Personalization: Enhancing User Experience in E-commerce. Retrieved from <https://www.ijfmr.com/papers/2023/6/8899.pdf>

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, 62(1), 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>

Goutte, C., Gaussier, E. (2005). A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. In: Losada, D.E., Fernández-Luna, J.M. (eds) *Advances in Information Retrieval. ECIR 2005. Lecture Notes in Computer Science*, vol 3408. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-31865-1_25

Guercini, S., Bernal, P. M., & Prentice, C. (2018). New marketing in fashion e-commerce. *Journal of Global Fashion Marketing*, 9(1), 1–8. <https://doi.org/10.1080/20932685.2018.1407018>

Hachcham, A. (2023). XGBoost: Everything You Need to Know. Retrieved from Neptune.ai website: <https://neptune.ai/blog/xgboost-everything-you-need-to-know>

Hao, J., & Ho, T. K. (2019). Machine Learning Made Easy: A Review of Scikit-learn Package in Python Programming Language. *Journal of Educational and Behavioral Statistics*, 44(3), 348-361. <https://doi.org/10.3102/1076998619832248>

Harreis, H., Koullias, T., Roberts, R., & Te, K. (2023). Generative AI in Fashion | McKinsey. Retrieved from www.mckinsey.com website: <https://www.mckinsey.com/industries/retail/our-insights/generative-ai-unlocking-the-future-of-fashion>

IBM. (2023). What is Naïve Bayes | IBM. website: <https://www.ibm.com/topics/naive-bayes>

“In Crescita Del 13,5% Anche Febbraio-Marzo 2023: Inditex: Nel 2022 I Ricavi Salgono Del 17,5% a 32,6 Miliardi.” *Www.fashionmagazine.it*, 15 Mar. 2023, www.fashionmagazine.it/business/vendite-in-crescita-del-135-tra-febbraio-e-marzo-inditex-nel-2022-i-ricavi-salgono-del-175-a-326-miliardi-111260. Accessed 17 July 2024.

Indeemo. (2024). The Importance of Occasions in Consumer Research. Retrieved from Indeemo website: <https://indeemo.com/blog/consumer-research-occasions>

“Inditex: 36 Miliardi Di Ricavi Nel 2023, Dividendo +28%.” *Www.teleborza.it*, 13 Mar. 2024, finanza.lastampa.it/News/2024/03/13/inditex-36-miliardi-di-ricavi-nel-2023-dividendo-+28percento/MjdfMjAyNC0wMy0xMjM19UTEI#:~:text=Inditex%2C%20la%20societ%C3%A0%20spagnola%20dell. Accessed 21 July 2024.

Jin, Y. (2023). Analyze the Effect of Customer Retention Rate on Customer Lifetime Value(CLV) - Based on the Regression Model. <https://doi.org/10.4108/eai.18-11-2022.2326838>

Kelly Di Blas. (2024). Rapporto 2023 sull'innovazione digitale in Italia: alcuni progressi inaspettati - BuoneNotizie.it. Retrieved from BuoneNotizie.it website: <https://www.buonenotizie.it/innovazione/2024/02/26/rapporto-2023-sullinnovazione-digitale-in-italia-alcuni-progressi-inaspettati/kelly-di>

Korniichuk, R., & Boryczka, M. (2021). Conversion Rate Prediction Based on Text Readability Analysis of Landing Pages. *Entropy*, 23(11), 1388. <https://doi.org/10.3390/e23111388>

Knight, W. (2017). Amazon Has Developed an AI Fashion Designer. Retrieved from MIT Technology Review website: <https://www.technologyreview.com/2017/08/24/149518/amazon-has-developed-an-ai-fashion-designer/>

Krabuanrat, K., & Phelps, R. (1998). Heuristics and rationality in strategic decision making: An exploratory study. *Journal of Business Research*, 41(1), 83–93. [https://doi.org/10.1016/s0148-2963\(97\)00014-3](https://doi.org/10.1016/s0148-2963(97)00014-3)

Lao, B., & Jagadeesh, K. (2015). Convolutional neural networks for fashion classification and object detection. *CCCV 2015 Comput. Vis*, 546, 120-129.

Loretelli, V. (2023). Paloma Elsesser: “Ho sempre amato la moda, ma non avevo il coraggio di entrare nei negozi”. *Harper's Bazaar Italia*. <https://www.harpersbazaar.com/it/moda/a46187930/paloma-elsesser/>.

Loretelli, V. (2023). Paloma Elsesser: la bio e la carriera. Retrieved from Harper’s BAZAAR website: <https://www.harpersbazaar.com/it/moda/a46187930/paloma-elsesser/>

Marketing personalizzato – un approccio alla personalizzazione della comunicazione con i clienti - Copymate. (2024). Retrieved from Copymate website: <https://copymate.app/it/blog/multi/marketing-personalizzato-un-approccio-alla-personalizzazione-della-comunicazione-con-i-clienti/#:~:text=Il%20marketing%20personalizzato%20pu%C3%B2%20aumentare>

May Britt Bjerke, Ralph Renger (2017), Being smart about writing SMART objectives, *Evaluation and Program Planning*, Volume 61, Pages 125-127, ISSN 0149-7189, <https://doi.org/10.1016/j.evalprogplan.2016.12.009>.

McKinney, W. (2011). pandas: a foundational Python library for data analysis and statistics. *Python for high performance and scientific computing*, 14(9), 1-9.

- Minini, A. (n.d.). La storia del linguaggio Python. website: <https://www.andreaminini.com/python/la-storia-del-linguaggio-python>
- Morelli, A. P. (2022). I DATI PERSONALI: IL NUOVO PETROLIO DEGLI HACKER. Retrieved July 23, 2024, from Mondo 27001 website: <https://www.mondo27001.it/i-dati-personali-il-nuovo-petrolio-degli-hacker/#:~:text=Nel%202006%2C%20Clive%20Humby%5B2>
- Mutanen, T. (2006). Customer churn analysis -a case study Author Teemu Mutanen. Retrieved from https://publications.vtt.fi/julkaisut/muut/2006/customer_churn_case_study.pdf
- Naresh K. Malhotra. (1982). Information Load and Consumer Decision Making, *Journal of Consumer Research*, Volume 8, Issue 4, Pages 419–430, <https://doi.org/10.1086/208882>
- Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*, 7, 21.
- Niciporuc, T. (2014). Comparative analysis of the engagement rate on Facebook and Google Plus social networks. Retrieved from ideas.repec.org website: <https://ideas.repec.org/p/sek/iacpro/0902287.html>
- Parmar, A., Katariya, R., Patel, V. (2019). A Review on Random Forest: An Ensemble Classifier. In: Hemanth, J., Fernando, X., Lafata, P., Baig, Z. (eds) *International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018*. ICICI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 26. Springer, Cham. https://doi.org/10.1007/978-3-030-03146-6_86
- Politecnico Milano. (2020). Linguaggio Python | Caratteristiche e Utilizzo di questo linguaggio. Retrieved from TechCamp@POLIMI website: <https://techcamp.polimi.it/linguaggio-python/#:~:text=Uno%20dei%20motivi%20del%20suo>
- Puspitasari, N. B., Pramono, S. W., Rinawati, D. I., & Fidiyanti, F. (2020). Online consumer segmentation study based on factors affecting e-commerce selection. In *IOP Conference Series: Materials Science and Engineering* (Vol. 722, No. 1, p. 012036). IOP Publishing.
- Rachid, A. D., Abdellah, A., Belaid, B., & Rachid, L. (2018). Clustering Prediction Techniques in Defining and Predicting Customers Defection: The Case of E-Commerce Context. *International Journal of Electrical and Computer Engineering (IJECE)*, 8(4), 2367. <https://doi.org/10.11591/ijece.v8i4.pp2367-2383>

- Roncato, S. (2024). La virata androgina di Gucci. Milano Finanza. Recuperato da <https://www.milanofinanza.it/fashion/la-virata-androgina-di-gucci-1954811>.
- Rapporto McKinsey sulla moda: previsioni sul 2023. Assomac. <https://assomac.it/it/news/news-dal-mondo/rapporto-mckinsey-sulla-moda-previsioni-sul-2023/>
- Redazione AGI. (2023). Il settore moda chiuderà il 2023 al +4%, a settembre contrazione. AGI. https://www.ansa.it/sito/notizie/economia/2023/12/13/il-settore-moda-chiudera-il-2023-al-4-a-settembre-contrazione_51c143d3-c9a5-4a62-809c-37a86c8e7fe5.html.
- Sabanoglu, T. (2023). Personal luxury goods market value by market segment worldwide 2020. Retrieved from Statista website: <https://www.statista.com/statistics/883902/personal-luxury-goods-market-value-by-product-category-worldwide/>
- Şen, A. (2008). The US fashion industry: A supply chain review. *International Journal of Production Economics*, 114(2), 571–593. <https://doi.org/10.1016/j.ijpe.2007.05.022>
- Shen, D., Ruvini, J. D., & Sarwar, B. (2012). Large-scale item categorization for e-commerce. In *Proceedings of the 21st ACM international conference on Information and knowledge management* (pp. 595-604).
- Smith, P. (2023). Revenue of the Global Apparel Market 2014-2027. Retrieved from Statista website: <https://www.statista.com/forecasts/821415/value-of-the-global-apparel-market>
- Statista. “Brand Value of the Leading 10 Apparel Brands Worldwide 2019 | Statista.” Statista, Statista, 2019, www.statista.com/statistics/267931/brand-value-of-the-leading-10-apparel-brands-worldwide/.
- Swaminathan, K., & Venkitasubramony, R. (2023). Demand forecasting for fashion products: A systematic review. *International Journal of Forecasting*, 40(1). <https://doi.org/10.1016/j.ijforecast.2023.02.005>
- Toolify AI. (2023). Unlocking fashion's future: AI vs. human creativity. Toolify AI. <https://www.toolify.ai/ai-news/unlocking-fashions-future-ai-vs-human-creativity-3079643>
- Tu, Y., Tang, X., & Ling, K. (2021). Using machine learning to forecast fashion trends from social media images. *Fashion and Textiles*, 40(4), 507-522.
- Van Gelder, K. (2024). U.S. online apparel and fashion market size 2022. Retrieved from Statista website: <https://www.statista.com/statistics/278890/us-apparel-and-accessories-retail-e-commerce-revenue/>

- Vinerean, S., & Opreana, A. (2021). Measuring Customer Engagement in Social Media Marketing: A Higher-Order Model. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(7), 2633–2654. MDPI. <https://doi.org/10.3390/jtaer16070145>
- Wikipedia Contributors. (2021). Python (programming language). Retrieved from Wikipedia website: https://en.wikipedia.org/wiki/Python_%28programming_language%29
- Wu, K., & Chi, K. (2024). Enhanced E-commerce Customer Engagement: A Comprehensive Three-Tiered Recommendation System. *Journal of Knowledge Learning and Science Technology* ISSN: 2959-6386 (Online), 2(3), 348–359. <https://doi.org/10.60087/jklst.vol2.n2.p359>
- Xia, Q., Lv, J., Ma, S., Gao, B., & Wang, Z. (2020). A New Information-Theoretic Method for Advertisement Conversion Rate Prediction for Large-Scale Sparse Data Based on Deep Learning. *Entropy*, 22(6), 643. <https://doi.org/10.3390/e22060643>
- Yadav, M. S. (1995). Bundle Evaluation in Different Market Segments: The Effects of Discount Framing and Buyers' Preference Heterogeneity. *Journal of the Academy of Marketing Science*, 23(3), 206–215. <https://doi.org/10.1177/0092070395233005>
- Yue, S., Li, P., & Hao, P. (2003). SVM classification: Its contents and challenges. *Applied Mathematics-A Journal of Chinese Universities*, 18(3), 332–342. <https://doi.org/10.1007/s11766-003-0059-5>
- Zhang, Yanbo, and Chuanlan Liu. (2024). "Unlocking the Potential of Artificial Intelligence in Fashion Design and E-Commerce Applications: The Case of Midjourney" *Journal of Theoretical and Applied Electronic Commerce Research* 19, no. 1: 654-670. <https://doi.org/10.3390/jtaer19010035>