AI, BigData & Advanced Analytics in FashionTech.

# 1. Modulo 1: Introduzione.

* **Приветствие и представление:** Кратко расскажите о целях воркшопа.
* **Обзор темы:** Объясните важность Big Data и AI Analytics в индустрии моды.
* **Цели воркшопа:** Что студенты узнают и смогут сделать к концу занятия.
* Структура вокршопа: ссылки на остальные модули.

### 

## 1.2. L'importanza di Big Data e AI Analytics nella moda.

1. **L'industria della moda e i dati.**Oggi la moda non è solo creatività e design, ma anche dati. Ogni interazione online, ogni acquisto, ogni post sui social genera dati preziosi che possono essere analizzati per prendere decisioni strategiche.
2. **Comprendere il comportamento dei consumatori.**Big Data permette di monitorare preferenze, abitudini di acquisto e tendenze in tempo reale. Per esempio, i dati raccolti dai social media possono aiutare i brand a prevedere quali capi saranno più popolari nella prossima stagione.
3. **Ottimizzazione della supply chain.**Con l'analisi dei dati, i brand possono ridurre gli sprechi, ottimizzare i tempi di produzione e distribuzione, e rispondere più rapidamente alla domanda del mercato.
4. **Personalizzazione dell'esperienza cliente.**Grazie all'AI, i brand possono creare esperienze su misura per ogni cliente. Sistemi di raccomandazione personalizzati aumentano la fidelizzazione e migliorano le vendite.
5. **Anticipare le tendenze.**L'AI analizza milioni di dati per identificare pattern e prevedere le tendenze future, aiutando i designer e i manager a prendere decisioni informate.
6. **Competitività sul mercato.**In un settore così competitivo, chi usa Big Data e AI ha un vantaggio strategico. Può prendere decisioni basate su dati concreti, riducendo i rischi e aumentando le opportunità.
7. **Sostenibilità.**Analizzare i dati aiuta anche a promuovere pratiche più sostenibili, ad esempio riducendo la sovrapproduzione e migliorando l'efficienza nell'uso delle risorse.

# 

# 2. Modulo 2: Основные Macro-Use-Case сценарии в Fashion Tech.

* **Пояснение технологий:**
  + **Big Data инструменты:** Hadoop, Spark, SQL-базы данных.
  + **AI и аналитические модели:** Готовые модели машинного обучения для классификации, регрессии и анализа текста.
* **Презентация списка Use Case:**
  + **Анализ тенденций моды:** Прогнозирование модных трендов на основе данных социальных сетей и поисковых запросов.
  + **Управление цепочкой поставок:** Оптимизация логистики и инвентаризации с помощью анализа больших данных.
  + **Персонализация предложений:** Рекомендательные системы для персонализированного шопинга.
  + **Sentiment Analysis бренда:** Анализ отзывов и упоминаний бренда в социальных сетях.
  + **Прогнозирование продаж:** Использование исторических данных для прогнозирования будущих продаж
* **Разделение ролей Big Data и ML:**
  + **Big Data:** Сбор, хранение и предварительная обработка данных.
  + **Модели ML:** Применение готовых моделей для извлечения инсайтов из данных.

## 

## **2.1. Termini Pratici.**

### Definizioni Chiave.

* **Machine Learning vs. Deep Learning vs. Neural Network:**
  + **Machine Learning:** Campo dell'AI che sviluppa algoritmi per permettere ai computer di apprendere dai dati.
  + **Deep Learning:** Sottoinsieme del machine learning che utilizza reti neurali profonde per modellare dati complessi.
  + **Neural Network:** Struttura di calcolo ispirata al cervello umano, utilizzata nel deep learning per riconoscere pattern complessi.
* **Machine Learning Model, Training e Inference:**
  + **Machine Learning Model:** Rappresentazione matematica che fa previsioni o decisioni basate sui dati.
  + **Training:** Processo di insegnare al modello attraverso l'esposizione a dati.
  + **Inference:** Utilizzo del modello addestrato per fare previsioni su nuovi dati.
* **JSON, Hadoop, Spark, SQL:**
  + **JSON:** Formato per lo scambio di dati tra applicazioni.
  + **Hadoop:** Framework per l'elaborazione distribuita di grandi quantità di dati.
  + **Spark:** Framework per l'elaborazione rapida dei Big Data, supporta l'elaborazione in-memory.
  + **SQL:** Linguaggio per gestire e interrogare database relazionali.

### Esempi Pratici.

* **Clustering:** Utilizzato per segmentare i clienti in gruppi basati sui loro comportamenti di acquisto, permettendo strategie di marketing mirate.
* **A/B Testing:** Testare due diverse versioni di una campagna pubblicitaria per determinare quale genera più conversioni.
* **Natural Language Processing (NLP):** Analizzare i commenti sui social media per estrarre opinioni e sentimenti riguardo a un brand di moda.
* **ARIMA:** Prevedere le vendite future basate su dati storici delle vendite mensili.
* **Filtraggio Collaborativo:** Raccomandare nuovi prodotti ai clienti basandosi sugli acquisti di utenti con interessi simili.
* **Sentiment Analysis con ChatGPT:** Utilizzando ChatGPT per analizzare i commenti sui social media e determinare il sentiment verso un nuovo brand di moda, identificando feedback positivi e negativi.
* **JSON:** I dati degli ordini online possono essere trasferiti in formato JSON tra il front-end del sito web e il back-end del sistema di gestione degli ordini per un'elaborazione efficiente.

### Termini Aggiuntivi Utili.

* **Data Lake:** Un repository centralizzato che consente di immagazzinare tutti i dati strutturati e non strutturati a qualsiasi scala.
* **Data Warehouse:** Un sistema utilizzato per la reportistica e l'analisi dei dati, strutturato per facilitare l'accesso e l'analisi.
* **ETL (Extract, Transform, Load):** Processo di estrazione dei dati da diverse fonti, trasformazione in un formato utilizzabile e caricamento in un data warehouse.
* **API (Application Programming Interface):** Interfaccia che permette a diverse applicazioni di interagire tra loro, ad esempio, recuperare dati da un social network per analisi.
* **Dashboard:** Strumento di visualizzazione che mostra metriche chiave e indicatori di performance in modo interattivo e facilmente comprensibile.

## 

## **2.2. Macro-Use-Case** in **Fashion Tech.**

Queste descrizioni dettagliate dei **Macro-Use-Case** in **Fashion Tech** forniscono una panoramica completa e strutturata, facilitando la comprensione per studenti e futuri manager della moda. Ogni Use Case è progettato per evidenziare come Big Data e AI possano trasformare i vari aspetti dell'industria della moda, migliorando l'efficienza, la precisione delle decisioni e la soddisfazione dei clienti.

### 1. Analisi delle Tendenze di Moda

#### **1. Nome e Descrizione**

* **Nome Use Case:** Previsione dei Trend di Moda
* **Descrizione:** Utilizzo di dati provenienti dai social media e dalle ricerche online per identificare e prevedere i trend emergenti nel settore della moda, permettendo alle aziende di adattare le loro collezioni in anticipo.

#### **2. Stakeholder Coinvolti**

* **Roli che partecipano:**
  + Marketing
  + Design
  + Analisi Dati
  + Ricerca e Sviluppo
* **Partecipanti principali:**
  + Team di Social Media Marketing
  + Designer di moda
  + Analisti di dati
  + Consulenti esterni specializzati in trend analysis

#### **3. Processi Aziendali Coinvolti e Modificati**

* **Processi coperti:**
  + Ricerca e sviluppo di nuovi prodotti
  + Pianificazione delle collezioni stagionali
  + Marketing e promozione
* **Cambiamenti post-implementazione:**
  + Maggiore tempestività nell'adattamento delle collezioni ai trend
  + Miglior coordinamento tra i team di design e marketing
  + Decisioni basate su dati reali anziché su intuizioni

#### **4. Problemi nei Processi Tradizionali**

* **Problemi principali:**
  + Lentezza nel rilevare e reagire ai cambiamenti nei trend
  + Affidamento eccessivo su intuizioni soggettive dei designer
  + Mancanza di dati concreti per supportare le decisioni
* **Conseguenze negative:**
  + Ritardi nel lancio delle collezioni
  + Riduzione della competitività sul mercato
  + Aumento del rischio di insuccesso commerciale delle nuove linee

#### **5. Soluzioni con Big Data e AI**

* **Come le tecnologie aiutano:**
  + Raccolta e analisi automatizzata di grandi volumi di dati dai social media e dalle ricerche online
  + Identificazione di pattern e trend emergenti attraverso l'analisi predittiva
* **Esempi di miglioramenti:**
  + Previsione accurata dei trend stagionali
  + Adattamento rapido delle strategie di design e marketing
  + Miglioramento della soddisfazione del cliente grazie a collezioni più allineate ai desideri del mercato

#### **6. Tecnologie, Strumenti e Approcci**

* **Strumenti principali:**
  + Python (librerie come Pandas, BeautifulSoup)
  + API di Google Trends
  + Strumenti di NLP (Natural Language Processing) come NLTK o SpaCy
  + Tableau per la visualizzazione dei dati
* **Approcci:**
  + Analisi delle tendenze tramite clustering e analisi delle frequenze
  + Analisi del sentiment per comprendere le preferenze dei consumatori
  + Machine Learning per la previsione dei trend futuri

#### **7. Difficoltà e Risorse Richieste**

* **Livello di complessità:** Medio
* **Competenze necessarie:**
  + Analisi dei dati
  + Conoscenza di strumenti di programmazione (es. Python)
  + Comprensione delle dinamiche di mercato della moda
* **Composizione del team:**
  + Analista dei dati
  + Data Scientist
  + Responsabile Marketing
  + Designer di moda
* **Durata del progetto:** 3-6 mesi

#### **8. Fattori Critici di Successo**

* **Elementi chiave:**
  + Accesso a fonti di dati aggiornate e rilevanti
  + Collaborazione efficace tra i vari dipartimenti aziendali
  + Qualità e accuratezza dei modelli predittivi utilizzati
* **Necessità per il successo:**
  + Investimento in tecnologie di analisi dei dati
  + Formazione continua del personale coinvolto
  + Supporto e impegno della leadership aziendale

### 2. Gestione della Catena di Approvvigionamento

#### **1. Nome e Descrizione**

* **Nome Use Case:** Ottimizzazione della Logistica e dell'Inventario
* **Descrizione:** Utilizzo di Big Data e analisi avanzate per migliorare l'efficienza della catena di approvvigionamento, riducendo i costi logistici e ottimizzando i livelli di inventario.

#### **2. Stakeholder Coinvolti**

* **Roli che partecipano:**
  + Logistica
  + Acquisti
  + IT
  + Finanza
* **Partecipanti principali:**
  + Manager della supply chain
  + Responsabili degli acquisti
  + Analisti logistici
  + Fornitori e partner logistici

#### **3. Processi Aziendali Coinvolti e Modificati**

* **Processi coperti:**
  + Pianificazione degli acquisti
  + Gestione dell'inventario
  + Distribuzione e logistica
* **Cambiamenti post-implementazione:**
  + Miglioramento della previsione della domanda
  + Riduzione dei tempi di consegna
  + Ottimizzazione dei livelli di stock, evitando eccessi e carenze

#### **4. Problemi nei Processi Tradizionali**

* **Problemi principali:**
  + Previsioni di domanda inaccurate
  + Elevati costi di magazzinaggio
  + Ritardi nelle consegne e inefficienze logistiche
* **Conseguenze negative:**
  + Perdita di vendite dovuta a stockouts
  + Aumento dei costi operativi
  + Riduzione della soddisfazione del cliente

#### **5. Soluzioni con Big Data e AI**

* **Come le tecnologie aiutano:**
  + Analisi predittiva per migliorare la precisione delle previsioni di domanda
  + Ottimizzazione delle rotte di distribuzione tramite algoritmi di ottimizzazione
  + Monitoraggio in tempo reale dell'inventario e delle spedizioni
* **Esempi di miglioramenti:**
  + Riduzione dei costi di magazzino del 20%
  + Miglioramento dei tempi di consegna del 15%
  + Aumento della disponibilità dei prodotti più richiesti

#### **6. Tecnologie, Strumenti e Approcci**

* **Strumenti principali:**
  + Hadoop e Spark per la gestione dei Big Data
  + Python e R per l'analisi dei dati
  + Sistemi ERP integrati (es. SAP, Oracle)
  + Algoritmi di Machine Learning per la previsione e l'ottimizzazione
* **Approcci:**
  + Analisi dei pattern di domanda storica
  + Ottimizzazione delle reti logistiche
  + Automazione dei processi di riordino

#### **7. Difficoltà e Risorse Richieste**

* **Livello di complessità:** Alto
* **Competenze necessarie:**
  + Ingegneria dei dati
  + Analisi predittiva e ottimizzazione
  + Conoscenza dei sistemi ERP
* **Composizione del team:**
  + Data Engineer
  + Data Scientist
  + Responsabile Logistica
  + Specialista ERP
* **Durata del progetto:** 6-12 mesi

#### **8. Fattori Critici di Successo**

* **Elementi chiave:**
  + Integrazione efficace dei sistemi IT esistenti
  + Precisione e qualità dei dati raccolti
  + Collaborazione stretta tra i team logistici e di analisi dati
* **Necessità per il successo:**
  + Investimenti in tecnologie avanzate
  + Formazione del personale sui nuovi strumenti e processi
  + Supporto continuo della direzione aziendale

### 3. Personalizzazione delle Offerte

#### **1. Nome e Descrizione**

* **Nome Use Case:** Sistemi di Raccomandazione per lo Shopping Personalizzato
* **Descrizione:** Implementazione di sistemi di raccomandazione basati su AI per offrire ai clienti suggerimenti personalizzati di prodotti, migliorando l'esperienza di acquisto e aumentando le vendite.

#### **2. Stakeholder Coinvolti**

* **Roli che partecipano:**
  + Marketing
  + Vendite
  + IT
  + Customer Service
* **Partecipanti principali:**
  + Team di e-commerce
  + Responsabili del marketing digitale
  + Data Scientist
  + Esperti di UX/UI

#### **3. Processi Aziendali Coinvolti e Modificati**

* **Processi coperti:**
  + Esperienza utente sul sito e-commerce
  + Strategie di marketing e promozione
  + Gestione delle relazioni con i clienti
* **Cambiamenti post-implementazione:**
  + Aumento delle conversioni e delle vendite
  + Miglioramento della soddisfazione e fidelizzazione dei clienti
  + Personalizzazione delle campagne di marketing

#### **4. Problemi nei Processi Tradizionali**

* **Problemi principali:**
  + Offerte generiche che non rispondono alle esigenze individuali dei clienti
  + Bassa conversione delle campagne di marketing
  + Difficoltà nel mantenere alto il livello di soddisfazione e fedeltà del cliente
* **Conseguenze negative:**
  + Riduzione delle vendite e del valore medio dell'ordine
  + Alta percentuale di abbandono del carrello
  + Scarsa differenziazione rispetto ai concorrenti

#### **5. Soluzioni con Big Data e AI**

* **Come le tecnologie aiutano:**
  + Analisi dei comportamenti di acquisto e delle preferenze dei clienti
  + Generazione di raccomandazioni personalizzate in tempo reale
  + Segmentazione avanzata del mercato per campagne mirate
* **Esempi di miglioramenti:**
  + Incremento del 25% nelle vendite cross-sell e up-sell
  + Aumento del 30% nella fidelizzazione dei clienti
  + Miglioramento della soddisfazione del cliente grazie a esperienze di acquisto personalizzate

#### **6. Tecnologie, Strumenti e Approcci**

* **Strumenti principali:**
  + Piattaforme di raccomandazione (es. Amazon Personalize, Google Recommendations AI)
  + Python e librerie di Machine Learning (es. TensorFlow, Scikit-learn)
  + Sistemi di gestione dei dati clienti (CRM) integrati
  + Strumenti di analisi dei dati (es. Tableau, Power BI)
* **Approcci:**
  + Filtraggio collaborativo e basato sul contenuto
  + Analisi predittiva dei comportamenti di acquisto
  + A/B testing per ottimizzare le raccomandazioni

#### **7. Difficoltà e Risorse Richieste**

* **Livello di complessità:** Medio
* **Competenze necessarie:**
  + Data Analysis
  + Machine Learning
  + UX/UI Design
  + Marketing digitale
* **Composizione del team:**
  + Data Scientist
  + Sviluppatore Front-end
  + Responsabile Marketing
  + Designer UX/UI
* **Durata del progetto:** 4-8 mesi

#### **8. Fattori Critici di Successo**

* **Elementi chiave:**
  + Accuratezza e rilevanza delle raccomandazioni
  + Integrazione fluida con le piattaforme di e-commerce
  + Adattabilità del sistema alle evoluzioni delle preferenze dei clienti
* **Necessità per il successo:**
  + Accesso a dati di alta qualità e aggiornati
  + Test continui e iterazioni basate sul feedback dei clienti
  + Collaborazione tra team tecnici e di marketing

### 4. Sentiment Analysis del Brand

#### **1. Nome e Descrizione**

* **Nome Use Case:** Analisi delle Opinioni e delle Menzioni del Brand
* **Descrizione:** Utilizzo di tecniche di sentiment analysis per monitorare e analizzare le opinioni dei clienti riguardo al brand sui social media e altre piattaforme online, al fine di migliorare la reputazione e le strategie di marketing.

#### **2. Stakeholder Coinvolti**

* **Roli che partecipano:**
  + Marketing
  + Customer Service
  + Comunicazione
  + IT
* **Partecipanti principali:**
  + Social Media Manager
  + Responsabile della Comunicazione
  + Analista di dati
  + Team di Customer Service

#### **3. Processi Aziendali Coinvolti e Modificati**

* **Processi coperti:**
  + Gestione della reputazione online
  + Strategie di comunicazione e marketing
  + Servizio clienti e gestione delle crisi
* **Cambiamenti post-implementazione:**
  + Miglioramento della reattività alle opinioni dei clienti
  + Personalizzazione delle campagne di comunicazione
  + Proattività nella gestione delle crisi di reputazione

#### **4. Problemi nei Processi Tradizionali**

* **Problemi principali:**
  + Difficoltà nel monitorare manualmente tutte le menzioni del brand
  + Ritardi nella risposta alle opinioni negative
  + Mancanza di insights reali sulle percezioni dei clienti
* **Conseguenze negative:**
  + Deterioramento della reputazione del brand
  + Perdita di clienti a causa di risposte inadeguate
  + Opportunità mancate per migliorare i prodotti e i servizi

#### **5. Soluzioni con Big Data e AI**

* **Come le tecnologie aiutano:**
  + Raccolta automatizzata di dati da diverse piattaforme online
  + Analisi del sentiment per identificare opinioni positive, negative e neutre
  + Monitoraggio in tempo reale delle menzioni del brand
* **Esempi di miglioramenti:**
  + Riduzione del tempo di risposta alle opinioni negative del 50%
  + Aumento della soddisfazione del cliente grazie a interventi tempestivi
  + Miglioramento delle strategie di marketing basate sui feedback reali

#### **6. Tecnologie, Strumenti e Approcci**

* **Strumenti principali:**
  + Strumenti di social listening (es. Brandwatch, Hootsuite Insights)
  + Python e librerie di NLP (es. NLTK, SpaCy)
  + API di sentiment analysis (es. IBM Watson, Google Cloud Natural Language)
  + Dashboard di visualizzazione dati (es. Tableau, Power BI)
* **Approcci:**
  + Analisi del linguaggio naturale per comprendere il contesto delle menzioni
  + Classificazione del sentiment mediante modelli di Machine Learning
  + Integrazione con sistemi CRM per azioni automatizzate

#### **7. Difficoltà e Risorse Richieste**

* **Livello di complessità:** Medio
* **Competenze necessarie:**
  + Natural Language Processing
  + Data Analysis
  + Social Media Management
  + Comunicazione aziendale
* **Composizione del team:**
  + Data Scientist
  + Social Media Manager
  + Analista di Comunicazione
  + Sviluppatore IT
* **Durata del progetto:** 3-6 mesi

#### **8. Fattori Critici di Successo**

* **Elementi chiave:**
  + Accuratezza nell'analisi del sentiment
  + Copertura completa delle piattaforme di menzione del brand
  + Capacità di reagire tempestivamente ai feedback
* **Necessità per il successo:**
  + Accesso a strumenti avanzati di social listening
  + Formazione continua del personale coinvolto
  + Collaborazione tra team di marketing e customer service

### 5. Previsione delle Vendite

#### **1. Nome e Descrizione**

* **Nome Use Case:** Previsione delle Vendite Future
* **Descrizione:** Utilizzo di dati storici e tecniche di analisi avanzate per prevedere le vendite future, permettendo una pianificazione più accurata della produzione, del marketing e della gestione dell'inventario.

#### **2. Stakeholder Coinvolti**

* **Roli che partecipano:**
  + Vendite
  + Marketing
  + Finanza
  + Produzione
* **Partecipanti principali:**
  + Responsabile delle vendite
  + Analisti finanziari
  + Team di produzione
  + Responsabili del marketing

#### **3. Processi Aziendali Coinvolti e Modificati**

* **Processi coperti:**
  + Pianificazione delle vendite e del budget
  + Gestione dell'inventario
  + Pianificazione della produzione
* **Cambiamenti post-implementazione:**
  + Miglioramento della precisione delle previsioni di vendita
  + Ottimizzazione dei livelli di inventario
  + Pianificazione più efficiente della produzione in base alla domanda prevista

#### **4. Problemi nei Processi Tradizionali**

* **Problemi principali:**
  + Previsioni di vendita basate su dati storici limitati o intuizioni soggettive
  + Sovrapproduzione o sottoproduzione di prodotti
  + Difficoltà nel pianificare campagne di marketing efficaci
* **Conseguenze negative:**
  + Costi elevati dovuti a stock in eccesso o carenze di inventario
  + Perdite di vendite a causa di stockouts
  + Inefficienze nella gestione delle risorse produttive

#### **5. Soluzioni con Big Data e AI**

* **Come le tecnologie aiutano:**
  + Analisi dei dati storici di vendita e dei trend di mercato
  + Utilizzo di modelli predittivi per stimare la domanda futura
  + Integrazione con dati esterni come eventi stagionali e promozioni
* **Esempi di miglioramenti:**
  + Aumento della precisione delle previsioni di vendita del 20%
  + Riduzione dei costi di magazzino del 15%
  + Miglioramento della pianificazione delle campagne di marketing mirate

#### **6. Tecnologie, Strumenti e Approcci**

* **Strumenti principali:**
  + Python e R per l'analisi statistica
  + Librerie di Machine Learning (es. Scikit-learn, TensorFlow)
  + Sistemi ERP e CRM integrati
  + Strumenti di visualizzazione dei dati (es. Tableau, Power BI)
* **Approcci:**
  + Modelli di regressione lineare e multivariata
  + Serie temporali e ARIMA
  + Modelli di machine learning avanzati come le reti neurali

#### **7. Difficoltà e Risorse Richieste**

* **Livello di complessità:** Alto
* **Competenze necessarie:**
  + Analisi statistica e data science
  + Conoscenza dei sistemi ERP e CRM
  + Competenze di programmazione in Python o R
* **Composizione del team:**
  + Data Scientist
  + Analista Finanziario
  + Responsabile Vendite
  + Sviluppatore ERP
* **Durata del progetto:** 4-8 mesi

#### **8. Fattori Critici di Successo**

* **Elementi chiave:**
  + Qualità e completezza dei dati storici
  + Accuratezza dei modelli predittivi utilizzati
  + Collaborazione tra i dipartimenti di vendite, marketing e finanza
* **Necessità per il successo:**
  + Investimento in tecnologie di analisi dei dati
  + Formazione del personale sulle tecniche di previsione
  + Supporto continuo della direzione aziendale

### 2.3. Ruoli di Big Data e Modelli di Machine Learning.

#### **2.3.1. Big Data: Raccolta, Archiviazione e Pre-Elaborazione dei Dati.**

* **Ruolo:** Fornire una base solida e ben strutturata di dati per l'analisi e il machine learning.
* **Attività principali:**
  + Raccolta di dati da diverse fonti (social media, e-commerce, sensori IoT, etc.).
  + Archiviazione in Data Lake o Data Warehouse.
  + Pre-elaborazione per pulizia, normalizzazione e trasformazione dei dati.
* **Strumenti principali:**
  + **Hadoop:** Per la gestione distribuita di dati su larga scala.
  + **Spark:** Per l'elaborazione rapida e in-memory.
  + **SQL:** Per interrogare e gestire database relazionali.
  + **ETL:** Per estrazione, trasformazione e caricamento dei dati.

#### **2.3.2. Modelli di Machine Learning: Estrarre Insight dai Dati**

* **Ruolo:** Utilizzare algoritmi per identificare pattern e fare previsioni basate sui dati elaborati.
* **Attività principali:**
  + Addestramento di modelli su dati storici per creare modelli predittivi.
  + Inferenza: Applicazione di modelli per ottenere insight su nuovi dati.
  + Automazione decisionale attraverso algoritmi di classificazione, regressione e clustering.
* **Esempi di Modelli Utilizzati nei Macro-Use-Case:**
  + **Classificazione:** Analisi del sentiment per determinare l’opinione dei clienti.
  + **Regressione:** Previsione delle vendite future basate su trend passati.
  + **Clustering:** Segmentazione dei clienti per strategie di marketing personalizzate.
* **Strumenti principali:**
  + **Python (Scikit-learn, TensorFlow, PyTorch):** Per la costruzione e l'addestramento dei modelli.
  + **API di Machine Learning predefinite (es. Google Cloud AI, IBM Watson):** Per accelerare lo sviluppo con modelli già pronti.
  + **Framework di Visualizzazione (Tableau, Power BI):** Per rappresentare i risultati in modo comprensibile.

#### **2.3.3. Sinergia tra Big Data e Machine Learning**

* **Come interagiscono:**
  + Big Data raccoglie e prepara grandi volumi di dati che i modelli di machine learning utilizzano per identificare pattern e insight.
  + Machine Learning restituisce risultati che possono essere integrati nei sistemi di Big Data per migliorare continuamente la raccolta e l'elaborazione.
* **Obiettivo comune:** Creare un ciclo continuo di miglioramento basato sui dati per decisioni aziendali più informate ed efficienti.

#### 2.3.4. Fattori Critici di Successo

1. **Qualità dei dati:** Dati accurati e ben preparati migliorano significativamente le prestazioni dei modelli.
2. **Scalabilità:** Infrastrutture che supportano grandi volumi di dati e l’elaborazione rapida.
3. **Collaborazione tra team:** Comunicazione fluida tra data engineer, data scientist e decision-maker aziendali.

# 3. Modulo 3: Dimostrazione Pratica del Caso d'Uso Selezionato.

## 3.1. Descrizione dell’obiettivo.

* **Obiettivo:** Analizzare i sentimenti degli utenti sui social media relativi a un brand specifico.
* **Valore:** Comprendere l’opinione pubblica per migliorare la strategia di marketing e gestire la reputazione online.

## 3.2. Fasi del processo.

### 3.2.1. Raccolta dei dati (esempio con Twitter API):

* **Descrizione:** Dimostrazione del processo di raccolta dei dati da Twitter utilizzando l'API con la libreria tweepy.
* **Ottimizzazione:** Per risparmiare tempo e garantire una dimostrazione efficace, utilizziamo un file JSON precompilato con dati sintetici. Questo file contiene tweet simulati relativi al brand, con campi come testo, data di pubblicazione e ID.
* **Strumenti:**
  + Python.
  + Libreria tweepy (esempio di utilizzo API Twitter).
* **Output:** File JSON (tweets\_data.json) contenente tweet grezzi, pronto per la fase di pre-elaborazione.

### 3.2.2. Pre-elaborazione dei dati

* **Descrizione:** Pulizia e preparazione dei dati per l’analisi:
  + Rimozione di link, hashtag, emoji e simboli speciali.
  + Tokenizzazione e rimozione delle stop word.
* **Strumenti:**
  + Python: librerie pandas, re, nltk.
* **Output:** Testo strutturato e pronto per l’analisi.

## 3.3. Analisi del Sentiment: due approcci.

### 3.1. Utilizzo diretto di ChatGPT (senza API).

**Descrizione:**

* Inserire manualmente i testi pre-elaborati nella console di ChatGPT.
* Richiedere una classificazione dei sentimenti per ciascun testo (positivo, negativo, neutro).
* Ideale per dimostrare un metodo semplice e accessibile, senza configurazioni tecniche.

**Strumenti:**

* Console ChatGPT (interfaccia web).

**Procedura:**

1. Copiare i testi pre-elaborati dal file 03\_tweets\_dati\_cleaned.json.
2. Incollare uno o più testi nella console di ChatGPT.
3. Utilizzare un prompt simile per richiedere l'analisi.
4. Annotare i risultati forniti da ChatGPT.

**Output:**

* Classificazione manuale dei sentimenti (da annotare o salvare separatamente).

### 3.2. Utilizzo di ChatGPT tramite API.

**Descrizione:**

* Automatizzare l'analisi inviando i testi elaborati all'API di ChatGPT e ricevendo le classificazioni in formato strutturato.
* Dimostra un approccio scalabile e programmatico per analisi di grandi volumi di dati.

**Strumenti:**

* Python.
* API ChatGPT (chiave di accesso necessaria).

**Procedura:**

1. Leggere i testi pre-elaborati dal file 03\_tweets\_dati\_cleaned.json.
2. Inviare ciascun testo all'API di ChatGPT tramite script Python.
3. Raccogliere le classificazioni dei sentimenti e salvarle in un nuovo file JSON.

**Output:**

* File JSON contenente il testo e il relativo sentimento (positivo, negativo, neutro).

### 3.3. Utilizzo di modelli pre-addestrati (Hugging Face).

* **Descrizione:** Applicare un modello pre-addestrato della libreria transformers per l’analisi del sentiment.
* **Strumenti:**
  + Python: libreria transformers.
* **Output:** Classificazione automatica dei sentimenti.

### 4. Addestramento aggiuntivo del modello.

* **Descrizione:** Mostrare come addestrare ulteriormente un modello utilizzando un dataset etichettato.
* **Dataset:** JSON contenente testo dei tweet e le relative etichette di sentiment (positivo, neutro, negativo).
* **Strumenti:**
  + Python: librerie transformers, datasets.
* **Output:** Modello personalizzato per il sentiment analysis.

### 5. Confronto dei due approcci.

* **Metriche di valutazione:**
  + Accuratezza (Accuracy).
  + Report di classificazione (classification\_report da sklearn).
* **Risultati:**
  + **ChatGPT:** Ideale per analisi rapide e senza necessità di sviluppo.
  + **Hugging Face:** Più flessibile e personalizzabile per esigenze specifiche.

### 6. Output e Risorse Finali.

* **File JSON:**
  + Tweet grezzi per la pre-elaborazione.
  + Dataset etichettato per l’addestramento.
* **Script Python:**
  + Pulizia dei dati.
  + Integrazione con ChatGPT.
  + Utilizzo del modello Hugging Face.
  + Addestramento di un nuovo modello.
* **Presentazione:** Schema chiaro delle fasi con codice esemplificativo.

### 

# 4. Modulo 4: Заключение и обсуждение (10 минут).

* **Выводы:** Подведение итогов воркшопа, ключевые моменты.
* **Обсуждение:** Вопросы от студентов, обсуждение полученного опыта.
* **Дальнейшие шаги:** Рекомендации по самостоятельному изучению, полезные ресурсы.

### 

## Glossario dei Termini.

Questo glossario ampliato copre una vasta gamma di termini tecnologici necessari per comprendere e spiegare i **Macro-Use-Case** in **Fashion Tech**. Ogni definizione è pensata per essere chiara e accessibile a studenti e futuri manager dell'industria della moda, facilitando la comprensione delle tecnologie e degli approcci utilizzati per trasformare il settore attraverso Big Data e AI Analytics.

#### **A**

* **API (Application Programming Interface):** Un insieme di regole e protocolli che permettono a diverse applicazioni di comunicare tra loro. Ad esempio, un'API di Twitter consente di estrarre dati sui tweet per analisi.
* **A/B Testing:** Una metodologia sperimentale utilizzata per confrontare due versioni di una variabile (ad esempio, una pagina web) per determinare quale performa meglio. Nel contesto del marketing, può essere utilizzato per ottimizzare le campagne pubblicitarie.

#### 

#### **B**

* **Big Data:** Grandi volumi di dati strutturati e non strutturati che richiedono tecnologie avanzate per essere raccolti, archiviati e analizzati. Ad esempio, i dati generati dai social media, dagli acquisti online e dai sensori IoT nel settore moda.

#### 

#### **C**

* **ChatGPT:** Un modello di linguaggio sviluppato da OpenAI che utilizza l'intelligenza artificiale per comprendere e generare testo in linguaggio naturale, utile per analisi di sentiment e automazione del servizio clienti.
* **Clustering:** Una tecnica di machine learning non supervisionato che raggruppa un insieme di oggetti in cluster basati su caratteristiche simili. Nel fashion, può essere utilizzato per segmentare i clienti in gruppi con comportamenti di acquisto simili.
* **Content-Based Filtering (Filtraggio Basato sul Contenuto):** Un metodo di raccomandazione che suggerisce prodotti basati sulle caratteristiche dei prodotti stessi e sulle preferenze precedenti dell'utente. Ad esempio, raccomandare abiti simili a quelli che un cliente ha già acquistato.

#### **D**

* **Deep Learning (Apprendimento Profondo)**Un sottoinsieme del Machine Learning che utilizza reti neurali complesse con molti strati (deep neural networks) per modellare e risolvere problemi complessi, come il riconoscimento delle immagini nelle tendenze di moda.

#### 

#### **E**

* **ETL (Extract, Transform, Load):** Un processo utilizzato per estrarre dati da diverse fonti, trasformarli in un formato adatto per l'analisi e caricarli in un data warehouse. Fondamentale per la gestione dei Big Data.

#### 

#### **F**

* **Forecasting (Previsione):** L'uso di dati storici per prevedere eventi futuri. Nel fashion, può essere utilizzato per prevedere le vendite stagionali.

#### 

#### **H**

* **Hadoop**Un framework open-source per l'elaborazione distribuita di grandi set di dati su cluster di computer. Utilizzato per gestire e analizzare Big Data in modo efficiente.

#### **I**

* **Inference (Inferenza):** Il processo di utilizzo di un modello di machine learning addestrato per fare previsioni o prendere decisioni basate su nuovi dati.

#### 

#### **J**

* **JSON (JavaScript Object Notation):** Un formato leggero per lo scambio di dati, facilmente leggibile sia da umani che da macchine. Spesso utilizzato per trasferire dati tra server e applicazioni web.

#### 

#### **M**

* **Machine Learning (Apprendimento Automatico):** Un campo dell'intelligenza artificiale che consente ai computer di imparare dai dati e migliorare le proprie prestazioni nel tempo senza essere esplicitamente programmati. Ad esempio, utilizzare algoritmi di machine learning per raccomandare prodotti ai clienti.
* **Machine Learning Model (Modello di Apprendimento Automatico):** Una rappresentazione matematica creata attraverso l'addestramento su dati che può fare previsioni o decisioni senza essere programmata per compiere specifiche azioni.
* **Machine Learning Pipeline:** Una sequenza di passaggi che include la raccolta dei dati, la pulizia, la trasformazione, l'addestramento del modello e la sua messa in produzione. Essenziale per automatizzare e ottimizzare il processo di machine learning.

#### 

#### **N**

* **Neural Network (Rete Neurale)**Un modello di machine learning ispirato al cervello umano, composto da nodi interconnessi (neuroni) che elaborano i dati in vari strati. Utilizzato nel deep learning per riconoscere pattern complessi.
* **Natural Language Processing (NLP) (Elaborazione del Linguaggio Naturale):** Un campo dell'intelligenza artificiale che si occupa dell'interazione tra computer e linguaggio umano, permettendo ai computer di comprendere, interpretare e generare linguaggio naturale.

#### 

#### **O**

* **Optimization Algorithms (Algoritmi di Ottimizzazione):** Tecniche matematiche utilizzate per trovare la soluzione migliore possibile a un problema, spesso minimizzando o massimizzando una funzione obiettivo. Utilizzati per ottimizzare le rotte di distribuzione o i livelli di inventario.

#### 

#### **P**

* **Predictive Analytics (Analisi Predittiva)**L'uso di dati, algoritmi statistici e tecniche di machine learning per identificare la probabilità di risultati futuri basati su dati storici. Ad esempio, prevedere quali trend di moda saranno popolari nella prossima stagione.

#### **Q:**

* **Query (Interrogazione):** Una richiesta di informazioni da un database. Utilizzata per estrarre dati specifici necessari per l'analisi e la reportistica.

#### **R:**

* **Regression (Regressione):** Un tipo di algoritmo di machine learning utilizzato per prevedere un valore continuo, come le vendite future basate su dati storici.

#### 

#### **S:**

* **SQL (Structured Query Language):** Un linguaggio di programmazione standard per gestire e manipolare database relazionali. Utilizzato per estrarre e analizzare dati strutturati.
* **Spark:** Un framework di elaborazione dei Big Data noto per la sua velocità e facilità d'uso. Supporta l'elaborazione in-memory e viene utilizzato per analisi complesse e machine learning.

#### 

#### **T:**

* **Training (Addestramento):** Il processo di alimentare un modello di machine learning con dati per permettergli di imparare e migliorare le sue previsioni o decisioni.

#### 

#### **U**

* **Unstructured Data (Dati Non Strutturati):** Dati che non seguono uno schema predefinito, come testo libero, immagini e video. Richiedono tecniche avanzate per l'analisi e l'estrazione di informazioni utili.

#### 

#### **V**

* **Visualization (Visualizzazione):** La rappresentazione grafica dei dati per facilitare la comprensione e l'analisi. Strumenti come Tableau o Power BI sono comunemente utilizzati per creare dashboard interattive.

#### **W**

* **Web Scraping:** Tecniche utilizzate per estrarre dati da siti web. Utilizzato per raccogliere grandi quantità di dati dai social media, siti di e-commerce e altre fonti online per l'analisi.

#### **Y**

* **Yield Management:** Strategie utilizzate per ottimizzare la vendita di prodotti o servizi in base alla domanda prevista, massimizzando i ricavi e riducendo gli sprechi.

#### **Z**

* **Zero-Shot Learning:** Una tecnica di machine learning che permette a un modello di fare previsioni su classi di dati che non ha mai visto durante l'addestramento, utile per affrontare problemi con dati limitati.