



Basic Knowlwdge di riferimento in preparazione al Tavolo:

Data science e modelli predittivi per la Customer Experience

Milano, 23 marzo 2025, ore 10-13, in presenza e via web

Indice:

A - DATA SCIENCE: PUNTI-CHIAVE

B - MODELLI PREDITTIVI PER MIGLIORARE LA CX: DIFFICOLTA' E LINEE DI SOLUZIONE

C - RACCOLTA E GESTIONE DEI DATI NELLA CUSTOMER OBSESSION: PRINCIPI CHIAVE

D - COME COSTRUIRE MODELLI PREDITTIVI EFFICACI PER LA CUSTOMER EXPERIENCE

E - DIFFERENZA TRA MODELLI PREDITTIVI E PRESCRITTIVI PER LA CX

A – Data science: punti-chiave

I punti chiave della Data Science si articolano in diverse fasi e componenti fondamentali, che vanno dalla raccolta dei dati fino alla comunicazione degli insight. Ecco gli aspetti principali:

1. Raccolta e Gestione dei Dati

- **Origine dei dati:** Dati strutturati (database relazionali) e non strutturati (testo, immagini, audio, video).
 - **Data Engineering:** Pulizia, trasformazione e integrazione dei dati da più fonti (ETL – Extract, Transform, Load).
 - **Big Data:** Tecnologie come Hadoop, Spark, e database NoSQL per gestire grandi volumi di dati.
-

2. Analisi Esplorativa dei Dati (EDA - Exploratory Data Analysis)

- **Descrivere e visualizzare i dati:** Identificazione di pattern, trend, anomalie.
 - **Statistica descrittiva:** Media, mediana, varianza, distribuzione.
 - **Feature Engineering:** Creazione di nuove variabili per migliorare le prestazioni dei modelli.
-

3. Modellazione e Algoritmi di Machine Learning

- **Modelli supervisionati** (classificazione e regressione): Random Forest, SVM, Reti Neurali, XGBoost.
 - **Modelli non supervisionati** (clustering e riduzione della dimensionalità): K-Means, PCA, DBSCAN.
 - **Deep Learning:** CNN per immagini, RNN per serie temporali, Transformer per NLP.
 - **Validazione dei modelli:** Suddivisione train/test, cross-validation, metriche di performance (MSE, accuracy, precision-recall, AUC-ROC).
-

4. Deployment e Ottimizzazione dei Modelli

- **MLOps:** Implementazione e gestione dei modelli in produzione.
- **Monitoraggio e aggiornamento:** Controllo continuo della performance per evitare degrado del modello.
- **Automazione del flusso di lavoro:** Strumenti come Kubernetes, Docker, CI/CD per gestire pipeline di Data Science.

5. Interpretazione e Comunicazione dei Risultati

- **Data Storytelling:** Creazione di dashboard interattive con strumenti come Tableau, Power BI, Plotly.
- **Comunicazione ai decisori aziendali:** Traduzione di insight tecnici in strategie operative.
- **Etica e Bias nei dati:** Identificazione di pregiudizi nei modelli e creazione di sistemi equi e trasparenti.

6. Applicazioni della Data Science

- **Business Intelligence:** Analisi predittiva per il marketing, vendite, gestione del rischio
- **AI e NLP:** Chatbot, traduzione automatica, sentiment analysis
- **Finanza e Fintech:** Algoritmi di trading, rilevazione di frodi.
- **Industria 4.0:** Manutenzione predittiva, ottimizzazione della supply chain.
- **Sanità:** Diagnosi automatizzata, analisi di immagini mediche.

Punti Chiave della Data Science in Sintesi

- ✓ **Dati di qualità:** Senza dati affidabili, i modelli non funzionano.
- ✓ **Analisi e feature engineering:** Comprendere e preparare i dati è cruciale.
- ✓ **Scelta del modello giusto:** Non esiste un modello perfetto, ma il migliore per il caso specifico.
- ✓ **Monitoraggio e aggiornamento continuo:** I modelli possono deteriorarsi nel tempo.
- ✓ **Comunicazione efficace:** Il valore della Data Science dipende dall'impatto sulle decisioni aziendali.

B – Modelli predittivi per migliorare la CX: difficoltà e linee di soluzione

La costruzione e l'utilizzo di modelli predittivi **per migliorare la customer experience (CX)** presentano diverse difficoltà, che spaziano dalla qualità dei dati alla spiegabilità dei modelli e alla loro integrazione nei processi aziendali.

Ecco le principali sfide:

1. Qualità e Disponibilità dei Dati

- **Dati incompleti o mancanti:** Molti dati sui clienti possono essere errati, mancanti o incoerenti.
- **Dati non strutturati:** Commenti, recensioni e conversazioni con il customer service possono essere difficili da interpretare.
- **Privacy e conformità:** Rispettare le normative (GDPR, CCPA) senza compromettere la personalizzazione.

✓ **Soluzione:** Implementare strategie di **data cleaning**, utilizzare **tecniche di imputazione** per i dati mancanti e garantire la conformità normativa.

2. Complessità nella Modellazione e Feature Engineering

- **Scelta delle variabili** (feature selection): Determinare quali dati influenzano realmente la CX.
- **Overfitting vs. underfitting:** Modelli troppo complessi rischiano di adattarsi troppo ai dati di training, mentre modelli semplici potrebbero non catturare informazioni rilevanti.
- **Dati non bilanciati:** In alcuni settori, il numero di clienti insoddisfatti può essere molto inferiore rispetto a quelli soddisfatti, causando squilibri nei modelli di classificazione.

✓ **Soluzione:** Utilizzare **tecniche di bilanciamento** (es. oversampling/smote per classi minoritarie) e **selezione automatica** delle feature per migliorare l'accuratezza.

3. Spiegabilità e Interpretabilità dei Modelli

- **Black box AI:** Modelli avanzati di deep learning e ensemble methods (XGBoost, Random Forest) possono essere difficili da interpretare.
- **Difficoltà nel convincere il management:** Se il modello non è spiegabile, i decision-maker potrebbero non fidarsi dei risultati.

✓ **Soluzione:** Implementare tecniche di Explainable AI (XAI), come SHAP e LIME, per rendere i modelli più trasparenti e comprensibili.

4. Implementazione e Integrazione nei Processi Aziendali

- **Difficoltà nell'automazione:** Portare i modelli dalla fase sperimentale alla produzione può richiedere infrastrutture robuste (MLOps).
- **Scalabilità:** I modelli devono funzionare in tempo reale per personalizzare l'esperienza del cliente su larga scala.
- **Allineamento con i processi aziendali:** I modelli devono essere integrati nei CRM, chatbot e sistemi di marketing automation.

✓ **Soluzione:** Utilizzare **MLOps** e pipeline automatizzate per l'addestramento e il deployment dei modelli.

5. Adattabilità e Evoluzione del Modello nel Tempo

- **Customer behavior drift:** Le preferenze e il comportamento dei clienti cambiano nel tempo, rendendo obsoleti i modelli.
- **Feedback loop errato:** Se il modello usato per personalizzare le offerte, potrebbe generare un **effetto di auto-conferma**, alterando i dati futuri.

✓ **Soluzione:** Implementare un sistema di monitoraggio continuo della performance del modello e riaddestrarlo periodicamente con nuovi dati.

6. Aspetti Etici e Bias nei Modelli

- **Discriminazione involontaria:** Se il modello è addestrato su dati storici, potrebbe perpetuare pregiudizi (es. offerte diverse in base al genere o all'etnia).
- **Rischio di invasione della privacy:** Un'iper-personalizzazione può risultare inquietante per il cliente (effetto "creepy").

✓ **Soluzione:** Usare **auditing etico** dei modelli e applicare tecniche per la **riduzione dei bias**, come fairness-aware learning.

Riepilogo delle Principali Difficoltà e Soluzioni nella costruzione e utilizzo del Modelli predittivi per la CX

Difficoltà	Soluzione
Dati mancanti o sporchi	Data cleaning, tecniche di imputazione
Feature engineering complessa	Algoritmi di selezione automatica delle variabili
Interpretabilità ridotta	Explainable AI (XAI), SHAP, LIME
Integrazione nei processi aziendali	MLOps, API scalabili
Modelli obsoleti nel tempo	Retraining continuo, monitoraggio drift
Bias nei modelli	Auditing etico, fairness-aware learning

C – Raccolta e gestione dei dati nella Customer Obsession: principi chiave

Per un'azienda **customer-obsessed**, il processo di raccolta e gestione dei dati deve essere **centrato sulle esigenze del cliente**, garantendo personalizzazione, tempestività e valore in ogni interazione. Questo approccio trasforma la gestione dei dati da un'attività tecnica a una leva strategica per migliorare la customer experience (CX).

Ecco i **principi chiave** per caratterizzare questo processo:

1. Customer-Centric Data Collection

La raccolta dei dati non deve essere un'attività passiva, ma mirata a **comprendere il cliente in profondità**.

Dati da raccogliere:

- ✓ **Dati Comportamentali** (click, acquisti, navigazione web, interazioni con il servizio clienti).
- ✓ **Dati Transazionali** (storico acquisti, abbonamenti, feedback, tassi di reso).
- ✓ **Dati Social e Sentiment Analysis** (recensioni, commenti sui social, Net Promoter Score - NPS).
- ✓ **Dati contestuali e di localizzazione** (interazioni su mobile, geolocalizzazione, preferenze di canale).
- ✓ **Dati impliciti e predittivi** (tendenze di utilizzo, personalizzazione in tempo reale).

Esempio: Un e-commerce raccoglie dati di navigazione e abbandono del carrello per proporre offerte personalizzate in tempo reale via email o chatbot.

2. Qualità e Governance dei Dati

Una customer obsession efficace si basa su dati **affidabili, aggiornati e ben strutturati**.

Principi di gestione della qualità:

- ✓ **Single Customer View:** Unificazione delle informazioni su ogni cliente da diversi canali (CRM, e-commerce, social).
- ✓ **Data Cleansing:** Eliminare dati duplicati, errati o obsoleti per garantire accuratezza.
- ✓ **Real-Time Data Processing:** Raccogliere ed elaborare i dati in tempo reale per una CX personalizzata.
- ✓ **Privacy by Design:** Rispettare normative GDPR e CCPA senza compromettere l'esperienza utente.

Esempio: Amazon integra dati di acquisto, preferenze e ricerche per personalizzare raccomandazioni, evitando suggerimenti irrilevanti o ripetitivi.

3. Integrazione e Omnicanalità

Il cliente si muove tra più touchpoint (sito web, mobile, negozio fisico, social, assistenza) e si aspetta **coerenza e continuità**.

Best Practices:

- ✓ **Data Lake & Cloud Computing:** Unificare i dati da fonti diverse per un accesso rapido e scalabile.
- ✓ **API e Sistemi Connessi:** Integrare CRM, ERP, piattaforme di marketing automation per garantire una visione olistica.
- ✓ **Customer Journey Mapping:** Raccogliere dati su ogni fase del percorso cliente per ottimizzare l'esperienza.

Esempio: Starbucks utilizza dati omnicanale per offrire premi e promozioni in base alle abitudini di consumo, sia online che nei punti vendita fisici.

4. Machine Learning e Analisi Predittiva

Non basta raccogliere dati, serve **estrarre insight utili** per anticipare le esigenze del cliente.

Modelli chiave per la customer obsession:

- ✓ **Customer Churn Prediction:** Identificare segnali di possibile abbandono e intervenire in tempo.
- ✓ **Recommendation Systems:** Personalizzare offerte e prodotti sulla base delle preferenze.
- ✓ **Sentiment Analysis:** Analizzare feedback e recensioni per comprendere emozioni e aspettative.
- ✓ **Dynamic Pricing:** Adattare i prezzi in base alla domanda e al comportamento del cliente.

Esempio: Netflix utilizza l'analisi predittiva per suggerire contenuti personalizzati, migliorando il tempo di permanenza sulla piattaforma.

5. Etica, Fiducia e Trasparenza

La customer obsession non deve violare la fiducia del cliente: i dati devono essere usati in modo etico.

Aspetti chiave:

- ✓ **Opt-in e controllo dei dati:** Consentire ai clienti di decidere come i loro dati vengono usati.
- ✓ **Personalizzazione non invasiva:** Evitare il "creepy effect" (es. suggerire prodotti troppo specifici).
- ✓ **Data Security e Cybersecurity:** Proteggere i dati da violazioni e attacchi informatici.

Esempio: Apple enfatizza la privacy dei dati nei suoi prodotti, dando agli utenti il controllo sulle informazioni condivise.

Riepilogo - Customer Obsession nella Raccolta e Gestione dei Dati

Fase	Principi Chiave	Esempio
Raccolta Dati	Dati comportamentali, social, transazionali	Amazon raccoglie dati di navigazione per suggerimenti personalizzati
Qualità e Governance	Single Customer View, Data Cleansing, GDPR Compliance	Netflix pulisce e struttura i dati per migliorare i consigli
Omnicanalità	Integrazione CRM, data lake, API	Starbucks usa i dati per offrire promozioni omnicanale
Analisi Predittiva	Churn Prediction, Sentiment Analysis, Recommendation Systems	Spotify usa AI per creare playlist personalizzate
Etica e Privacy	Trasparenza, Opt-in, Sicurezza	Apple dà all'utente il controllo sui suoi dati

Conclusione

Un'azienda **ossessionata dal cliente** deve trasformare la raccolta e gestione dei dati da un processo tecnico a una leva strategica per **creare esperienze fluide, personalizzate e di valore**.

D – Come costruire modelli predittivi efficaci per la Customer Experience

Una volta superate le difficoltà legate ai dati, alla governance e alla cultura aziendale, la **costruzione di un modello predittivo efficace** richiede un processo strutturato e metodico. Ecco i passi fondamentali per garantire che il modello sia **accurato, interpretabile e utile per migliorare la customer experience**.

1. Definire il Problema e l'Obiettivo del Modello

Domande chiave da porsi:

- ✓ Quale aspetto della CX voglio migliorare? (churn, personalizzazione, lead scoring, sentiment analysis, ecc.)
- ✓ Quale azione dovrà compiere il business grazie al modello?
- ✓ Quali sono le metriche di successo? (Precision, Recall, F1-score, ROC-AUC, RMSE, ecc.)

Esempio di obiettivo ben definito:

*"Voglio costruire un modello per **anticipare il churn** e ridurre del 15% il tasso di abbandono nei prossimi 6 mesi, intervenendo con offerte mirate."*

2. Raccolta e Preparazione dei Dati

Il dataset è la base del modello: senza dati di qualità, anche il miglior algoritmo fallirà.

Dati necessari per la CX

- ✓ **Dati demografici** (età, genere, localizzazione)
- ✓ **Dati comportamentali** (click, acquisti, durata della sessione, azioni su mobile)
- ✓ **Dati transazionali** (importo medio speso, frequenza d'acquisto, storico ordini)
- ✓ **Dati di interazione** (feedback, recensioni, chiamate al servizio clienti)
- ✓ **Dati esterni** (trend di mercato, eventi speciali, social sentiment)

Data Cleaning e Feature Engineering

- ◇ **Eliminare** valori mancanti con imputazione strategica
- ◇ **Normalizzare e scalare** i dati numerici per migliorare l'addestramento
- ◇ **Creare nuove feature** basate su insight di business (es. "tempo dall'ultimo acquisto")
- ◇ **Encoding** per variabili categoriche (es. One-Hot Encoding, Label Encoding)
- ◇ **Strumento utile: Pandas Profiling** per l'analisi esplorativa dei dati

3. Scelta dell'Algoritmo e della Tecnica di Modellazione

Non esiste un modello universale: la scelta dipende dal **problema** e dalla **qualità dei dati**.

Se il problema è...

- ✓ **Classificazione** (es. "il cliente abbandonerà?")
→ **Random Forest, XGBoost, SVM, Reti Neurali**
- ✓ **Regressione** (es. "quanto spenderà un cliente?")
→ **Linear Regression, Gradient Boosting, Deep Learning**
- ✓ **Clustering** (es. "quali segmenti di clienti esistono?")
→ **K-Means, DBSCAN, PCA per riduzione dimensionalità**
- ✓ **Analisi del linguaggio naturale (NLP)** (es. "qual è il sentiment delle recensioni?")
→ **Transformer (BERT, GPT), LSTM, Sentiment Analysis con TF-IDF**

Strumento utile: AutoML per testare più modelli e trovare il migliore

4. Addestramento, Validazione e Ottimizzazione

Dopo aver scelto il modello, bisogna addestrarlo e testarlo in modo rigoroso.

Strategie di validazione

- ✓ **Train/Test Split** (80/20, 70/30, 60/40) per evitare overfitting
- ✓ **Cross-validation K-Fold** per testare il modello su più suddivisioni dei dati
- ✓ **Hyperparameter tuning** con Grid Search o Bayesian Optimization per migliorare le prestazioni

Metriche di valutazione

- **Accuracy** (solo se il dataset è bilanciato)
- **Precision, Recall, F1-score** (per problemi di classificazione)
- **ROC-AUC** (se il problema ha più classi o soglie di decisione)
- **RMSE, MAE** (per regressione)

Strumento utile: SHAP e LIME per interpretare il modello e spiegare le decisioni

5. Deployment del Modello e MLOps

Un buon modello deve funzionare **nel mondo reale**, non solo in fase di test.

Best Practices per il Deployment

- ✓ **Pipeline automatizzata** per aggiornare il modello con nuovi dati
- ✓ **Monitoraggio delle performance** per evitare il "model drift"
- ✓ **Strumenti MLOps (Docker, Kubernetes, MLflow)** per gestire il ciclo di vita del modello

Esempio: Un modello predittivo per il churn che **aggiorna le previsioni ogni notte** e invia notifiche automatiche al CRM per azioni di retention

6. Integrazione nei processi aziendali e Data-Driven Decision Making

Un modello predittivo è inutile se non viene **adottato** dai team aziendali.

Fattori critici di successo

- ✓ **Formazione interna:** I team devono capire e fidarsi del modello
- ✓ **Dashboard interattive per visualizzare le previsioni** (es. con Power BI, Tableau)
- ✓ **Automazione:** Le azioni devono scattare in tempo reale (es. invio di un'email di retention)

Strumento utile: API REST con Flask o FastAPI per integrare il modello nei sistemi aziendali

🔗 Riepilogo: Processo per costruire un Modello Predittivo Efficace

Fase	Attività chiave	Strumenti utili
1. Definizione dell'obiettivo	Churn prediction, Recommendation, Sentiment Analysis	Business Analysis, KPI
2. Raccolta e pulizia dati	Data Cleaning, Feature Engineering	Pandas, Scikit-Learn
3. Scelta del modello	ML Supervisionato, NLP, Deep Learning	XGBoost, TensorFlow, PyTorch
4. Addestramento e validazione	Train/Test Split, Hyperparameter Tuning	Cross-validation, Grid Search
5. Deployment e MLOps	Automazione e monitoraggio	MLflow, Docker, API REST
6. Adozione e business impact	Dashboard interattive, decisioni data-driven	Tableau, Power BI

Conclusione

Creare un modello predittivo per la CX è un processo **complesso ma fondamentale** per un'azienda customer-obsessed.

L'obiettivo non è solo **fare previsioni**, ma **migliorare l'esperienza del cliente in modo concreto e misurabile**.

E – Differenza tra modelli predittivi e prescrittivi per la Customer Experience

Nel contesto della Customer Experience (CX), sia i modelli predittivi che quelli prescrittivi sono strumenti di intelligenza artificiale e analisi dei dati, ma **servono a scopi differenti**.

Ecco la differenza chiave:

◆ **Modelli Predittivi** → "Cosa accadrà?"

- Analizzano i dati storici per **prevedere un evento futuro**.
- Rispondono a domande come:
 - ✓ Quali clienti stanno per abbandonare? (Churn Prediction)
 - ✓ Quanto è probabile che un cliente acquisti un prodotto?
 - ✓ Quale sarà il livello di soddisfazione di un cliente dopo un'interazione?

Esempio pratico: Un e-commerce usa un modello predittivo per identificare clienti che hanno **alta probabilità di abbandono** in base alla loro attività recente (es. diminuzione della frequenza di acquisto).

◆ **Modelli Prescrittivi** → "Cosa dovrei fare?"

- Non solo predicano un evento, ma **forniscono consigli su come agire** per ottenere il miglior risultato possibile.
- Utilizzano ottimizzazione e simulazioni per suggerire la strategia migliore.
- Rispondono a domande come:
 - ✓ Qual è l'azione migliore per trattenere un cliente a rischio di abbandono?
 - ✓ Quale offerta personalizzata ha la maggiore probabilità di conversione?
 - ✓ Qual è il miglior mix di canali per migliorare la soddisfazione del cliente?

Esempio pratico: Dopo aver previsto che un cliente sta per lasciare l'e-commerce, il modello prescrittivo **decide quale strategia** adottare (sconto personalizzato, email di re-engagement, chiamata da un operatore).

Tabella di Confronto

Caratteristica	Modelli Predittivi	Modelli Prescrittivi
Obiettivo	Prevedere eventi futuri	Suggerire azioni ottimali
Esempio in CX	Predire quali clienti abbandoneranno	Decidere la miglior strategia di retention
Strumenti usati	Machine Learning, Regressione, Random Forest	AI + Ottimizzazione, Simulazioni, Reinforcement Learning
Risultato	Probabilità o tendenza futura	Raccomandazione dell'azione migliore
Output tipico	"Il cliente X ha il 80% di probabilità di abbandonare"	"Invia un'offerta speciale al cliente X per trattenerlo"

Quando usare modelli predittivi e prescrittivi per la CX?

✓ **Usa modelli predittivi quando vuoi anticipare problemi** (es. churn, insoddisfazione, domanda di prodotti).

✓ **Usa modelli prescrittivi quando vuoi automatizzare decisioni ottimali** (es. suggerire offerte personalizzate, migliorare il servizio clienti).

Approccio combinato: Un buon sistema di customer experience **usa entrambi i modelli:**

1 Il **Modello predittivo** identifica i clienti a rischio.

2 Il **Modello prescrittivo** decide la strategia migliore per intervenire.