Fondamenti MLOps: parte 4

Concetti dalla scorsa settimana

Temi affrontati la scorsa settimana

1. **Spiegazione comportamento** modelli tramite XAI
2. **Analisi di incertezza** in ambito machine learning
3. **Anomaly detection** per la definizione di limiti operazionali

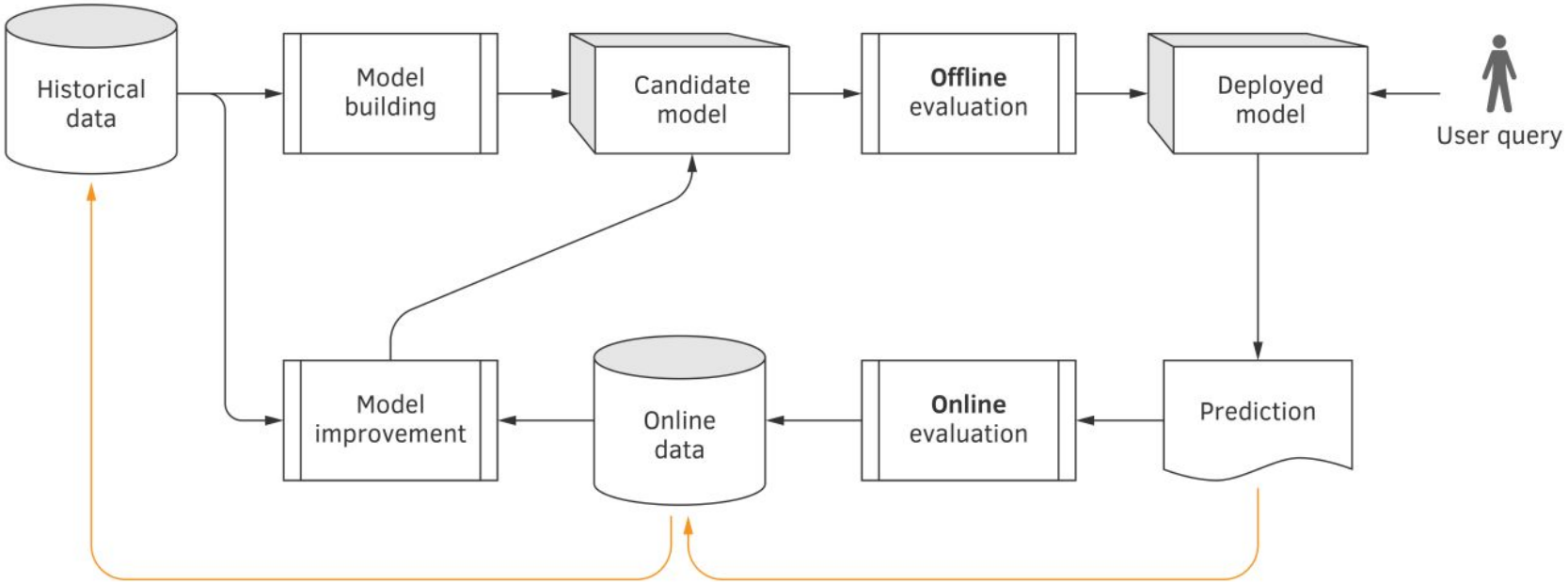
Programma di oggi

Gestione di modelli in produzione

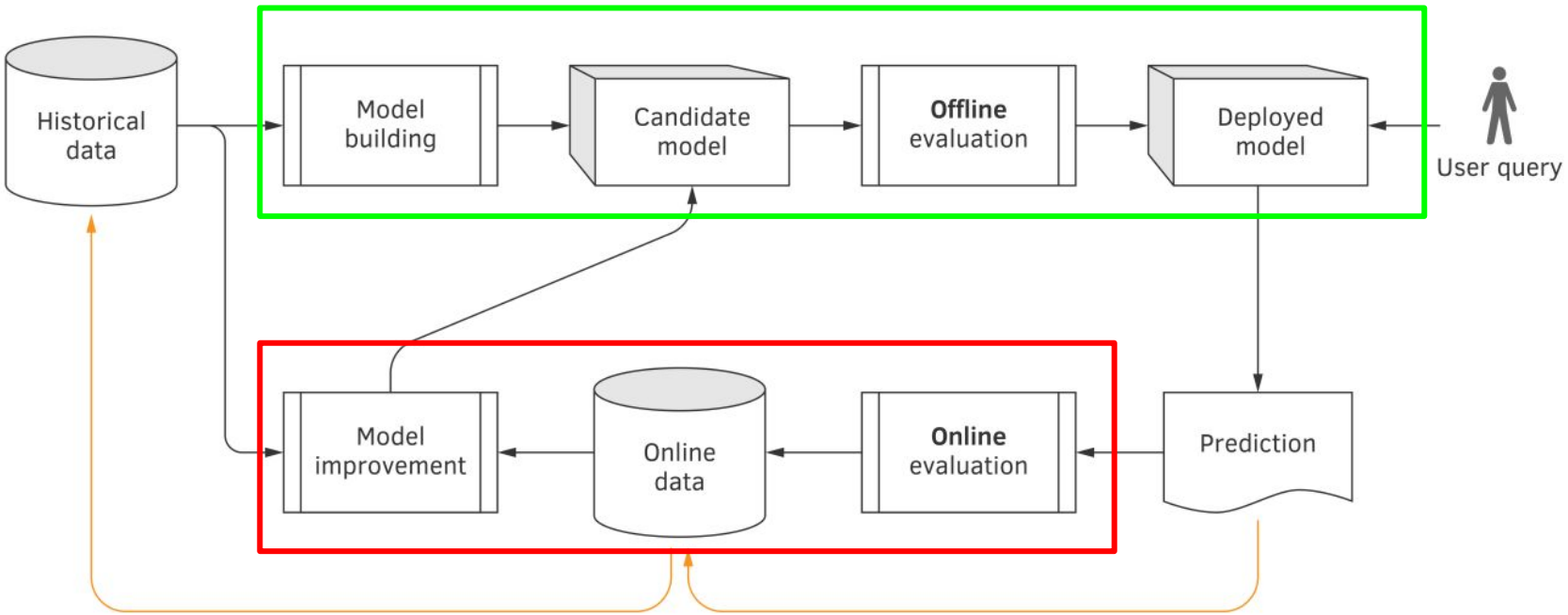
1. **Monitoraggio continuo** (esempio pratico con AWS SageMaker)

2. Miglioramento tramite **active learning**

Monitoraggio e miglioramento continuo



Monitoraggio e miglioramento continuo



Modelli in produzione

Tematiche di produzione affrontate fino ad ora.

Modalità messa in produzione : **Online** o **batch** serving.

Implementazione in cloud con soluzioni **serverless** (Lambda) a **orchestrate** (EC2/ECS).

Esistono anche soluzioni end-to-end dedicate al machine learning.

(**SageMaker**, H2O.ai, DataRobot, etc)

SageMaker

Amazon SageMaker

Servizio completamente dedicato al machine learning.

- Preparazione e processamento dati
- Creazione di feature stores
- Allenamento modelli su Jupyter notebook hostati o con funzioni **autopilot**
- Creazione endpoints per il serving
- Servizi di monitoraggio e etichettatura

SageMaker

Amazon SageMaker

Modelli possono essere allenati e messi in produzione in pochi click su macchine virtuali dedicate.

Contro:

Livello di astrazione introdotta limita il range operazionale dell'utente. In piu' a parita di macchine virtuali utilizzate per allenare e mettere in produzione i modelli il costo e' **30%-40% più alto**.

SageMaker

Esempio notebook SageMaker

Esempio SageMaker

Deploy di un modello su SageMaker tramite mlflow.

Step by step su:

https://github.com/Clearbox-AI/Corso_MLOps/blob/main/Esercizi_sessione2.md

Monitoraggio

Modelli in produzione devono essere tenuti sotto controllo e ri-allenati con una certa periodicità.

Motivo: dati cambiano nel tempo → Modelli rischiano di finire in condizioni operative diverse da quelle in cui sono stati allenati.

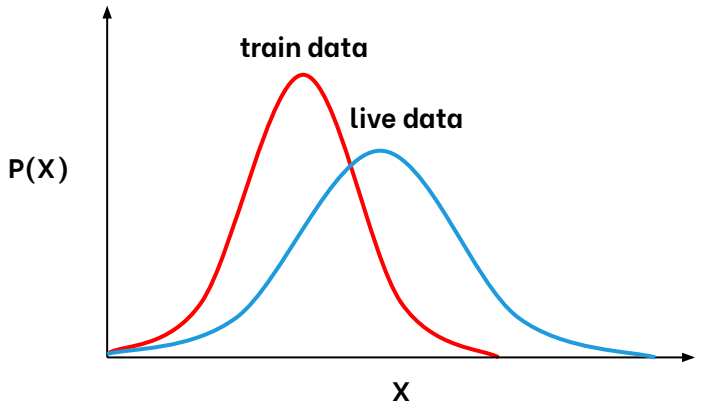
Due principali tipi di fenomeno:

1. **Data** drift
2. **Concept** drift

Data drift

Distribuzioni di probabilità delle features in input cambiano nel tempo.

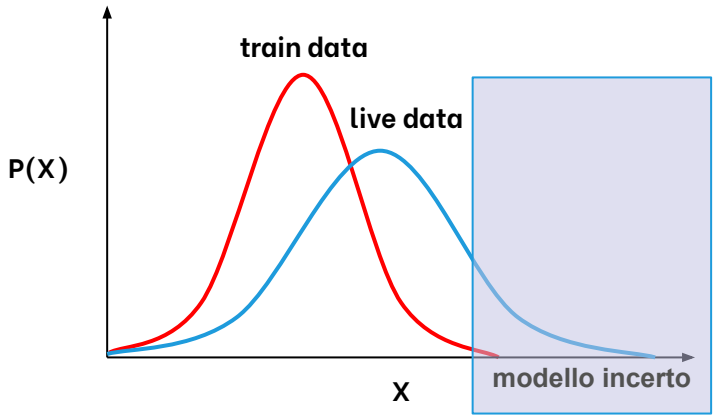
Motivi possono essere molteplici.



Data drift

Distribuzioni di probabilità delle features in input cambiano nel tempo.

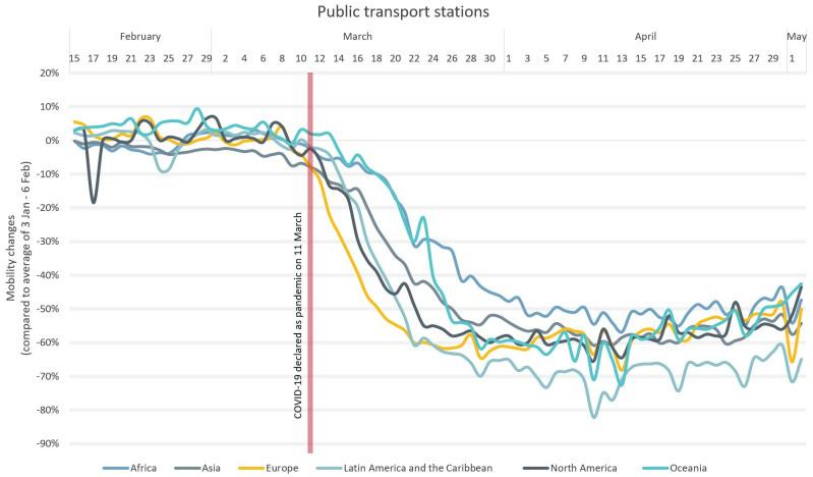
Motivi possono essere molteplici.



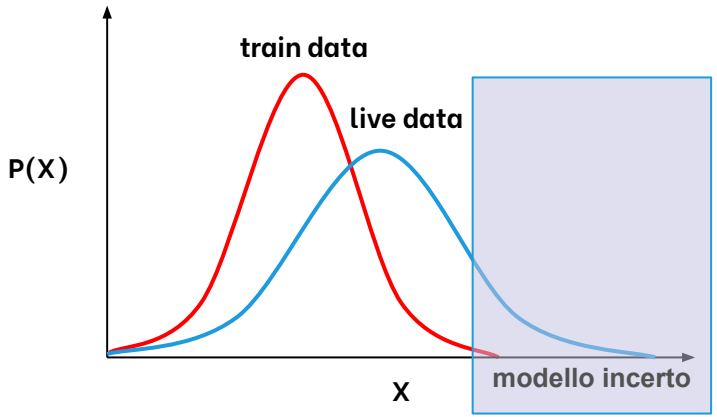
Data drift

Distribuzioni di probabilità delle features in input cambiano nel tempo.

Motivi possono essere molteplici.



Data drift



Distribuzioni di probabilità delle features in input cambiano nel tempo.

Motivi possono essere molteplici.

Statisticamente misurabile usando test come:

- Chi-square
- Maximum Mean Discrepancy (MMD)
- ...

Data drift

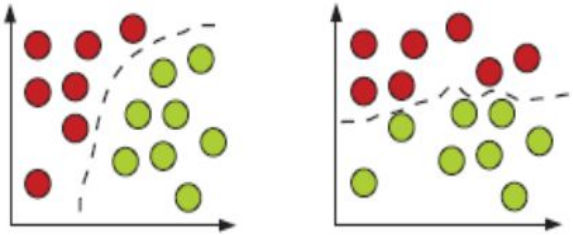
Problema: come misurare drift in combinazioni di variabili o in features nominali?

Failing Loudly: An Empirical Study of Methods for Detecting Dataset Shift, https://arxiv.org/pdf/1810.11953.pdf

Concept drift

Cambiamento nella relazione tra distribuzioni in input e **output** può cambiare considerevolmente nel tempo a seconda del problema in esame.

Distribuzioni nello spazio delle features possono rimanere simili quello che cambia in questo caso e' il decision boundary del problema



HAL Id : hal-02062610, version 1

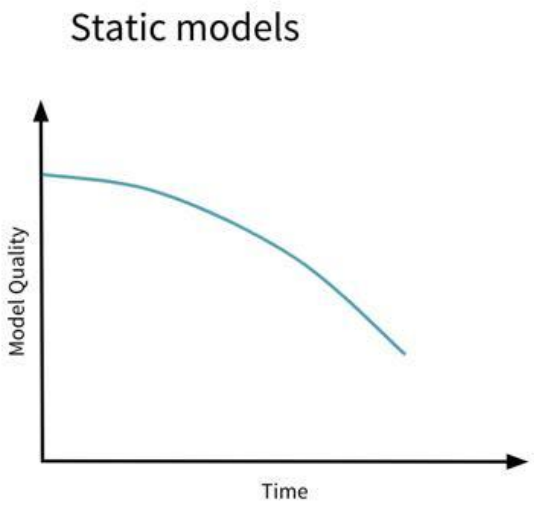
Concept drift

Esempi:

Churn prediction: nuovo provider telefonico low-cost entra nel mercato → churn rate cambia drasticamente a parità di distribuzioni in input

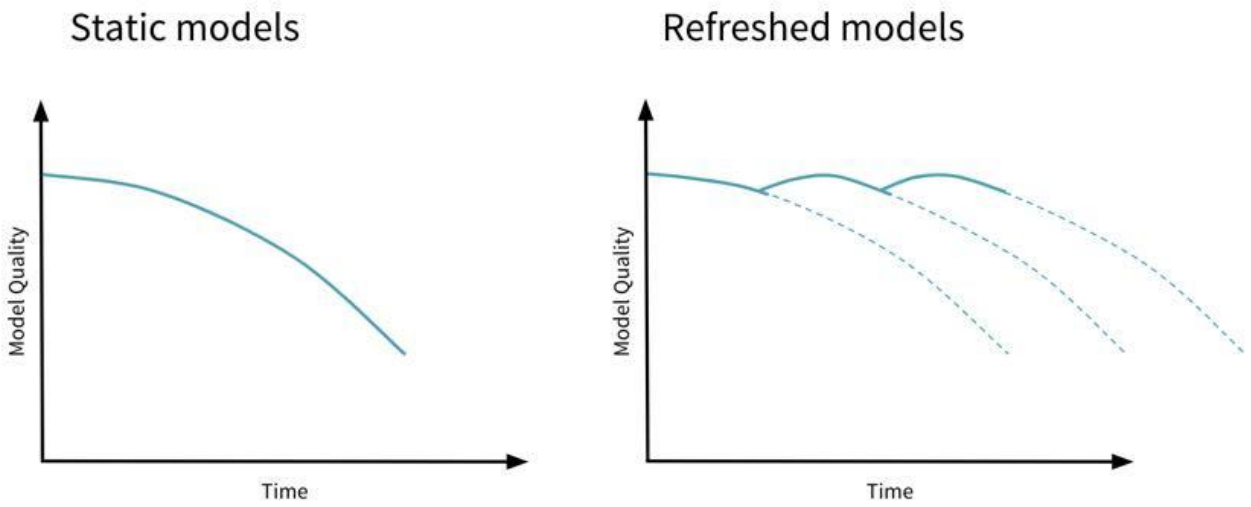
Sentiment analysis: '*This song is sick!*' avrebbe assunto un significato completamente diverso prima degli anni '90.

Concept e data drift



https://databricks.com/blog/2019/09/18/productionizing-machine-learning-from-deployment-to-drift-detection.html

Concept e data drift

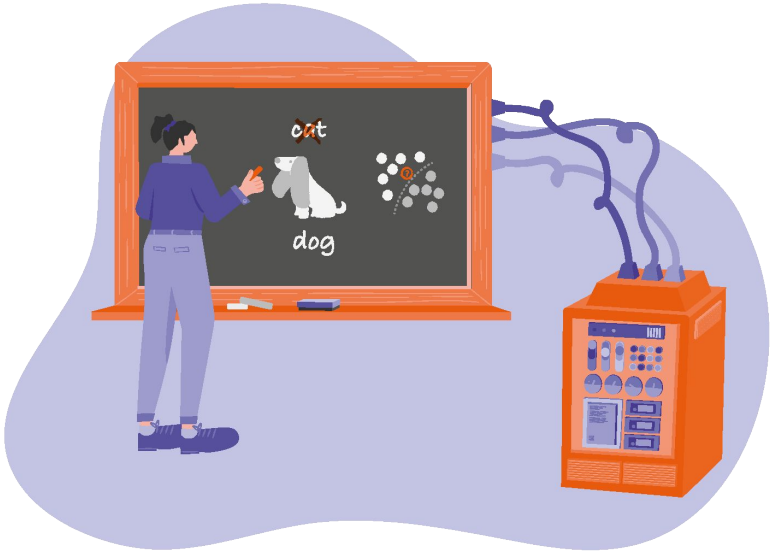


https://databricks.com/blog/2019/09/18/productionizing-machine-learning-from-deployment-to-drift-detection.html

Monitoraggio performance nel tempo

= Bisogno di etichettatura continua

Monitorate performance di modelli (accuracy, F1, etc) richiede la conoscenza della ground truth relativa ai dati che arrivano dal vivo.



Due possibilità:

- Processo automatico

  → problemi di forecasting
- Processo richiede etichettatura umana

  → Esempio, decision support system

Servizi etichettatura

AWS Mechanical Turk

AWS offre servizi di etichettatura → marketplace che offre accesso a migliaia di freelancers

SageMaker Model Monitor

Servizio monitoraggio di SageMaker.

- (al momento) Utilizzabile solamente quando in presenza di **dati strutturati**
- Permette di impostare ‘monitoring schedules’, check su dati e modello dal vivo effettuati secondo intervalli specificati dall’utente
- Richiede utilizzo di una macchina virtuale dedicata (mediamente più costosa)

SageMaker Model Monitor

Tipi monitoraggio

SageMaker offre 4 famiglie di moduli di monitoraggio

- **Data quality**
- Model quality
- Bias
- **Explanations**

SageMaker Data Quality Monitor

Installazione SageMaker client:

pip install sagemaker

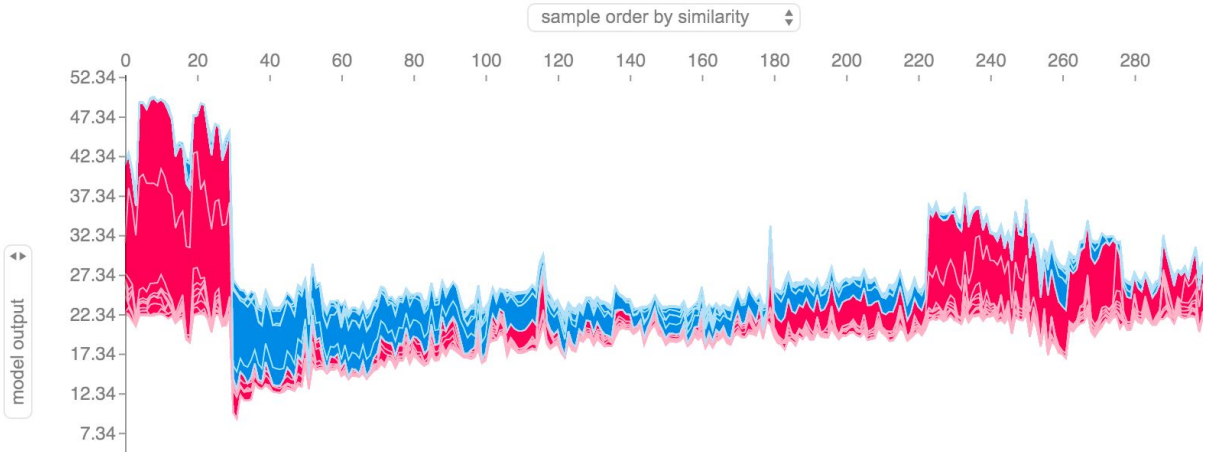
https://github.com/Clearbox-AI/Corso_MLOps/blob/main/sessione4/SageMaker_Monitoring.ipynb

Monitoraggio delle spiegazioni

delle decisioni del modello

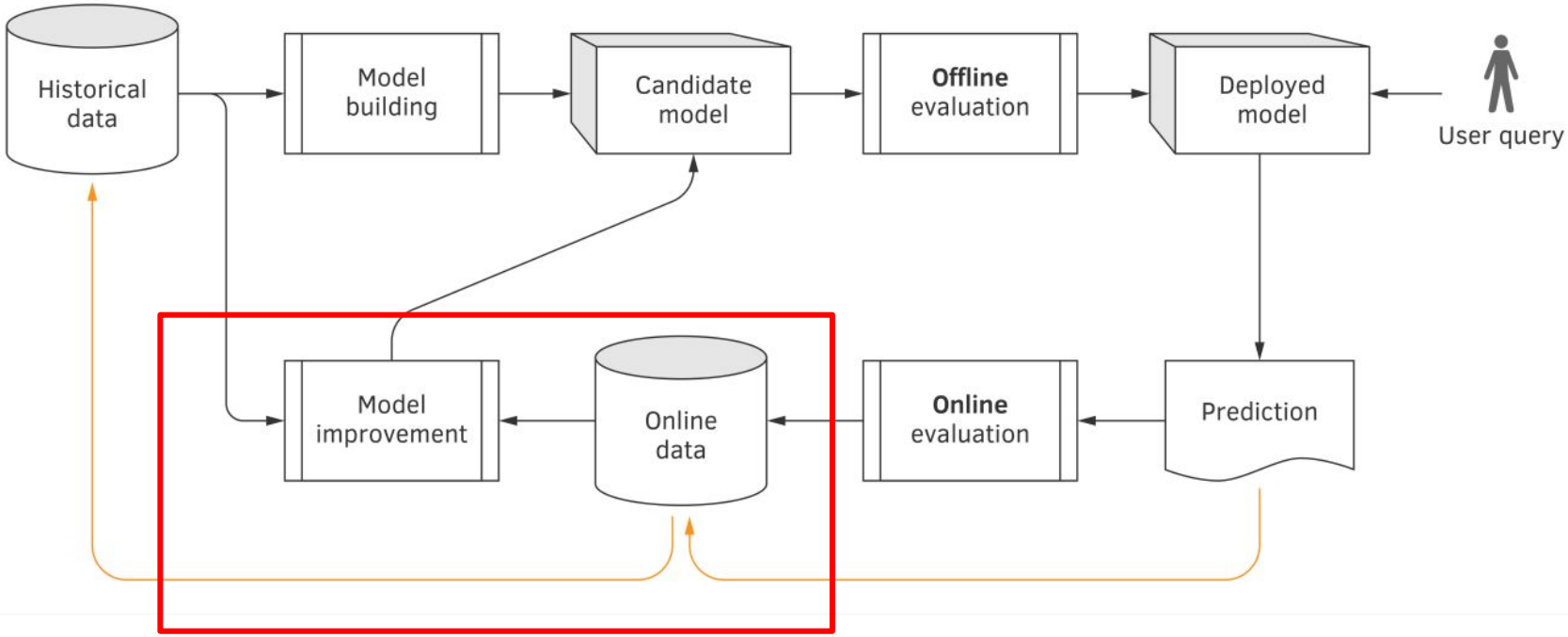
Spiegazioni delle decisioni del modello possono essere viste come un modo di descrivere quantitativamente il suo comportamento.

Lo stesso discorso può essere affrontato in ambito di monitoraggio: **come e quanto sta cambiando il comportamento del modello nel tempo**?



SageMaker Explainability Monitor

https://github.com/Clearbox-AI/Corso_MLOps/blob/main/sessione4/SageMaker_Explainable_Monitoring.ipynb


Historical data
Model building
Candidate model
Offline evaluation
Deployed model
User query
Model improvement
Online data
Online evaluation
Prediction

Active Learning

In molti casi etichettare dati può essere costoso in termini di tempo e denaro.

Lo scopo dell'active learning e' di definire strategie atte a migliorare il processo di ri-allenamento dei modelli usando il minor numero di nuove etichette possibile.

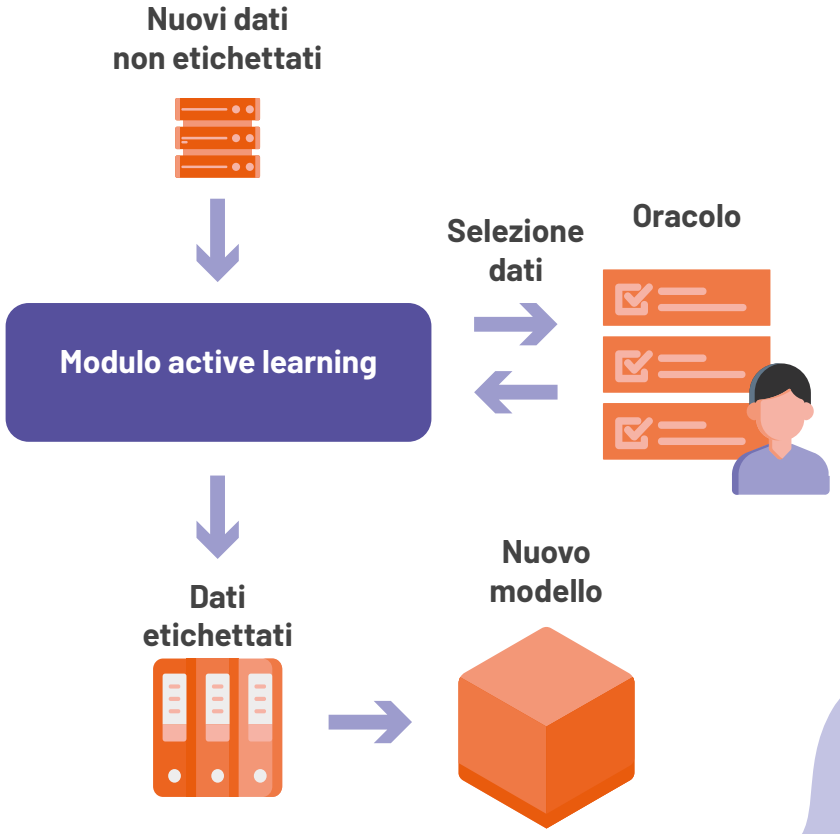
In fisica lo stesso tipo di problema viene definito come '*Design Of Experiments*'

Active Learning

In molti casi etichettare dati può essere costoso in termini di tempo e denaro.

Lo scopo dell'active learning e' di definire strategie atte a migliorare il processo di ri-allenamento dei modelli usando il minor numero di nuove etichette possibile.

In fisica lo stesso tipo di problema viene definito come '*Design Of Experiments*'



Active Learning

Come scegliere i punti da etichettare?

Diverse strategie possibili. Tra le più popolare strategie basate sulla riduzione dell'incertezza del modello che si vuole migliorare (**uncertainty sampling**)

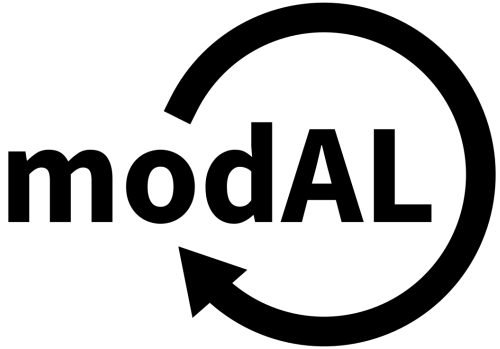
Smallest Margin Uncertainty

$$\phi_{SM}(x) = P_\theta(y_1^*|x) - P_\theta(y_2^*|x)$$

Lowest Confidence Uncertainty

$$\phi_{LC}(x) = 1 - P_\theta(y_1^*|x)$$

Active Learning



Framework per l'active learning in Python:

modAL

Libreria compatibile con il formato scikit-learn (BaseEstimator), permette di utilizzare diverse strategie per il sampling dei punti da etichettare.

https://github.com/modAL-python/modAL

Esercizio Active Learning

Installazione modAL client:

pip install modal

https://github.com/Clearbox-AI/Corso_MLOps/blob/main/sessione4/ActiveLearning.ipynb

Tecniche aggiornamento modelli

single deployment: nuovo modello direttamente sostituito al vecchio

silent deployment: nuovo modello fatto girare in parallelo per un intervallo di tempo ma utenti non esposti

canary deployment: nuovo modello fatto girare in parallelo ed esposto a piccola frazione di utenti

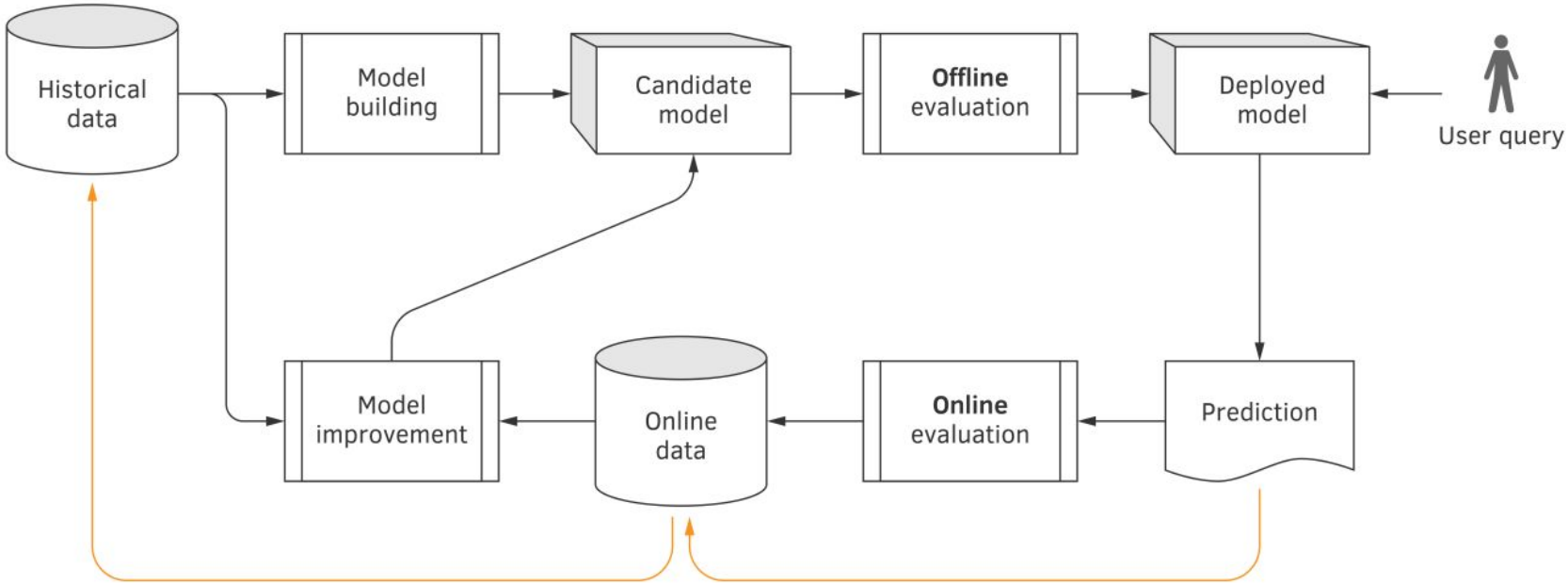
Tecniche aggiornamento modelli

multi-armed bandits

Problema → Come allocare risorse (queries assegnate ad un determinato modello) senza conoscere con certezza le metriche associate a ciascun modello.

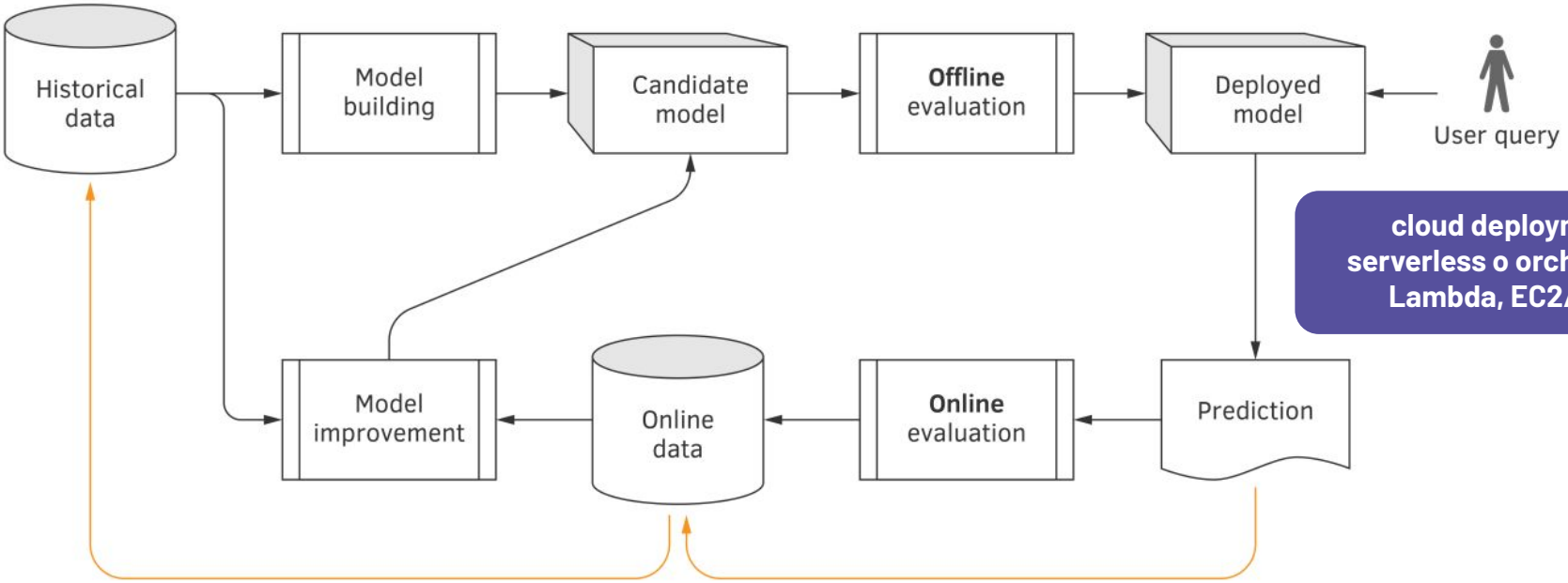
Exploration ↔ Exploitation

Esistono algoritmi dedicati a questo tipo di problema (es, Upper Confidence Bound) che determinano dal vivo l'assegnazione delle queries fino ad arrivare ad una scelta del modello da utilizzare.


Historical data
Model building
Candidate model
Offline evaluation
Deployed model
User query
Model improvement
Online data
Online evaluation
Prediction

experiment tracking, model testing
mlflow, github actions

spiegazioni, calibrazione
SHAP, sklearn.calibration



cloud deployment
serverless o orchestrato
Lambda, EC2/ECS

active learning
modAL

monitoraggio
Model Monitor

Esploriamo i contenuti dei bucket di monitoraggio

e fermiamo le istanze

Per vedere i task di monitoraggio attivi

aws sagemaker list-monitoring-schedules

Per fermare un determinato task

aws sagemaker delete-monitoring-schedule --monitoring-schedule-name nome-task

Extra

Esercizio sessione 2, export di un modello mlflow → SageMaker

Demo Clearbox model assessment tool.

clearbox AI
humans && machines

Thanks for Reading

Feel free to contact us:


www.clearbox.ai


shalini@clearbox.ai
giovannetti@clearbox.ai


@ClearboxAI