

Potenziare i Modelli di Intelligenza Artificiale tramite l'Ottimizzazione degli Iperparametri

Applicazioni di statistica avanzata in azienda

Valerio Perrone

Applied Science Manager, Amazon Web Services (AWS)

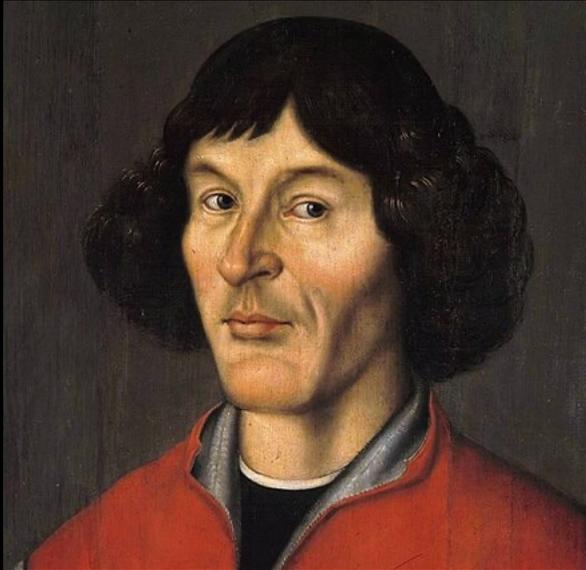
Disclaimer

Le opinioni, i pensieri e i punti di vista espressi in questa presentazione appartengono esclusivamente al presentatore e non riflettono necessariamente le opinioni di Amazon Web Services, Inc.

Agenda

- Statistica e intelligenza artificiale (IA)
- 3 casi studio: previsioni più rapide, accurate, a costi ridotti
- Cos'è il tuning degli iperparametri e perché è così importante?
- Come trovare gli iperparametri ottimali con la statistica avanzata?
- Q&A

Rivoluzione Copernicana

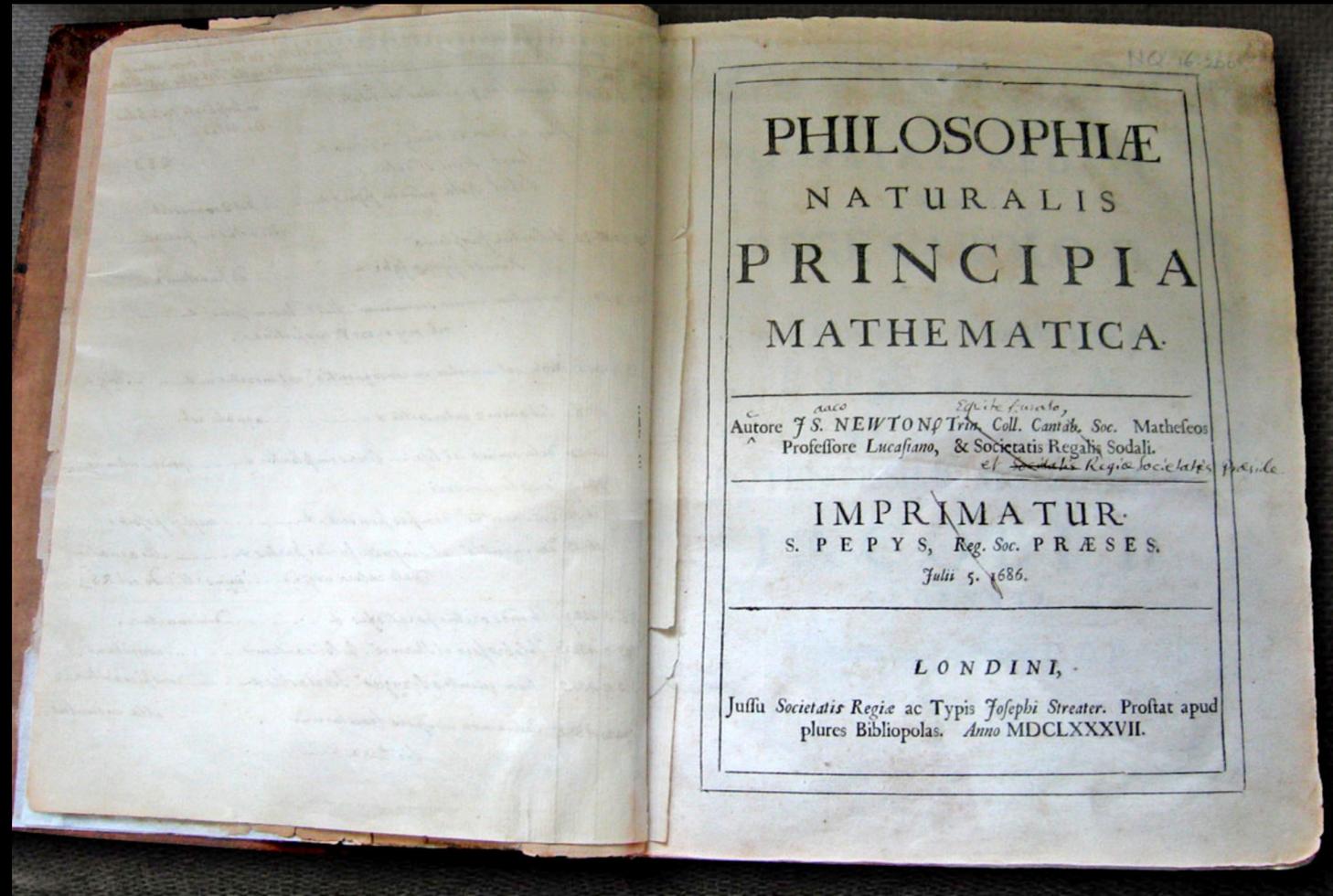


Niccolò Copernico
(1473 – 1543)

Modelli basati su...

Teoria → Dati

Modelli basati sulla teoria



Modelli basati sui dati

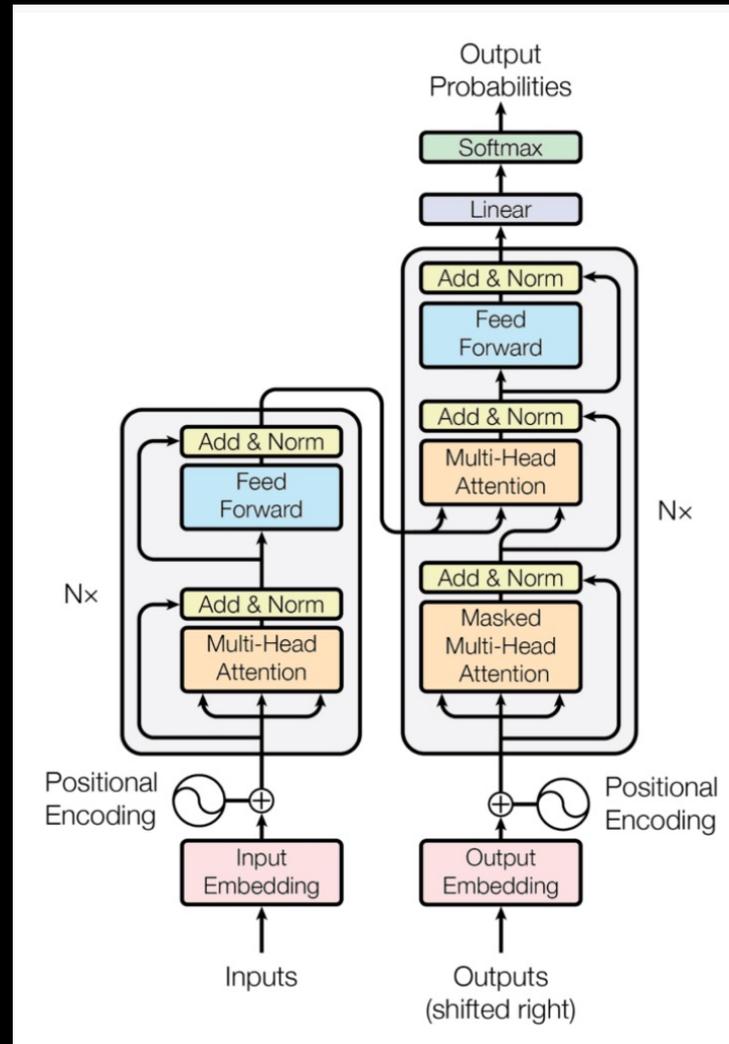
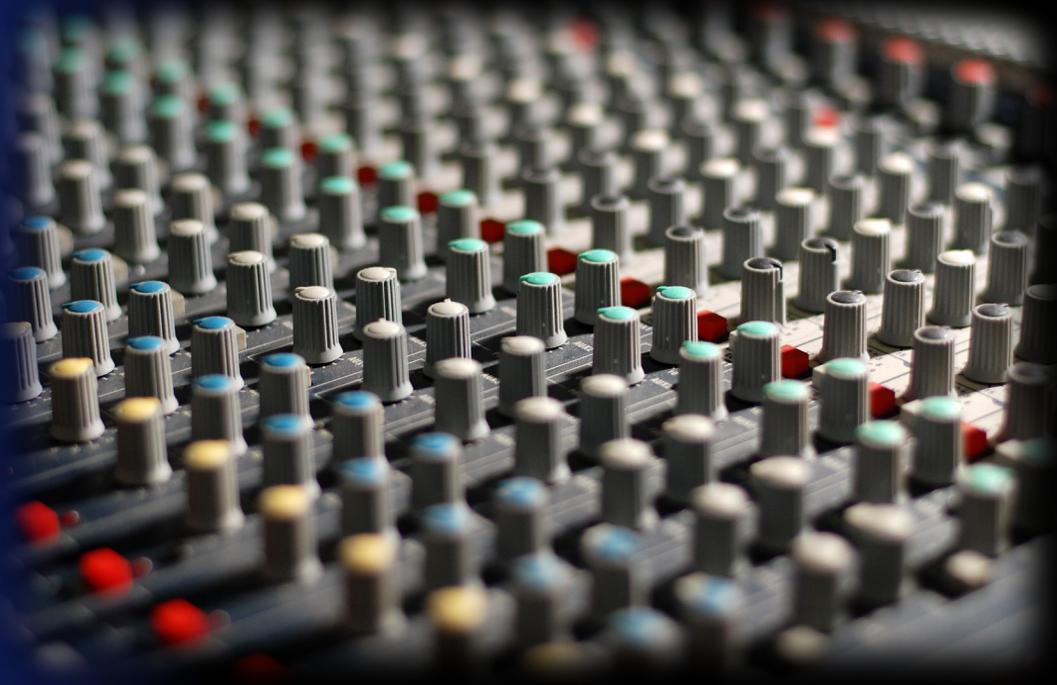


Immagine:
Articolo "Attention is
all you need"

Cos'è il tuning degli iperparametri?



Alberi di decisione

Tree depth
Max leaf nodes
Gamma
Eta
Lambda
Alpha

...

Reti neurali

Number of layers
Hidden layer width
Learning rate
Embedding dimensions
Dropout

...

Task generali

(preparazione dei dati, inferenza ...)

Parametro 1
Parametro 2
...

Iperparametri

Come il tuning degli iperparametri ha aiutato grandi aziende

Best Egg achieved three times faster ML model training with Amazon SageMaker Automatic Model Tuning

HAYAT HOLDING uses Amazon SageMaker to increase product quality and optimize manufacturing output, saving \$300,000 annually

How Light & Wonder built a predictive maintenance solution for gaming machines on AWS

Caso studio 1: Best Egg stima il credit risk 3 volte più rapidamente

Best Egg:

- Piattaforma finanziaria leader
- \$22 miliardi di prestiti personali erogati

Obiettivo:

- Analisti di rischio e valutazione del credito
- ML per predire il rischio di inadempienza

Sfida:

- Mancanza di metodo universale
- Tempi e risorse sprecati

Tuning degli iperparametri:

- Valutazione del rischio e assegnazione del credito 3 volte più rapida
- Aumento della produttività dei data scientist

Caso studio 2: Light & Wonder prevede quando va fatta manutenzione migliorando il recall del 121%

- Light & Wonder:

- Azienda leader globale nel settore dei giochi d'azzardo
- Monitoraggio dei dati delle slot machine

- Obiettivo:

- Prevedere quando fare manutenzione
- Ridurre i tempi di fermo

- Sfida:

- Grandi volumi di dati
- Precisione bassa? Decisioni di manutenzione errate
- Recall basso? Perdita di guadagni

Tuning degli iperparametri:

- Ottimizzazione di una rete neurale
- Recall migliorato del 121% con precision all'80%

Caso studio 3: Hayat Holding migliora la qualità dei prodotti risparmiando 300K dollari all'anno

Hayat Holding:

- Leader globale con 41 aziende in diversi settori
- Produttore di pannelli in legno

Obiettivo:

- Prevedere la qualità dei prodotti
- Raccomandare il giusto consumo di adesivi

Sfida:

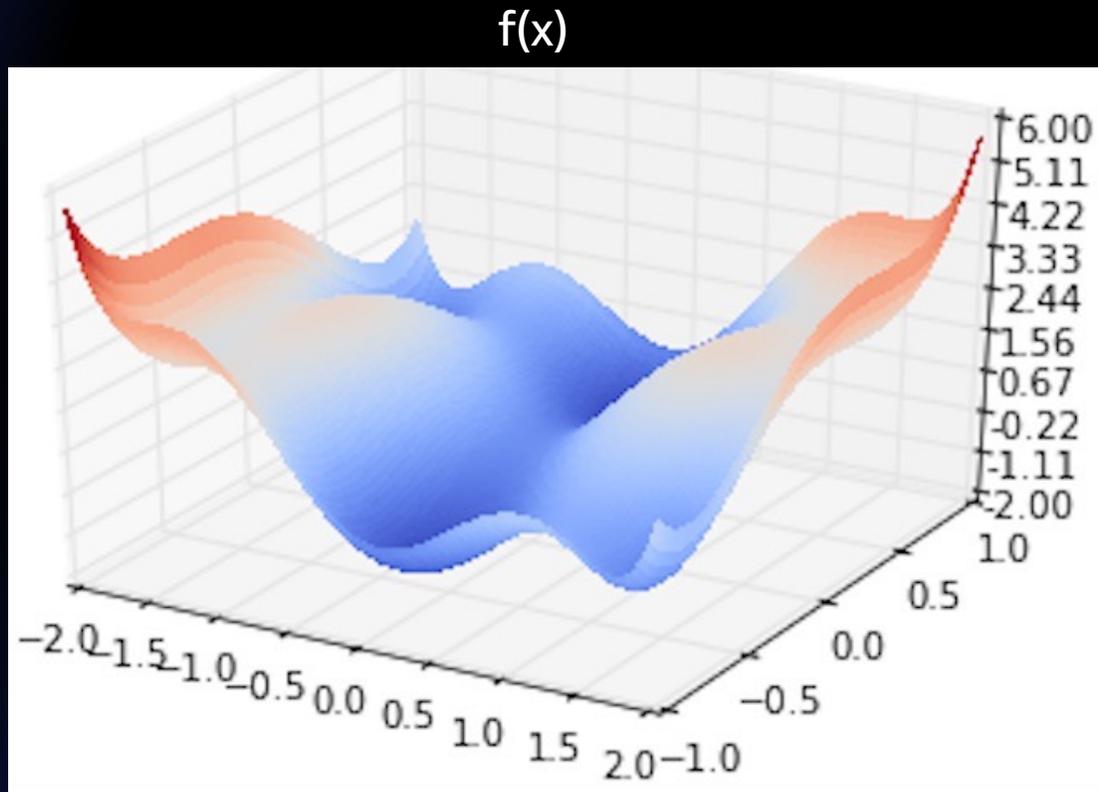
- Feedback lento con i test di laboratorio

Tuning degli iperparametri:

- Previsioni su qualità più accurate
- Risparmi annui di \$300.000
- Riduzione dell'impronta di carbonio

Ottimizzazione degli iperparametri: come funziona?

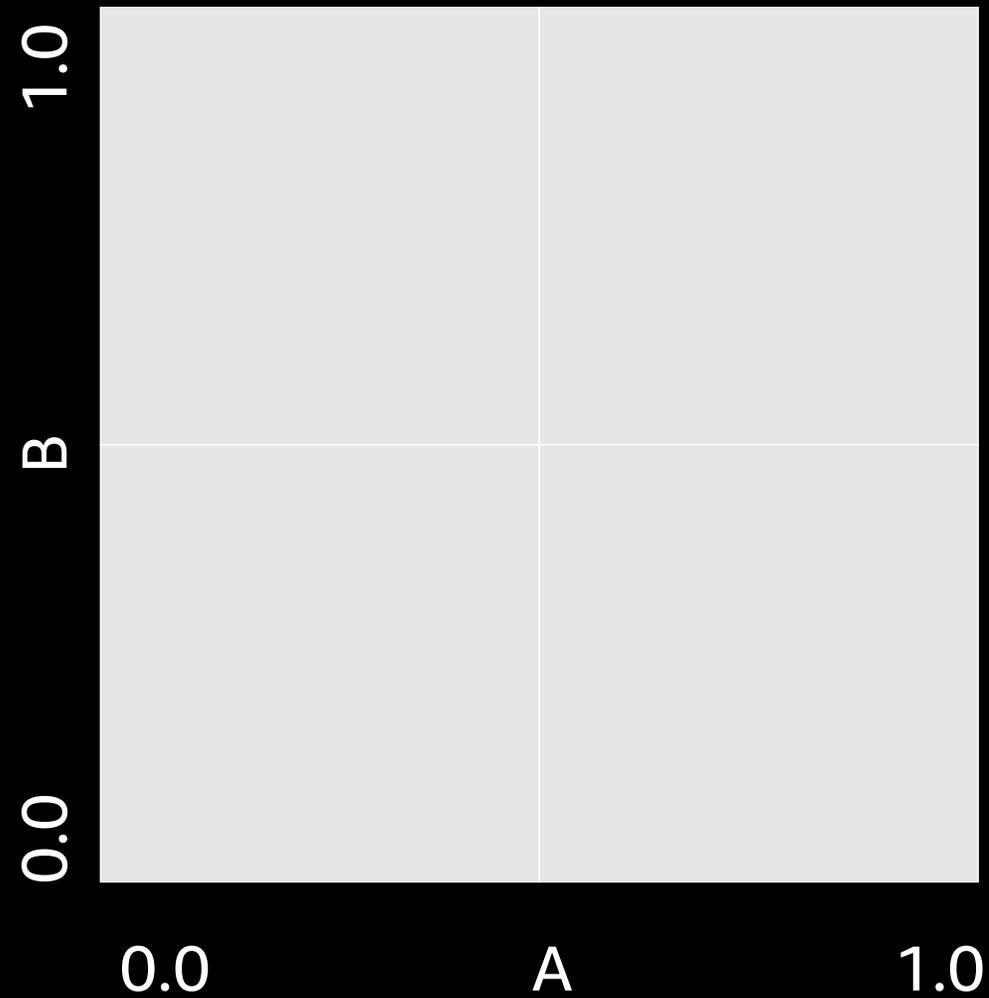
Trova la configurazione di iperparametri migliore x^* che minimizza la funzione obiettivo $f(x)$



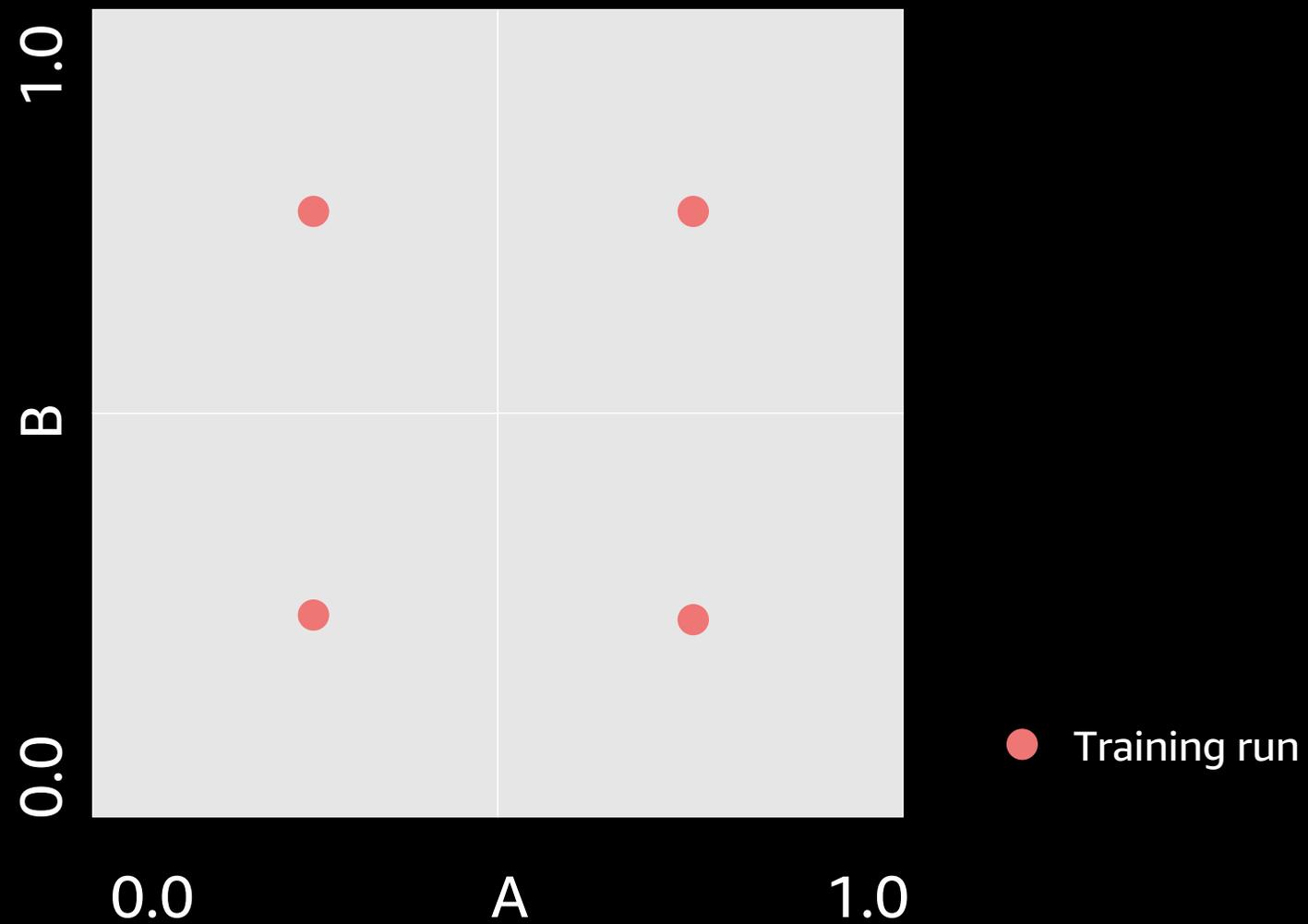
Sfide:

- $f(x)$ è costosa
- Non abbiamo la forma analitica o le derivate
- I segnali sono rumorosi

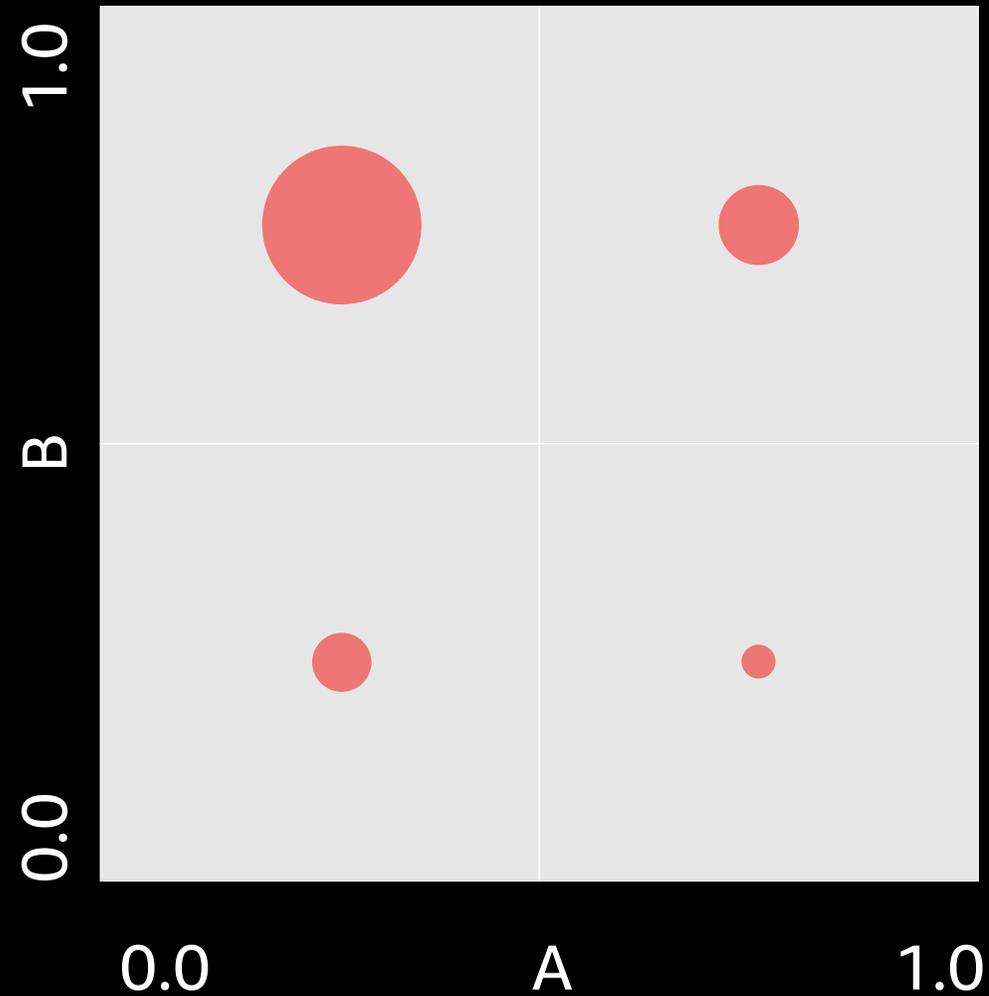
Come trovare gli iperparametri migliori?



Grid search



Grid search

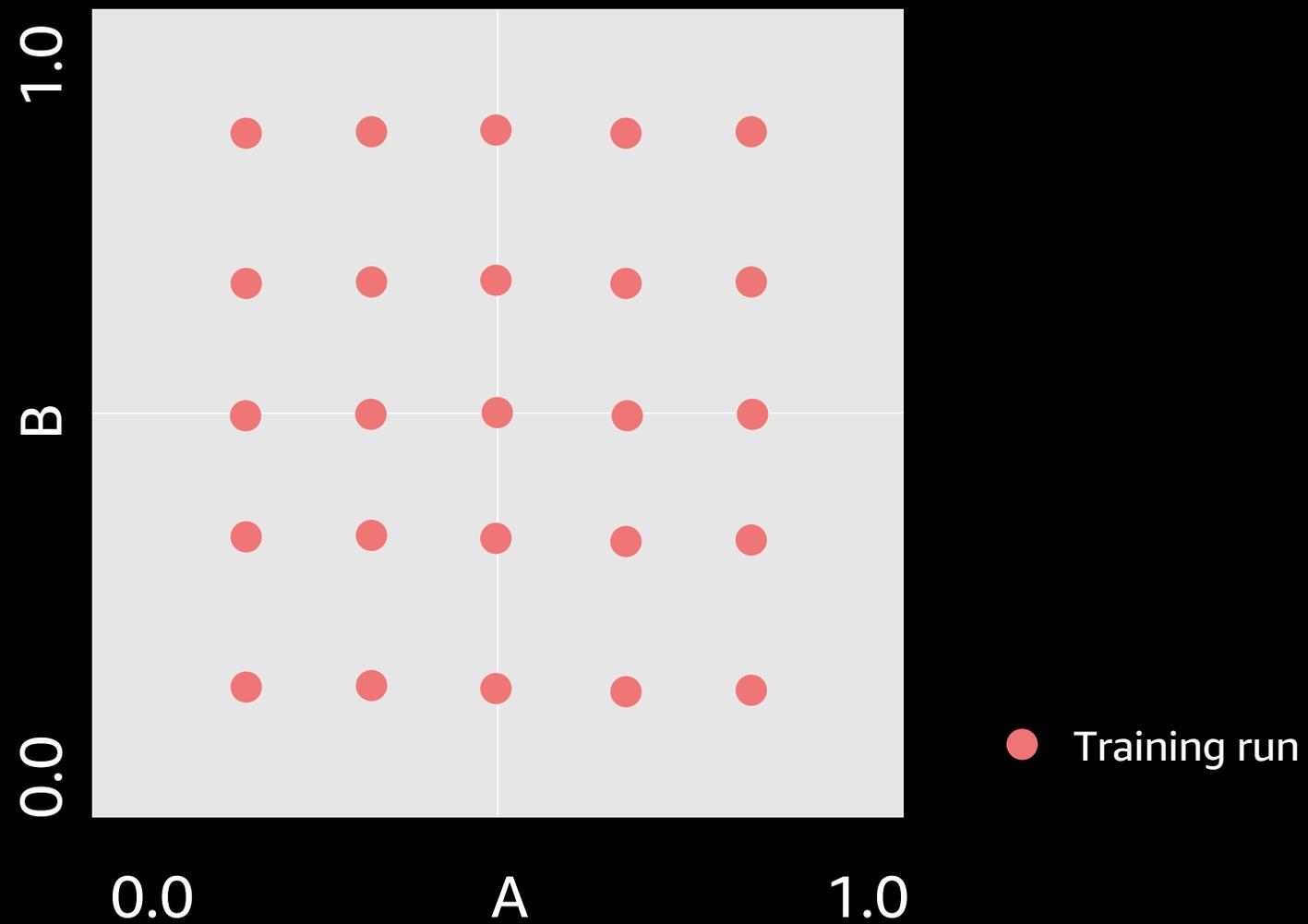


Funzione di perdita: misura l'efficacia del modello nel predire il risultato atteso. (più basso è, meglio è)

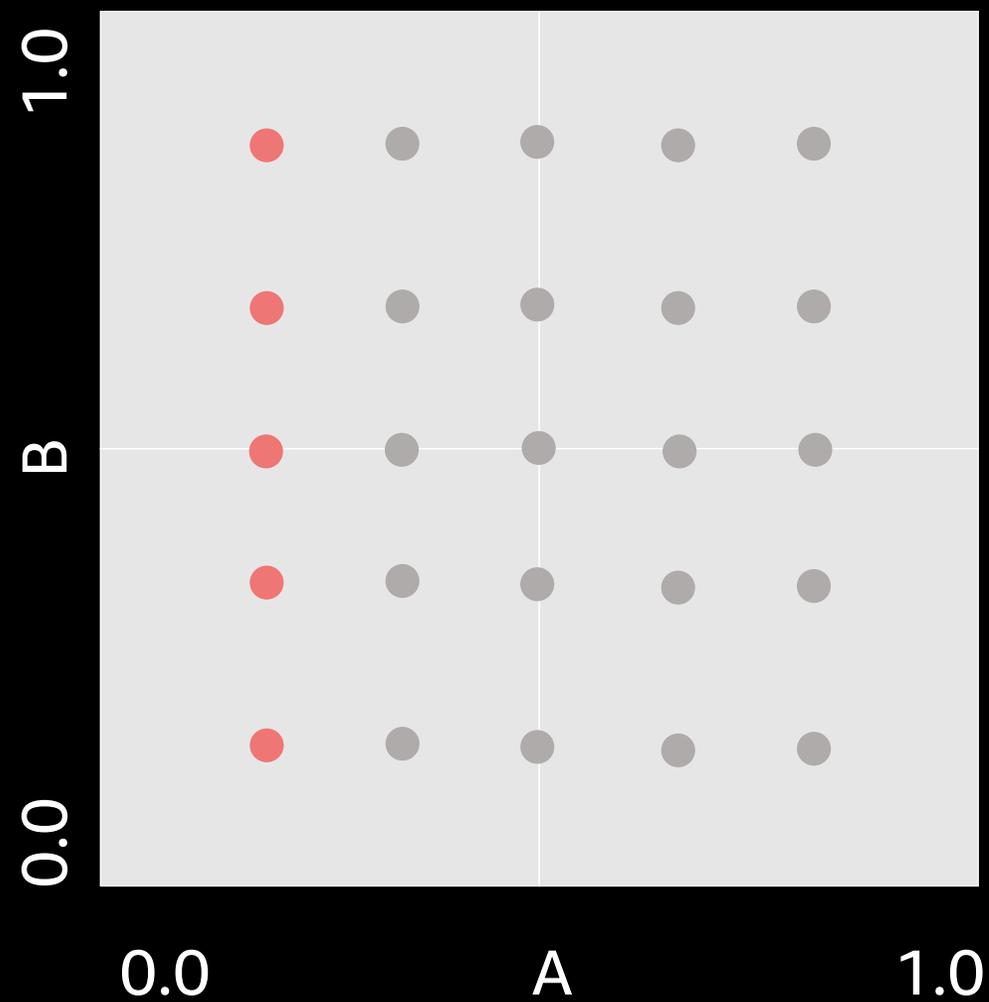
Esempio: errore di validazione

● Training
(Diametro = costo)

Grid search

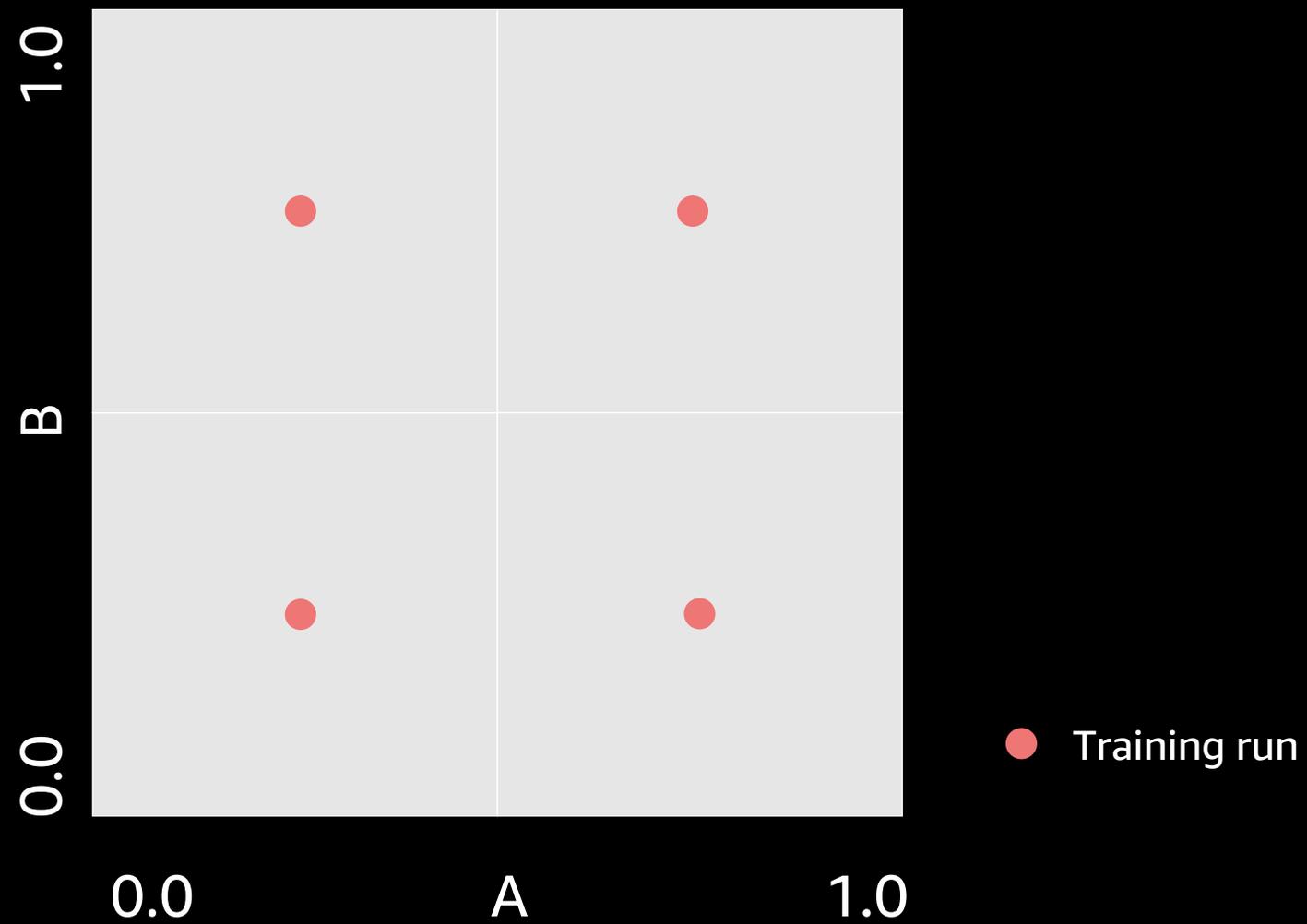


Grid search

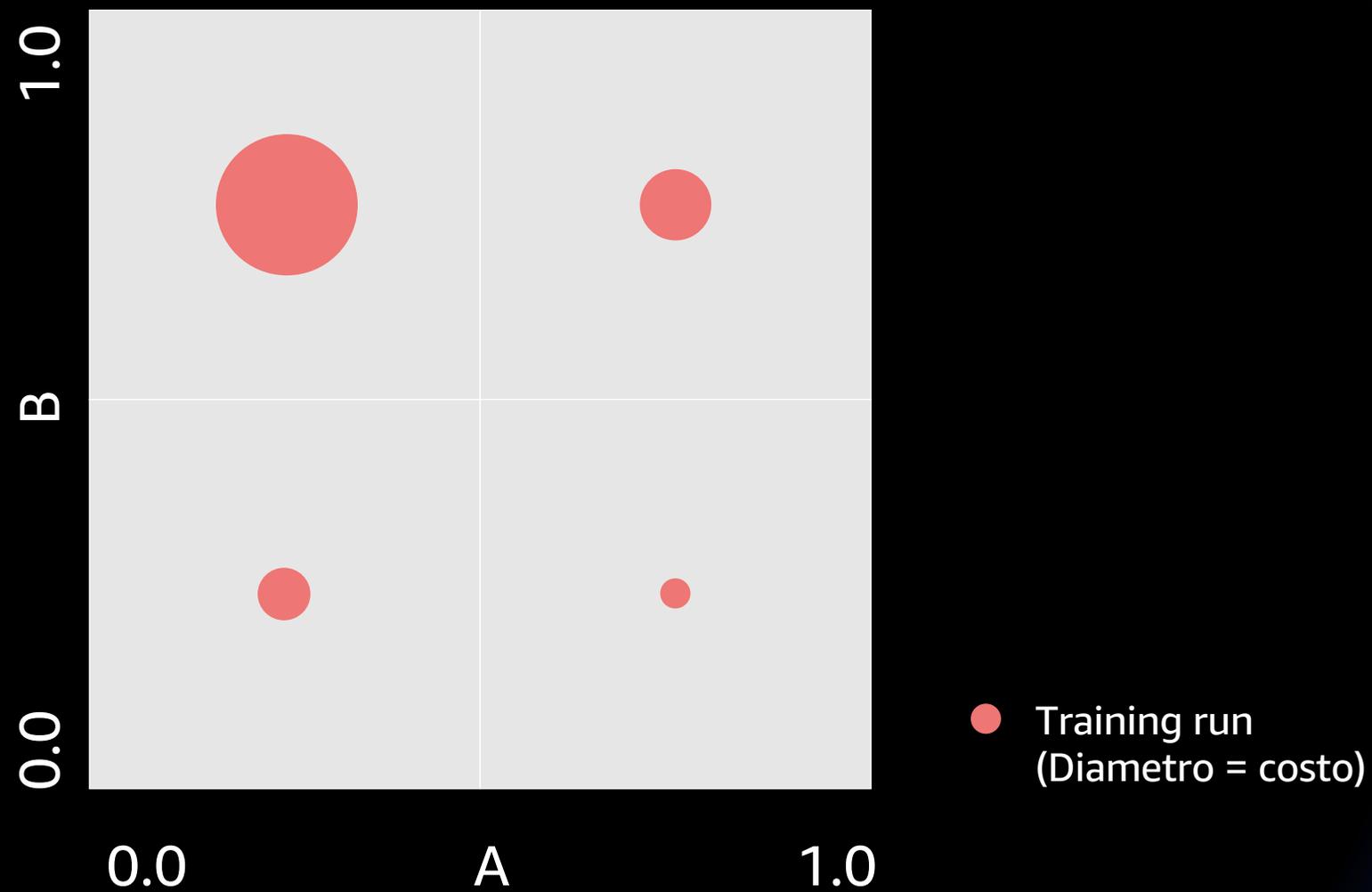


- Training run
- Training run (nessuna informazione nuova)

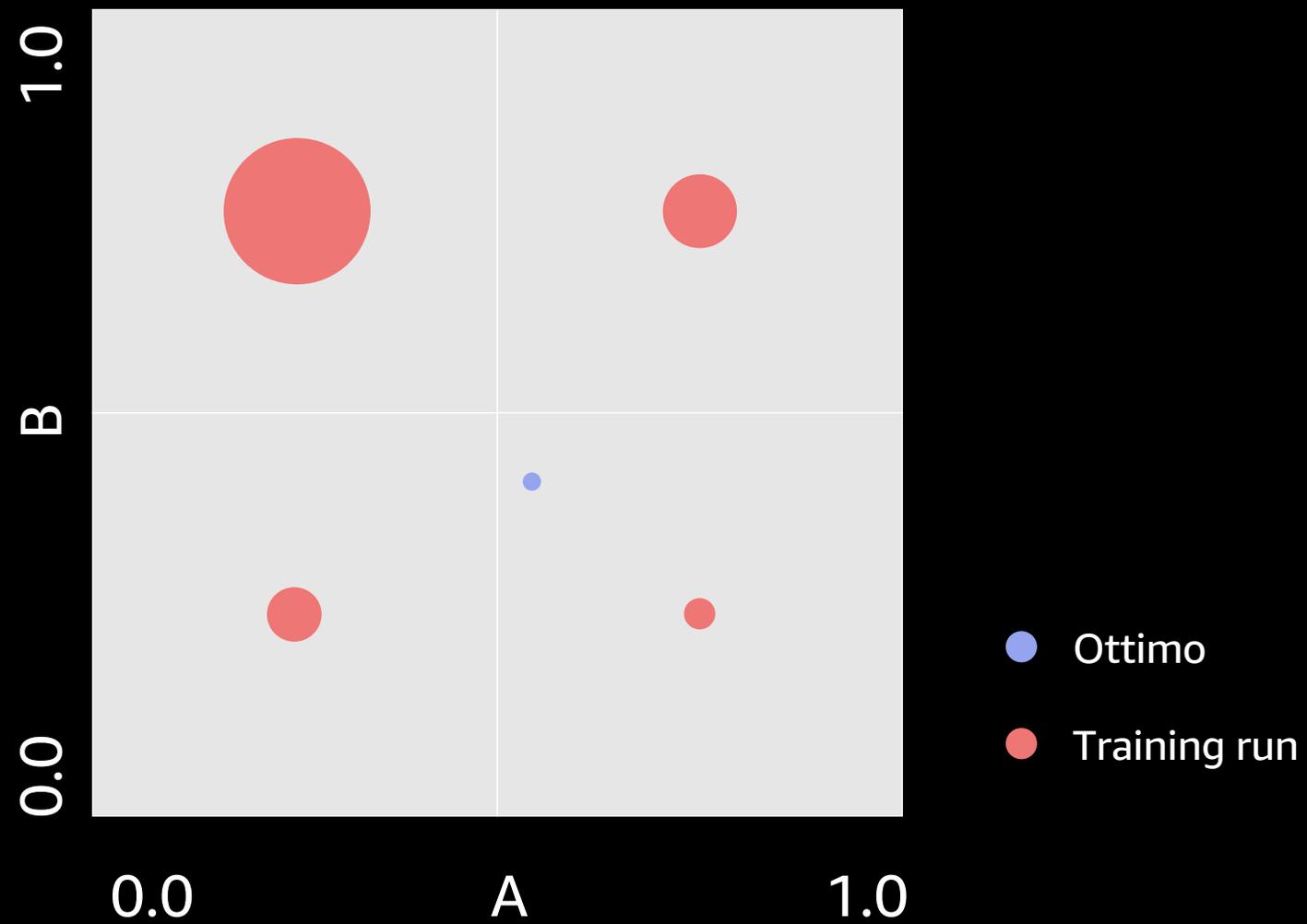
Grid search



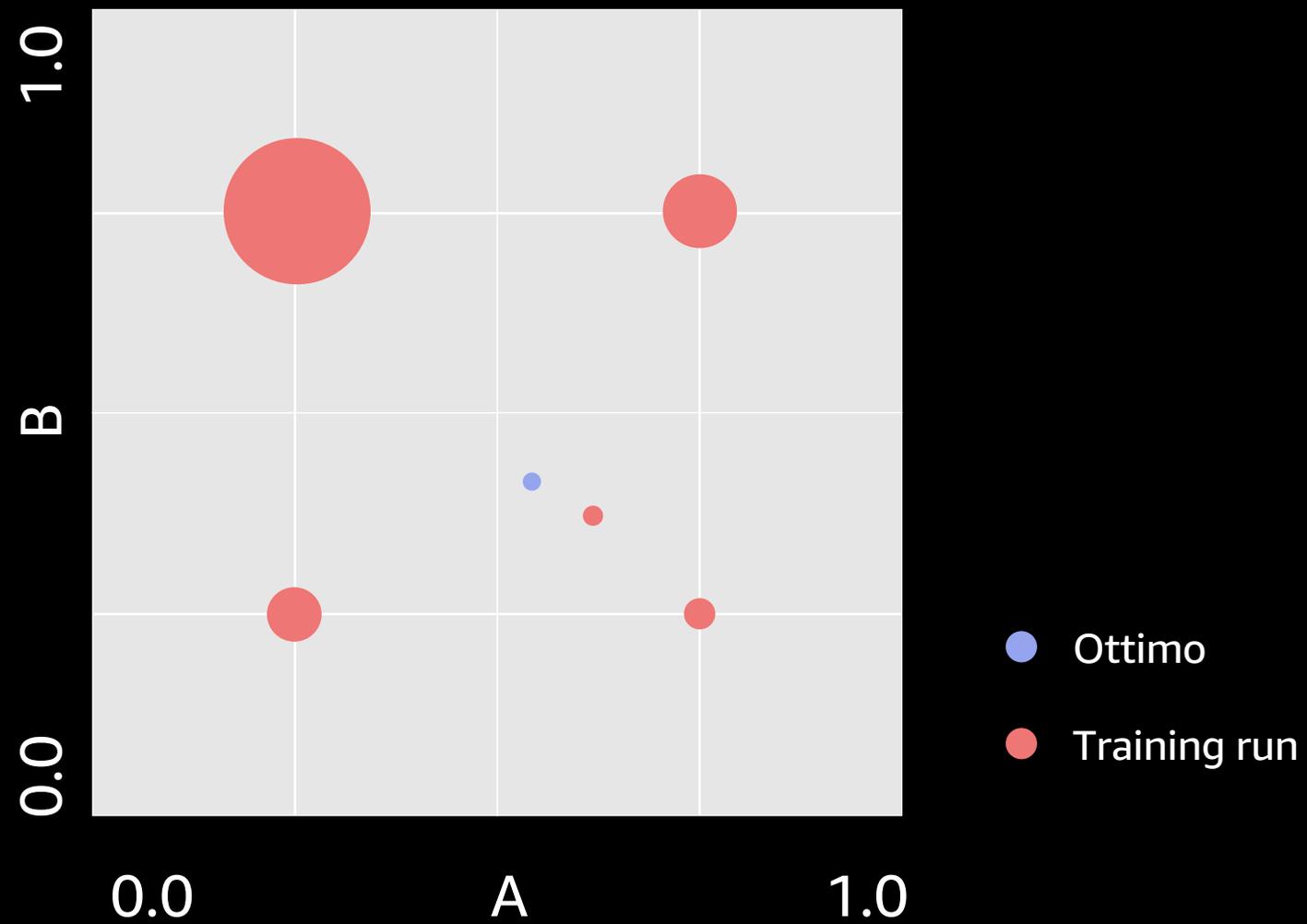
Grid search



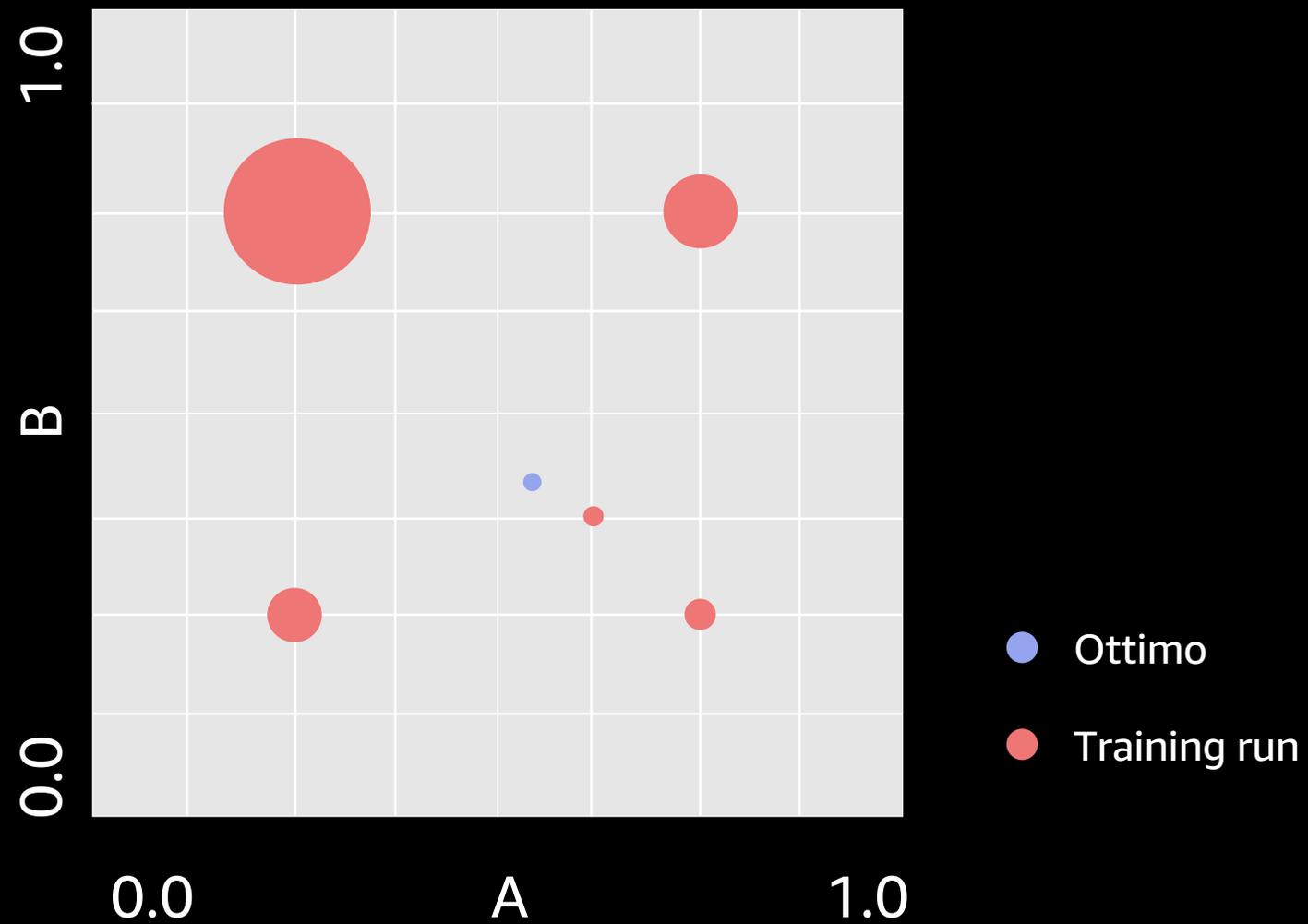
Grid search

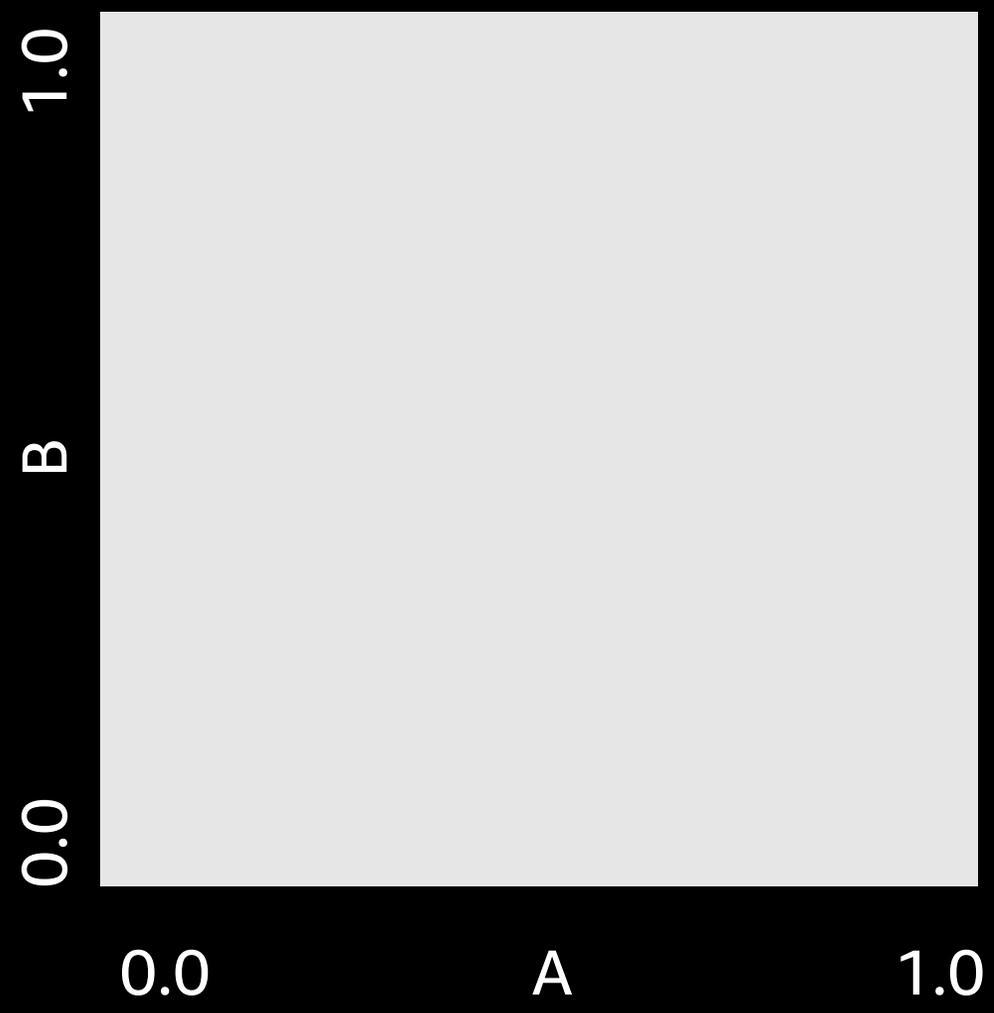


Grid search

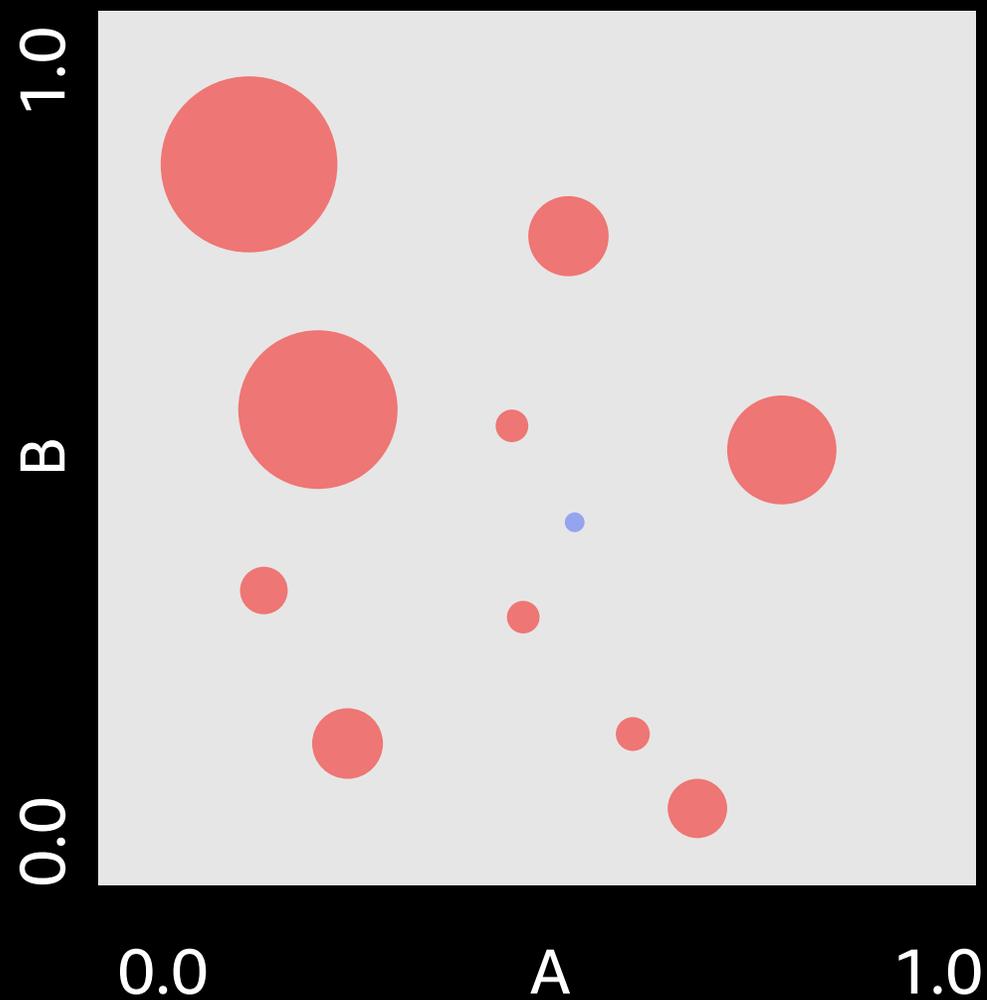


Grid search





Ricerca casuale



- Ottimo
- Training run
(Diameter = Loss)

Teorema di Bayes



Thomas Bayes
(1701 – 1761)

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

Ottimizzazione Bayesiana: 2 ingredienti

Ingrediente 1: un modello (processo Gaussiano)

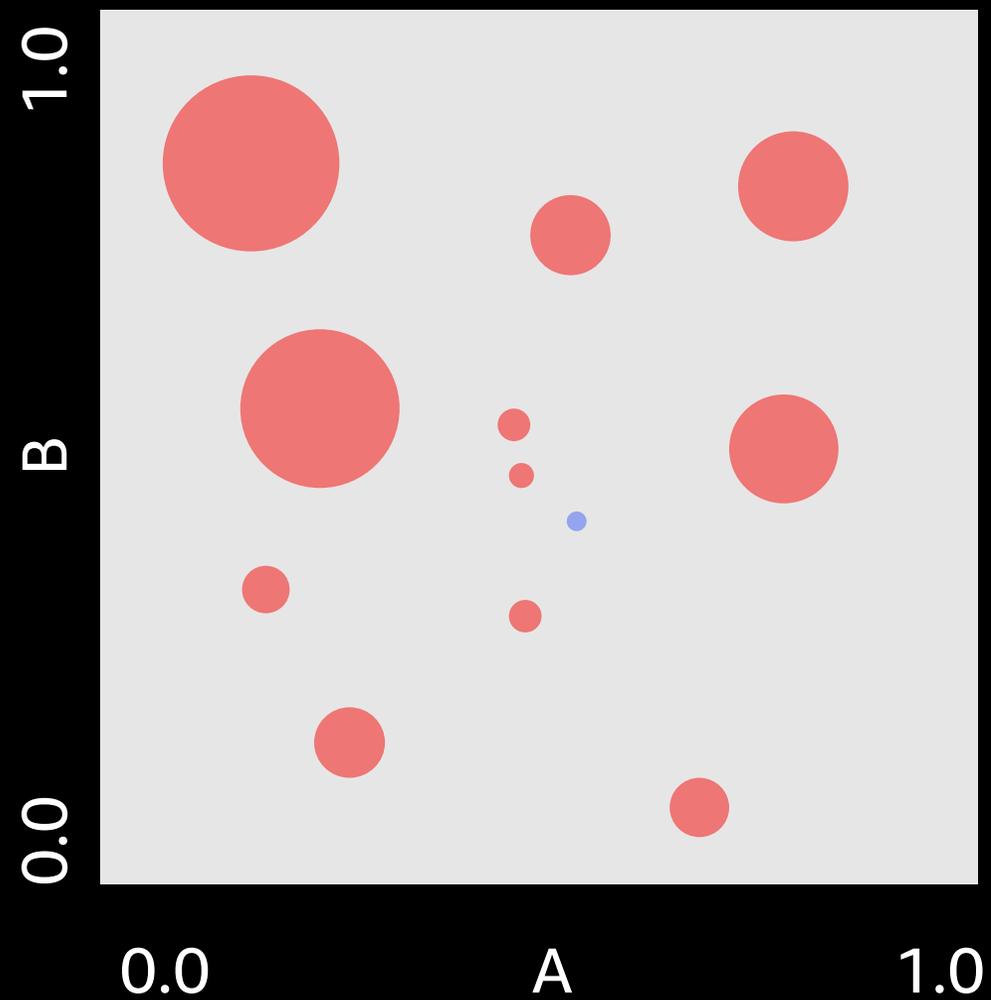
- Impara la relazione statistica tra iperparametri e funzione obiettivo

Ingrediente 2: funzione di acquisizione

- Dà un punteggio a ogni possibile configurazione di iperparametri
- Bilancia esplorazione e sfruttamento

Ricerca Bayesiana

ESPLORA VS. SFRUTTA



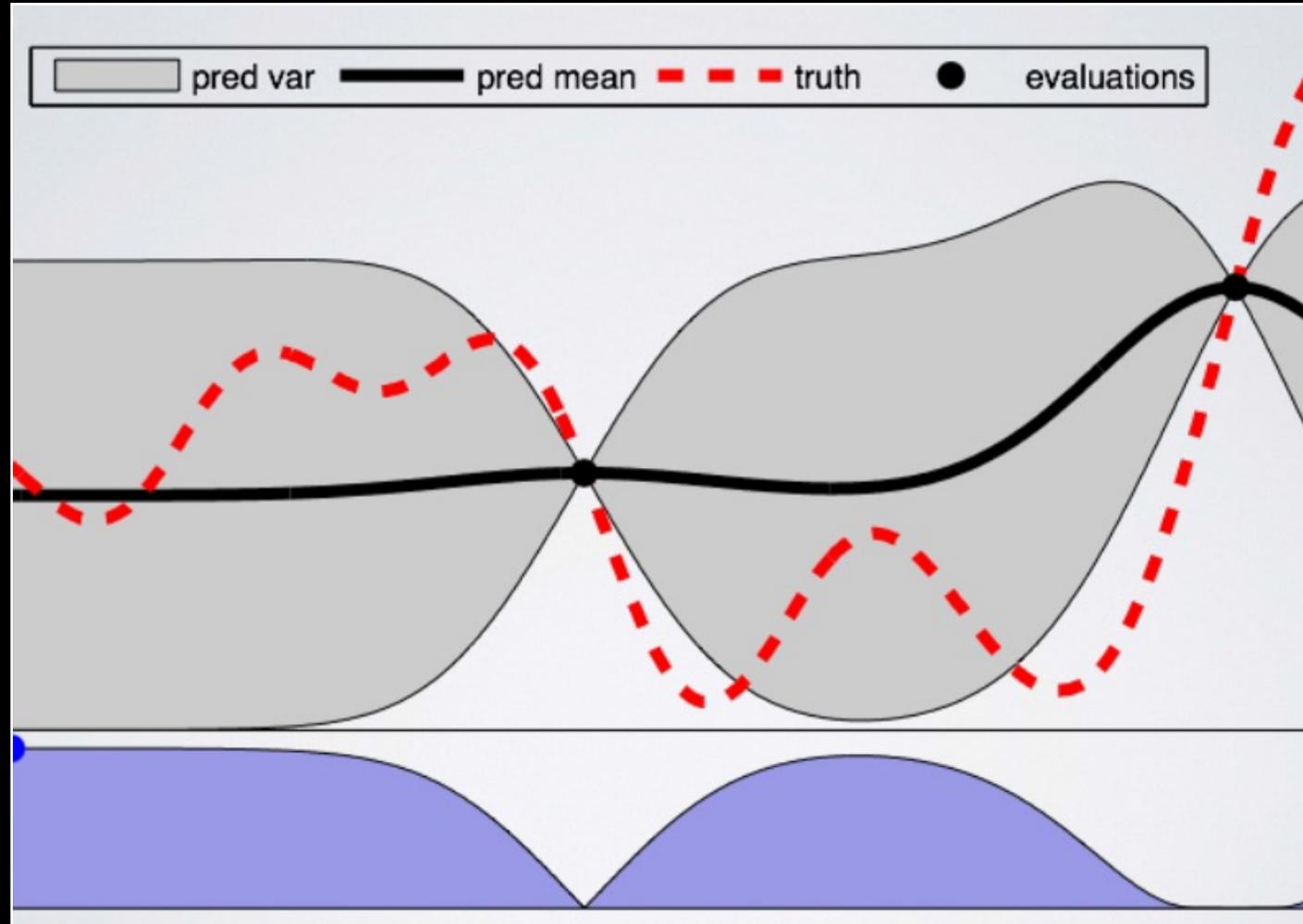
● Ottimo

● Training run
(Diameter = Loss)

Visualizziamo l'ottimizzazione Bayesiana

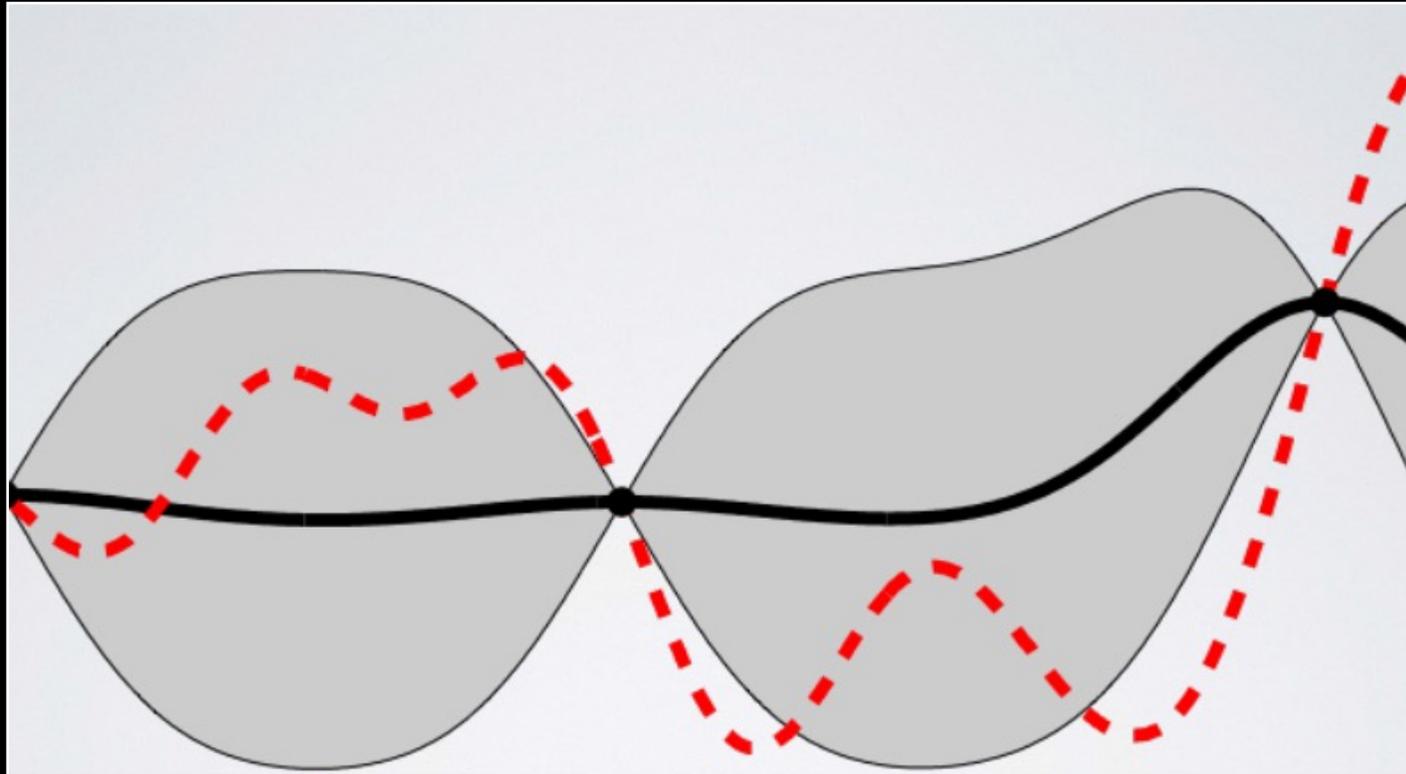
Modello

Funzione di acquisizione

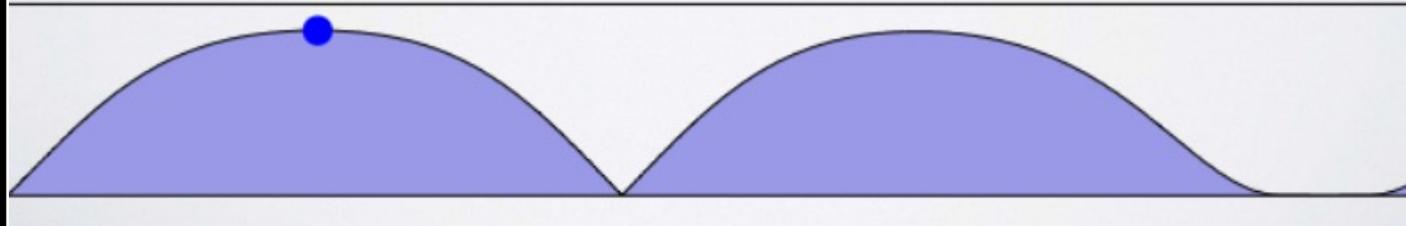


Visualizziamo l'ottimizzazione Bayesiana

Modello

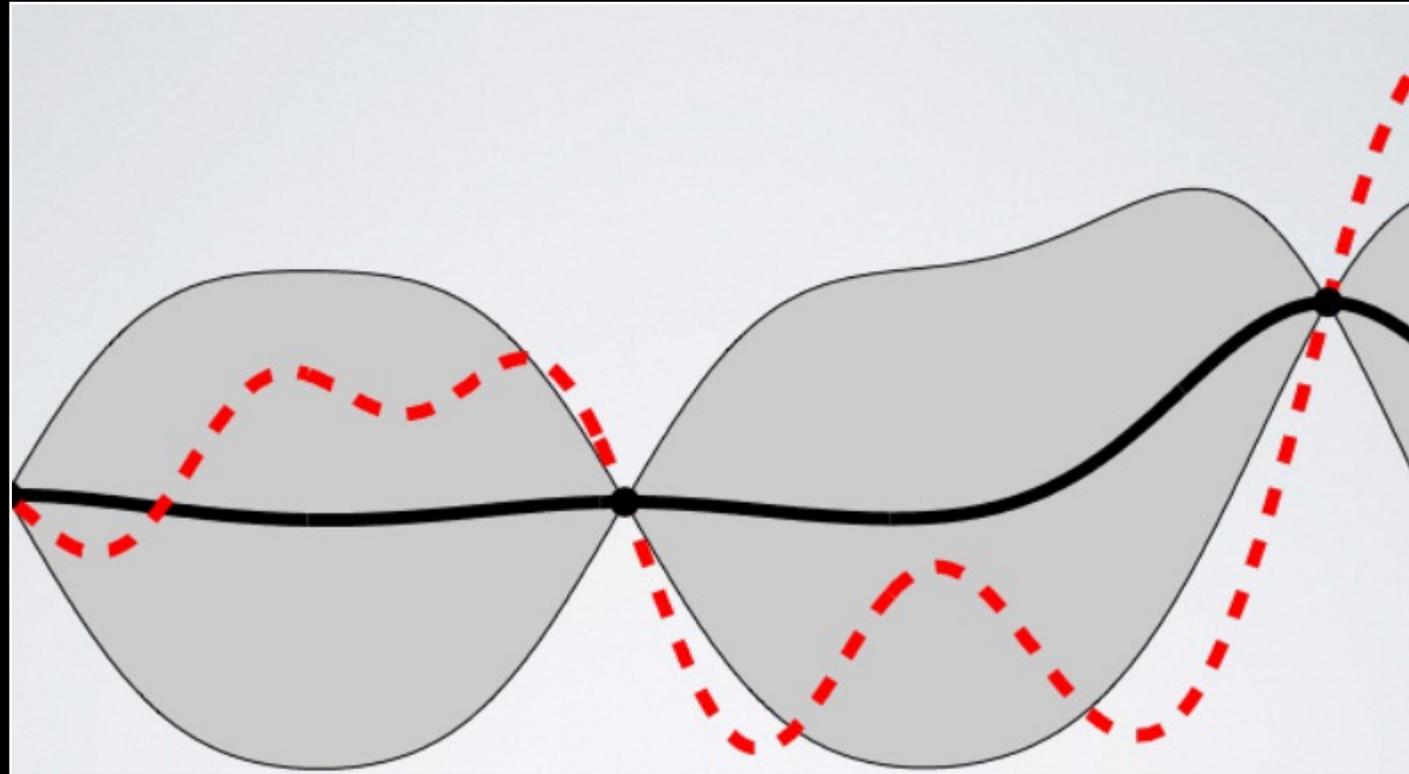


Funzione di acquisizione

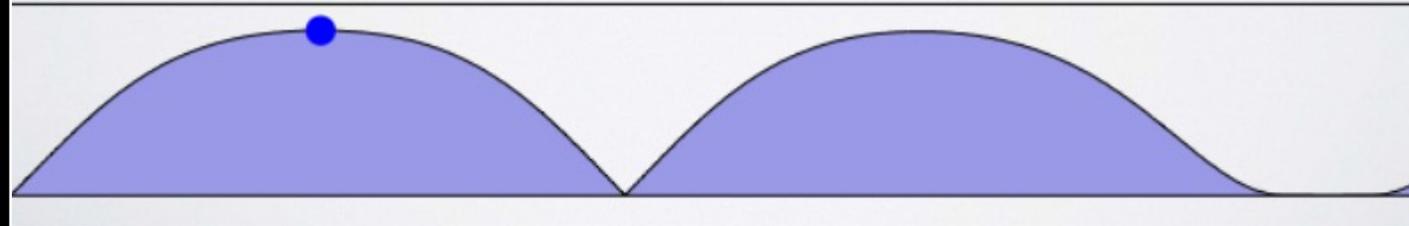


Visualizziamo l'ottimizzazione Bayesiana

Modello



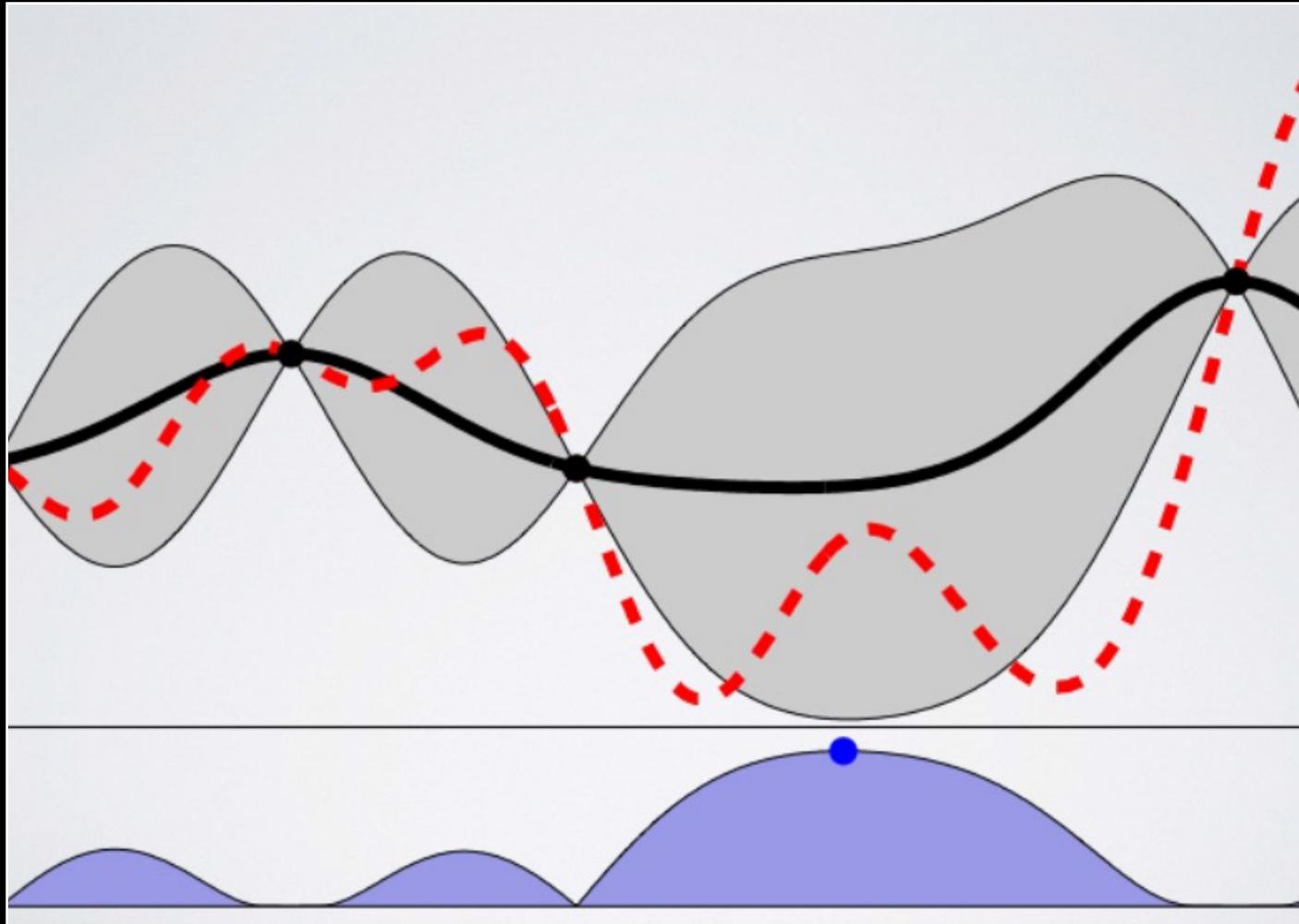
Funzione di
acquisizione



Visualizziamo l'ottimizzazione Bayesiana

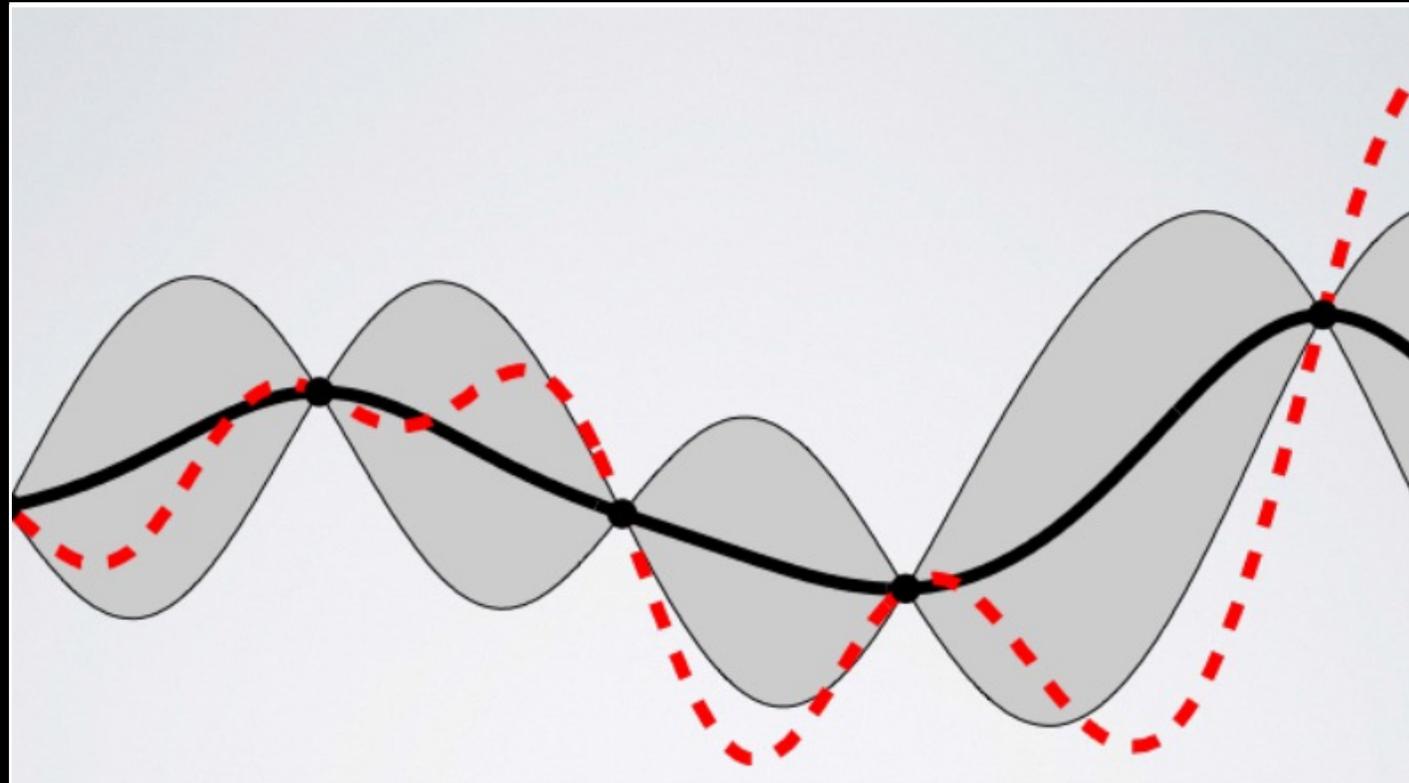
Modello

Funzione di
acquisizione

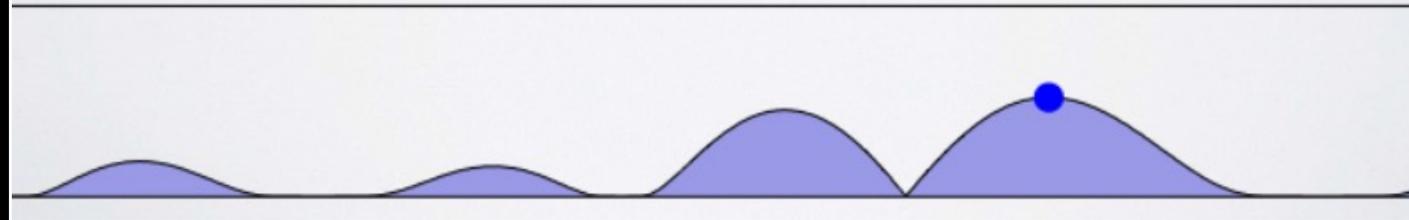


Visualizziamo l'ottimizzazione Bayesiana

Modello

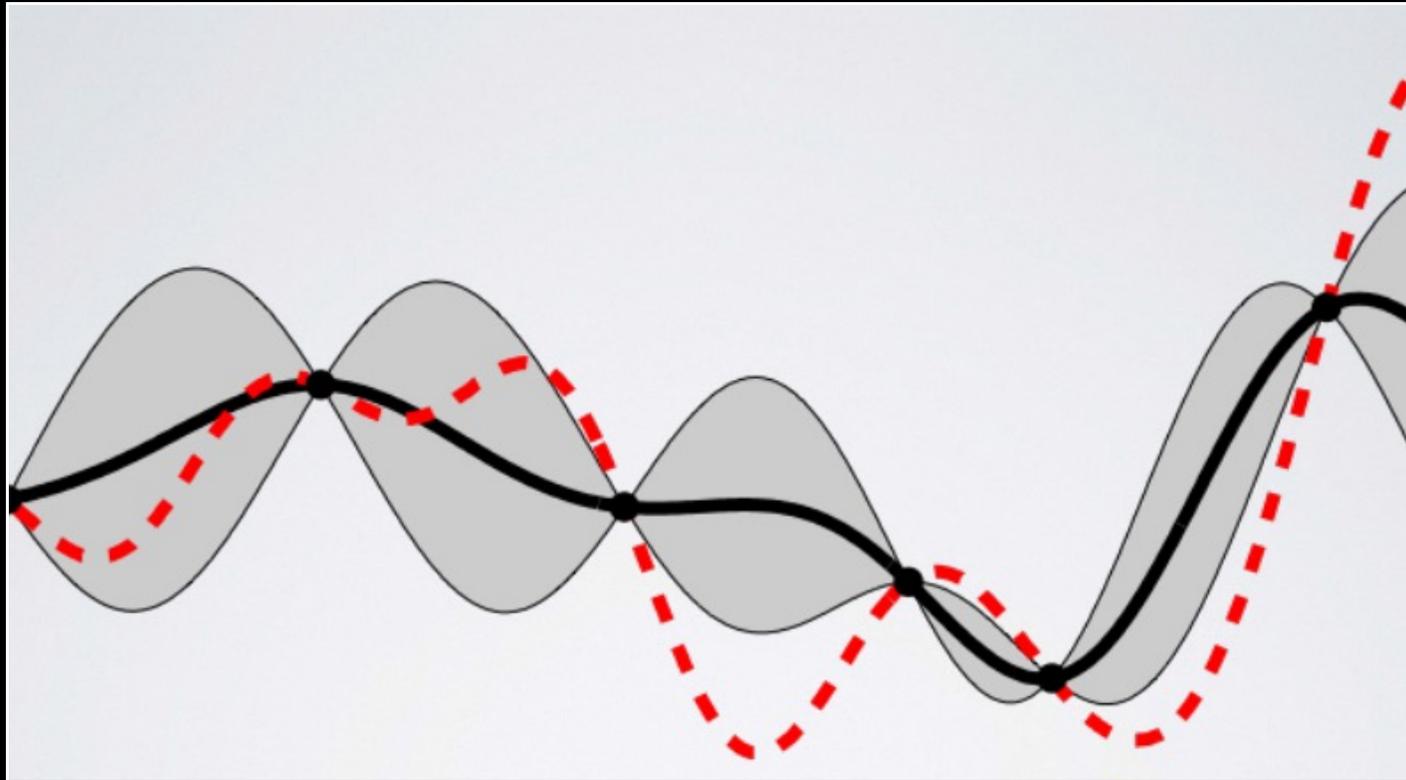


Funzione di
acquisizione

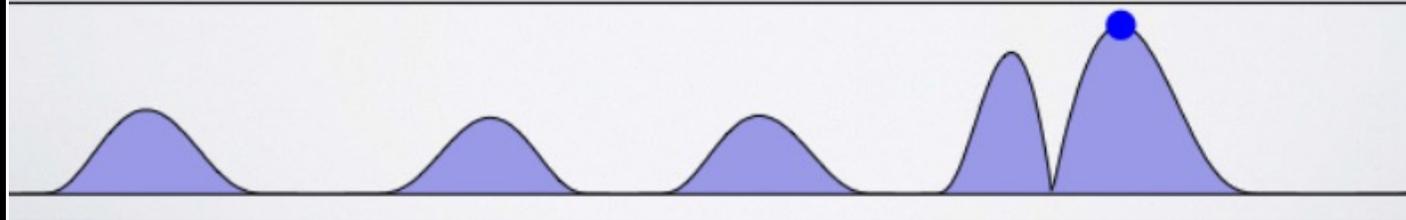


Visualizziamo l'ottimizzazione Bayesiana

Modello

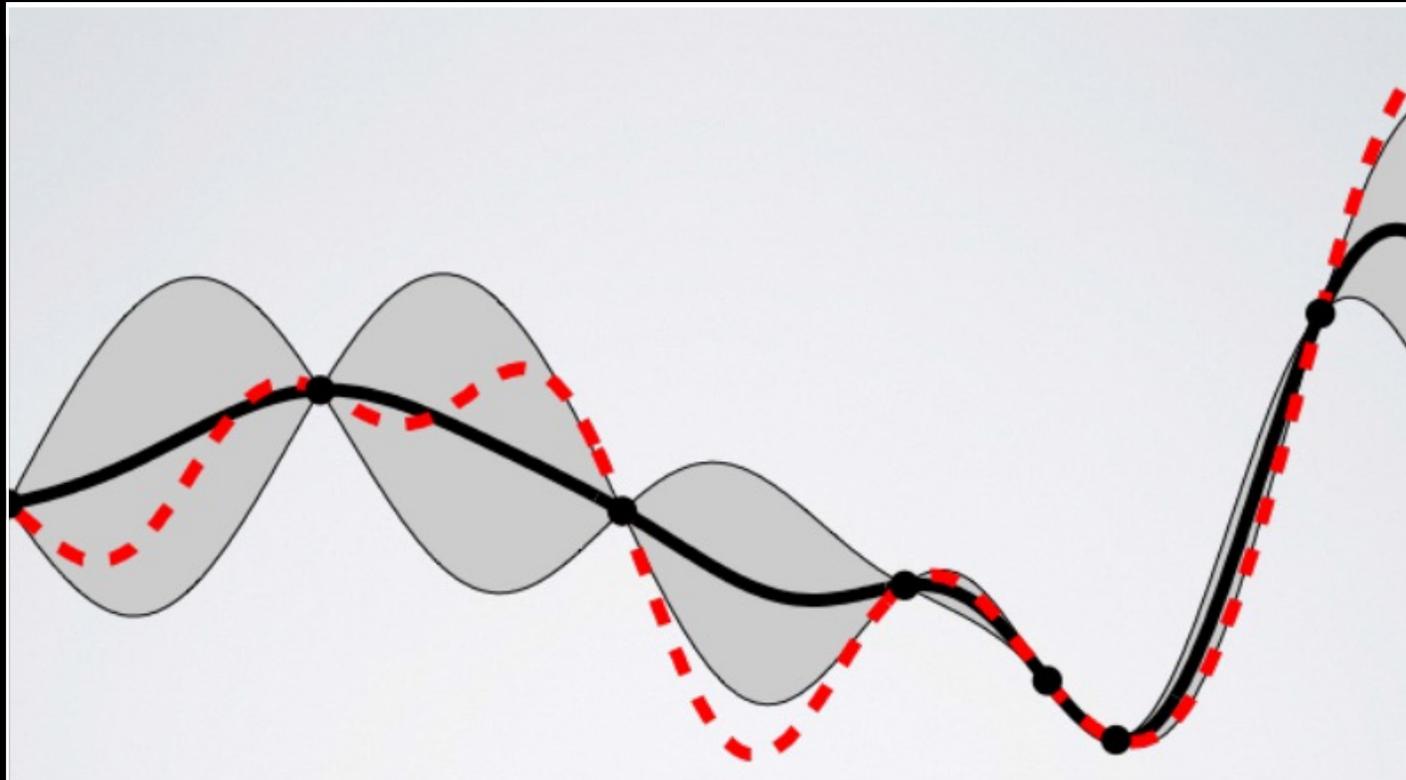


Funzione di
acquisizione

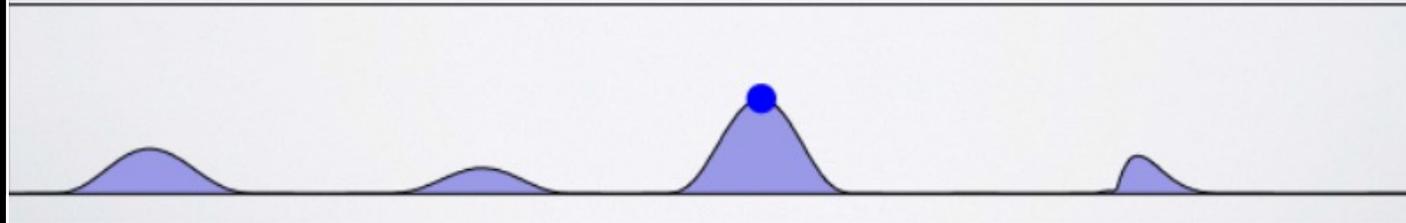


Visualizziamo l'ottimizzazione Bayesiana

Modello



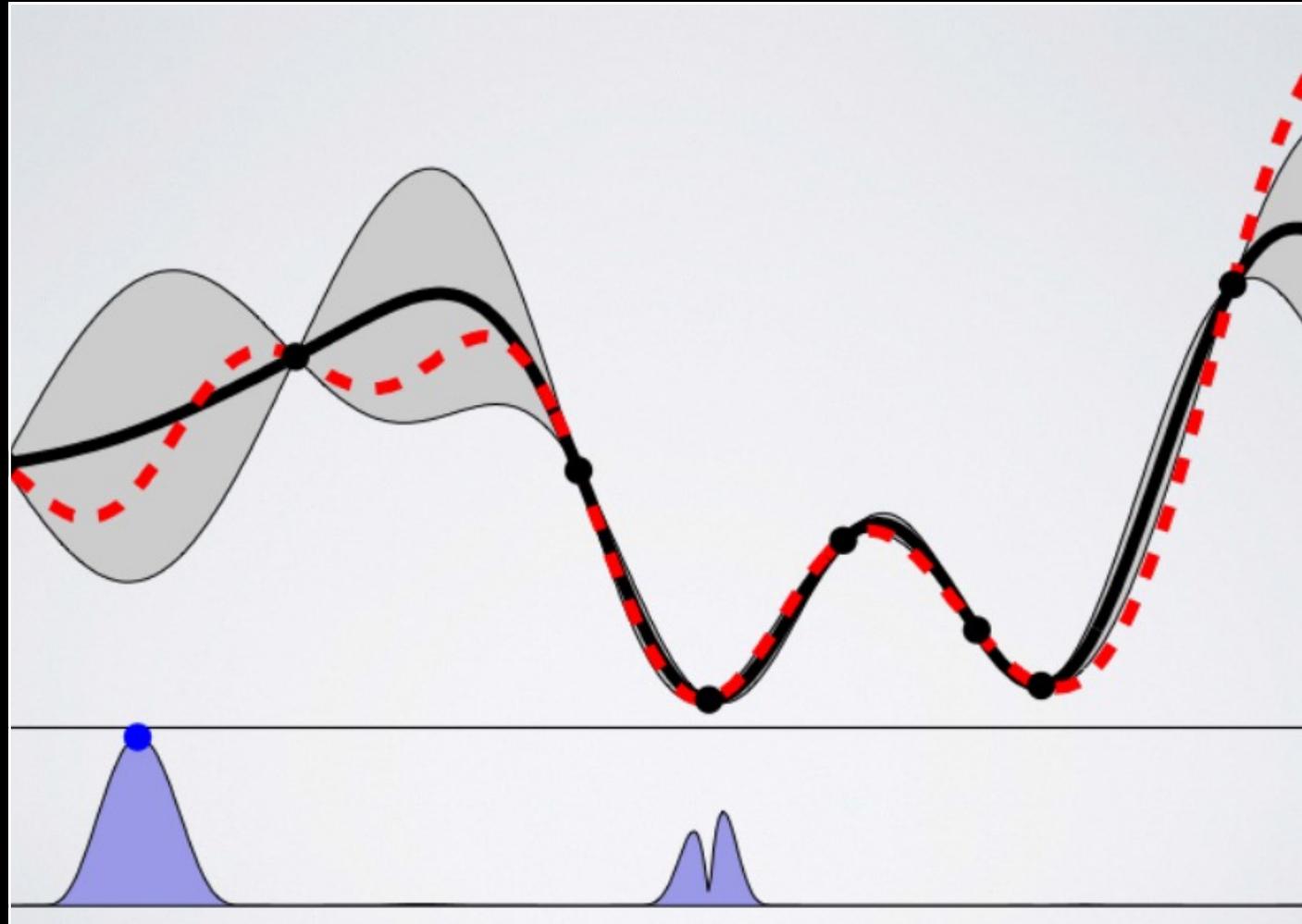
Funzione di
acquisizione



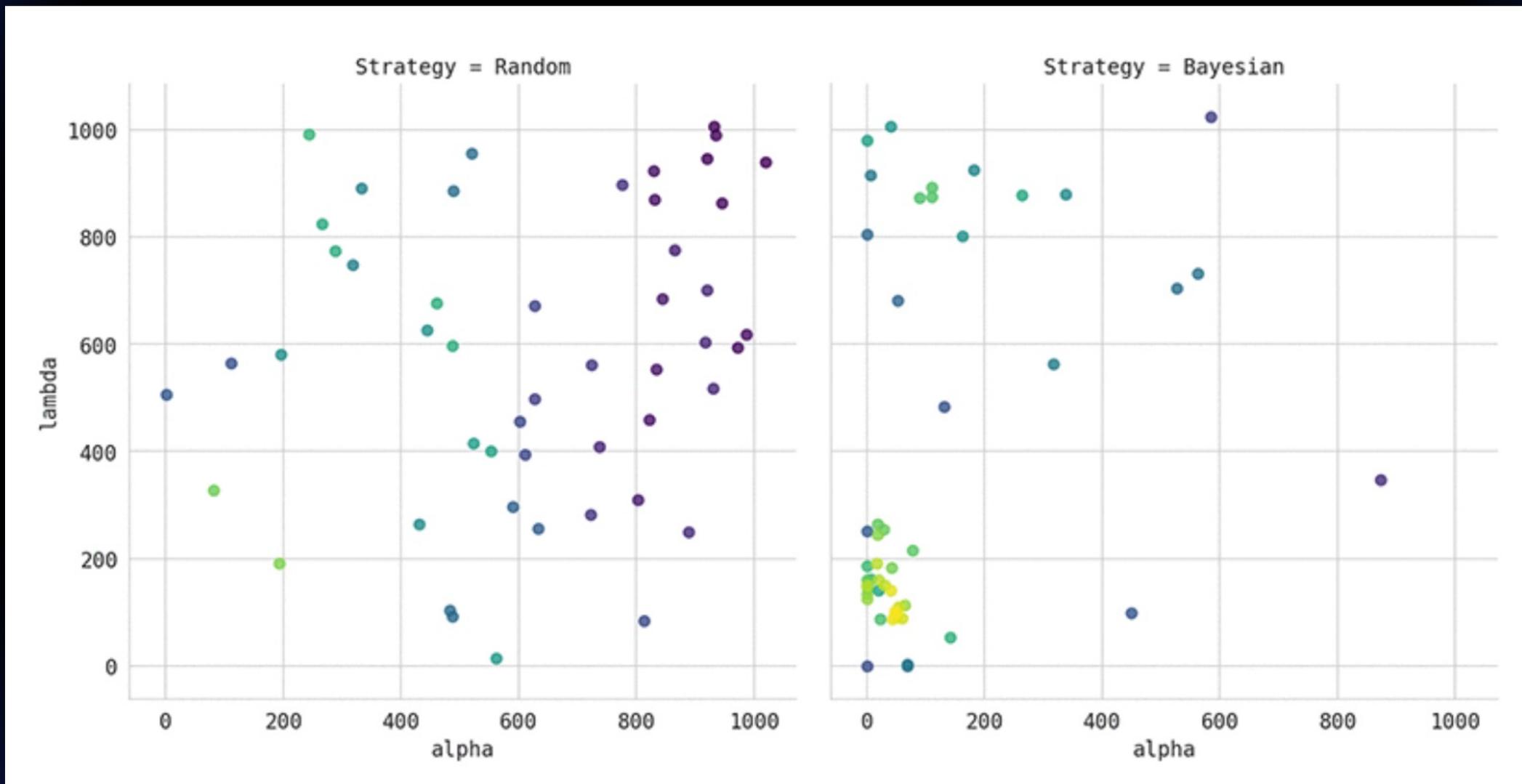
Visualizziamo l'ottimizzazione Bayesiana

Modello

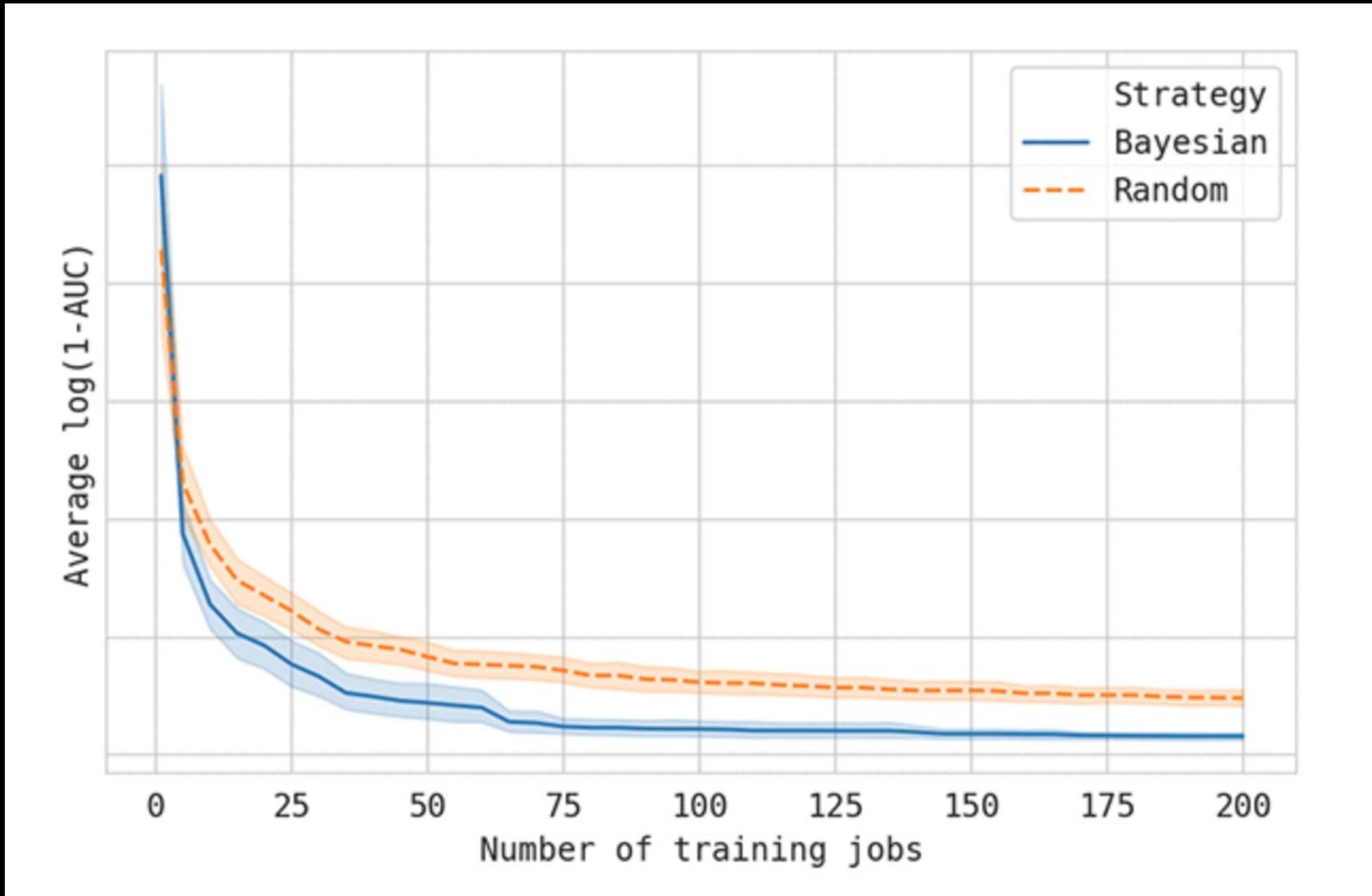
Funzione di
acquisizione



Ricerca Bayesiana e ricerca casuale



Bayesian vs. casuale: velocità di convergenza



Conclusioni

- **Statistica avanzata:** previsioni precise e veloci a costi ridotti
- **Tuning degli iperparametri:** cruciale nei modelli di intelligenza artificiale
- **Competenze statistiche:** essenziali guidare il successo aziendale nell'era dell'intelligenza artificiale

Grazie

Valerio Perrone

<https://www.linkedin.com/in/valerio-perrone/>

Articolo KDD'21 - Amazon SageMaker Automatic Model Tuning: Scalable Gradient-Free Optimization: <https://dl.acm.org/doi/abs/10.1145/3447548.3467098>