

Transfer Learning e Fine-Tuning



Transfer Learning e Fine-Tuning

Premessa

Trattando con le reti neurali, è comune imbattersi in problemi quali:

- *I dati che si hanno a disposizione non sono insufficienti.*
- *L'addestramento richiede troppo tempo.*
- *Le prestazioni non soddisfano i requisiti.*
- *Il modello creato non funziona correttamente.*

Per affrontare questi problemi è possibile sfruttare tecniche quali:

- **Transfer Learning.**
- **Fine-Tuning.**

In entrambi questi casi:

Si sfrutterà la conoscenza di un modello già addestrato.



Transfer Learning e Fine-Tuning

Premessa

Transfer Learning

Consiste nell'**utilizzare un modello pre-addestrato**, *pre-trained*, di una rete nota.

Il pre-addestramento è stato fatto, in genere, su di un **dataset di dimensioni superiori** a quello legato al proprio problema e **di tipologia simile**.

A questo modello vengono **aggiunti layer** nella parte terminale.

L'architettura iniziale viene resa non addestrabile, *frozen*, mentre **i nuovi layer saranno addestrati** a risolvere il task di partenza.



Transfer Learning e Fine-Tuning

Premessa

Fine-Tuning

Consiste nell'**utilizzare un modello pre-addestrato** di una rete nota.

A questo modello **è applicata la tecnica del *transfer learning*** qui accanto descritta.

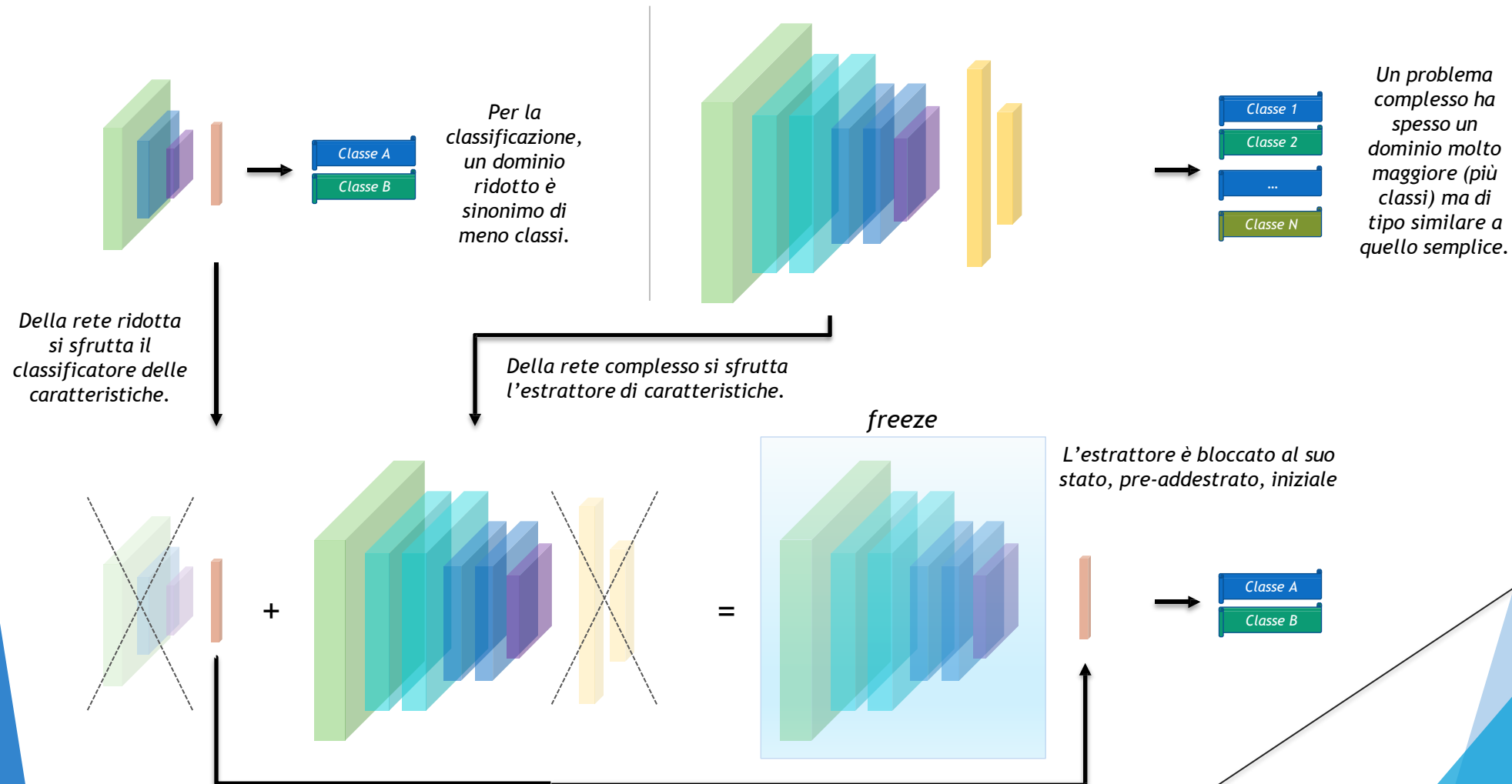
Il modello ottenuto dall'addestramento **viene totalmente sbloccato, unfreezed**, e riaddestrato sul task di partenza.

L'idea è quella di **iniziare l'addestramento da un punto di partenza stabile** ed ottimale.



Transfer Learning e Fine-Tuning

Idealmente





Transfer Learning e Fine-Tuning

Transfer Learning: per step

Quando i dati sono insufficienti per un addestramento affidabile:

1. Ottenere un modello pre-addestrato su dati simili: modelli addestrati su **ImageNet**, ad esempio, hanno una conoscenza generale delle caratteristiche del mondo reale.
2. Eseguirne il **freeze** dei pesi. I pesi diventeranno non-addestrabili.
Si evita quindi che la conoscenza che contengono venga annullata dall'esecuzione del transfer learning.
3. Aggiungere nuovi layers addestrabili nella parte terminale del modello.
4. Addestrare il nuovo modello sul nuovo dataset.



Transfer Learning e Fine-Tuning

Fine-Tuning: per step

A fronte di un *transfer learning*, il *fine-tuning*:

- ▶ *Influenza la conoscenza appresa da un modello su un dataset più ampio.*
- ▶ *Riadatta un modello su un nuovo dominio.*
- ▶ *Migliora le performance su un dato problema / task.*
- ▶ *Riduce il tempo e le risorse di calcolo richieste per addestrare un modello da zero.*

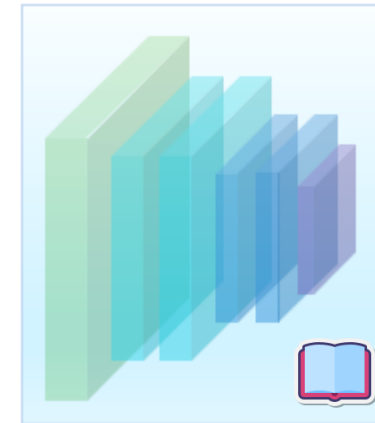
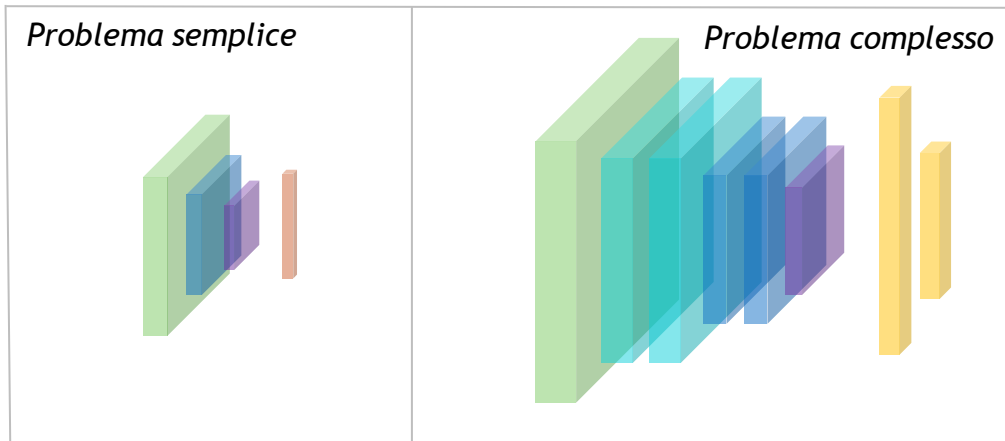
In seguito ai passi di **transfer learning**, si procede:

1. Eseguendo l'*unfreeze* del modello; permettendone l'addestramento completo.
2. Addestrando l'intero modello ed i suoi parametri sullo specifico task iniziale.

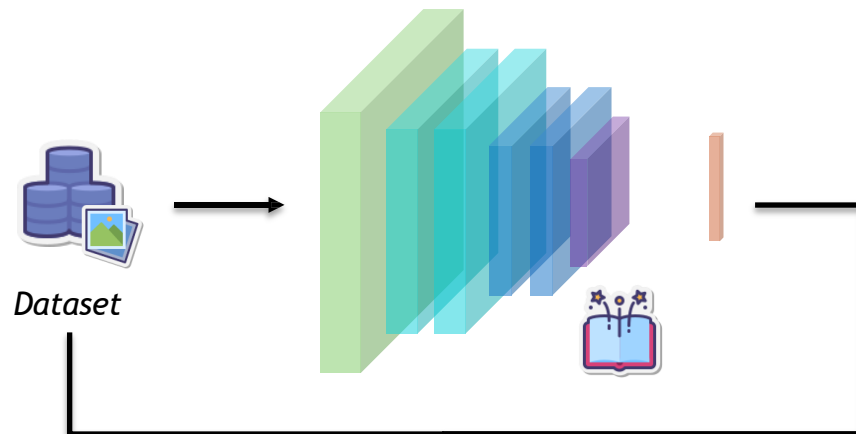


Transfer Learning e Fine-Tuning

Idealmente



Col **transfer learning** si ha un modello combinato il cui estrattore di feature è vincolato alla precedente conoscenza. Si assume quindi che **le caratteristiche apprese dalla conoscenza passata siano sufficienti**.



Qualora si voglia modificare la conoscenza e riadattarla al meglio al dominio semplificato, si riesegue un **training loop completo**. Nulla è **frozen**. La conoscenza è **nuova**.



Transfer Learning e Fine-Tuning

CNN Feature Extractor

Un'altra tecnica che si può utilizzare nel *transfer learning* è sfruttare la parte convoluzionale di un modello pre-addestrato per estrarre le features dal dataset di input.

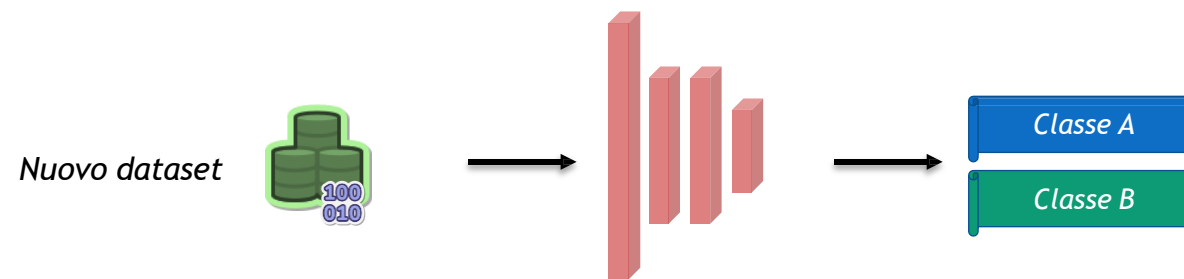
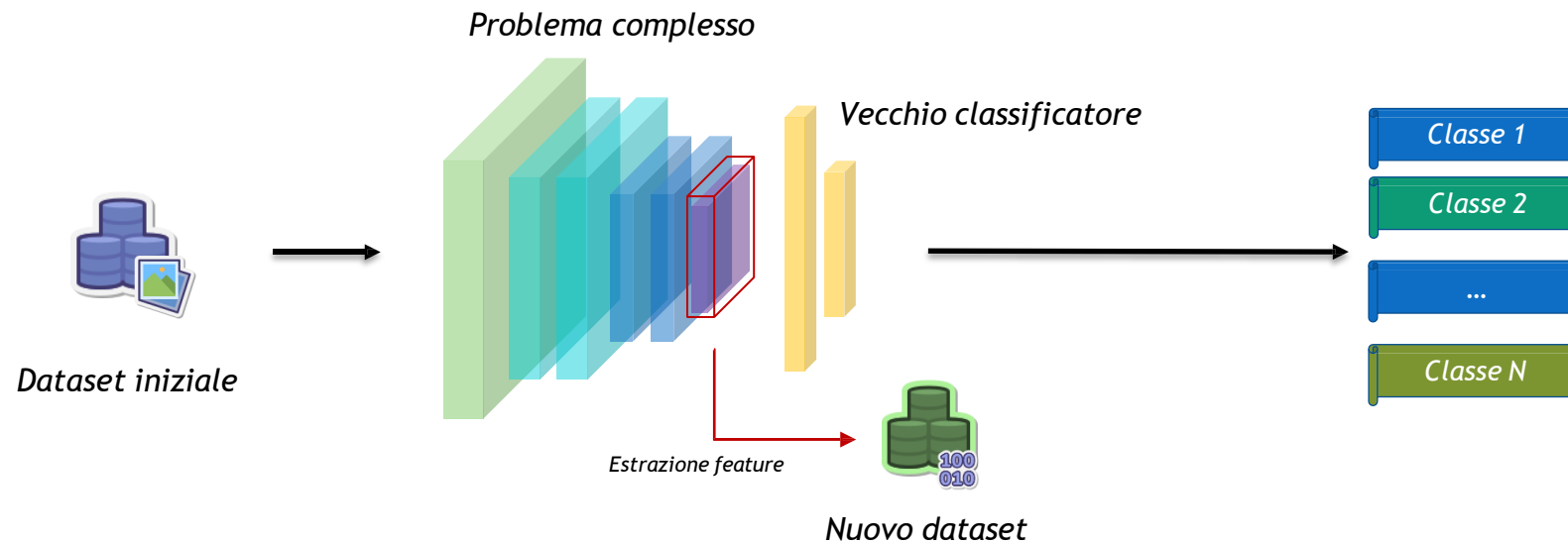
Nello specifico:

1. Si estraggono le features da uno specifico layer della **CNN** del modello pre-addestrato: *l'ultimo layer convoluzionale o uno strato completamente connesso*.
2. Le features estratte sul dataset di training/validazione vengono usate per creare un nuovo dataset, dove **il campione è la feature** e **l'etichetta è la stessa** del dataset iniziale.
3. Si addestra un nuovo classificatore usando questo dataset.
4. Anche per il dataset di test, ad essere classificati non saranno i dati originali ma la conversione dei dati in feature estratte dalla **CNN**.



Transfer Learning e Fine-Tuning

Idealmente



Nuovo classificatore: può essere un architettura di ML, una semplice SVM...



Transfer Learning e Fine-Tuning

CNN Feature Extractor

L'estrazione delle features ha il grande vantaggio che bisogna applicare il modello (**CNN**) solo una volta sui nuovi dati, con l'effetto di trasformarli in un nuovo dataset.

Lo svantaggio principale di questo sistema è che le features non possono essere modificate e la compatibilità tra i nuovi dati e quelli su cui era stato pre-addestrato il modello non è garantita.



Transfer Learning e Fine-Tuning

Un esempio

Di seguito un riferimento utile per l'accesso ad esercitazioni e soluzioni sviluppate con il framework **PyTorch**:

[GitHub - udacity: Projects and exercises](#)

Di seguito un riferimento specifico ad un esempio di transfer learning:

[GitHub - udacity: Transfer Learning](#)

Proviamo?

