

A* to train Neural Networks

Introduzione

- **Descrizione del progetto:** Ottimizzazione di una rete neurale tramite algoritmo A*.
 - **Obiettivo del progetto:** Esplorare configurazioni sub-ottimali di pesi discreti che riducono progressivamente l'errore quadratico medio (MSE) in una rete neurale.
-

Algoritmo Proposto

1. Inizializzazione

Creare una rete neurale con una struttura di base (multi-layer perceptron). I pesi sono inizializzati randomicamente con valori discreti compresi tra -1.000 e 1.000.

2. Funzione di Ricerca A*

➤ Definizione dello Spazio degli Stati

Nell'algoritmo A*, ogni stato rappresenta una specifica configurazione dei pesi della rete neurale; più precisamente, uno stato è un vettore che contiene i valori discreti attuali dei pesi.

➤ Euristic

La funzione euristica valuta ogni stato sulla base della funzione di perdita (loss function), in questo caso l'errore quadratico medio (MSE) calcolato sul training set.

Criteri di valutazione dell'euristica:

- Valore basso dell'MSE: La configurazione di pesi è promettente e vale la pena esplorare ulteriori vicini simili. MSE basso → Euristic
- Valore alto dell'MSE: La configurazione di pesi è meno promettente, e A* tenderà a non esplorare ulteriori stati vicini. MSE alto → Euristic

➤ Azioni Applicabili

Le azioni che portano da uno stato all'altro sono modifiche discrete ai pesi della rete neurale. Le azioni possibili includono:

- Aumentare il peso di 1 unità.
- Ridurre il peso di 1 unità.

3. Procedura

Si applicano le azioni descritte per generare nuovi stati vicini. Ogni volta che viene generata una nuova configurazione di pesi, si esegue un forward attraverso la rete, durante il quale l'output prodotto viene confrontato con il target desiderato, misurando il Mean Squared Error (MSE). Successivamente, l'algoritmo A* seleziona lo stato più promettente, ossia quello con il valore euristico più basso.

4. Criterio di Stop

A* continuerà a esplorare lo spazio di ricerca generando nuovi stati (configurazioni di pesi) finché si raggiunge uno dei seguenti criteri di stop:

- **Numero massimo di iterazioni** raggiunto.
- **Numero massimo di stati esplorati** raggiunto.

Esperimenti e Risultati

- Utilizzare un dataset per testare l'algoritmo.
- Le prestazioni del modello ottimizzato con A* verranno confrontate con un modello addestrato tramite la classica **discesa del gradiente** per valutare differenze di precisione e tempi di convergenza.

Strumenti Utilizzati

- **Python** (libreria **PyTorch** per la costruzione della rete e l'implementazione dell'algoritmo A*).
- **NumPy e Matplotlib** per il supporto all'elaborazione dei dati e la visualizzazione dei risultati.

PROGETTO AI

Sofia Albini