

Tecniche di Predizione Dual Trauma

Rete Neurale

Le reti neurali si fondano su algoritmi in grado di imitare la struttura neuronale del cervello umano, creando, appunto, reti neurali organizzate in diversi strati, dove ogni strato calcola i valori per quello successivo, affinché l'informazione venga via via processata in maniera sempre più completa e profonda.

In linea generale, le reti neurali si compongono di tre strati, capaci di coinvolgere migliaia di neuroni e decine di migliaia di connessioni: lo strato di ingresso (I – Input), lo strato “nascosto” e lo strato di uscita (O – Output).

Il primo è quello che ha il compito di ricevere ed elaborare segnali e dati provenienti dall'esterno; il secondo strato, chiamato strato H – Hidden, ha in carico il processo di elaborazione vero e proprio, mentre lo strato di uscita raccoglie i risultati dell'elaborazione dello strato H, i quali vengono poi adattati alle richieste del successivo livello della rete neurale.

Il processo prevede, in sostanza, che quanto viene prodotto “in uscita” dal primo strato di neuroni artificiali faccia poi da input allo strato successivo e così via, in un ciclo continuo.

Decision Tree

Un decision tree è un sistema con n variabili in input e m variabili in output. Le variabili in input (attributi) sono derivate dall'osservazione dell'ambiente. Le ultime variabili in output, invece, identificano la decisione / azione da intraprendere.

Il processo decisionale è rappresentato con un albero logico rovesciato dove ogni nodo è una funzione condizionale.

Ogni nodo verifica una condizione (test) su una particolare proprietà dell'ambiente (variabile) e ha due o più diramazioni verso il basso in funzione. Il processo consiste in una sequenza di test. Comincia sempre dal nodo radice, il nodo genitore situato più in alto nella struttura, e procede verso il basso. A seconda dei valori rilevati in ciascun nodo, il flusso prende una direzione oppure un'altra e procede progressivamente verso il basso.

La decisione finale si trova nei nodi foglia terminali, quelli più in basso. In questo modo, dopo aver analizzato le varie condizioni, l'agente giunge alla decisione finale.

SVM con Funzioni RBF, Polinomiale e Lineare

Il Support Vector Machine (SVM) è un algoritmo di apprendimento automatico supervisionato che può essere utilizzato sia per scopi di classificazione che di regressione.

L'algoritmo SVM ottiene la massima efficacia nei problemi di classificazione binari come Dual Trauma, anche se può venire utilizzato per problemi di classificazione multiclasse.

Il Support Vector Machine ha l'obiettivo di identificare l'iperpiano che meglio divide i vettori di supporto in classi. Per farlo esegue i seguenti step:

- Cerca un iperpiano linearmente separabile o un limite di decisione che separa i valori di una classe dall'altro. Se ne esiste più di uno, cerca quello che ha margine più alto con i vettori di supporto, per migliorare l'accuratezza del modello.
- Se tale iperpiano non esiste, SVM utilizza una mappatura non lineare per trasformare i dati di allenamento in una dimensione superiore (se siamo a due dimensioni, valuterà i dati in 3 dimensioni). In questo modo, i dati di due classi possono sempre essere separati da un iperpiano, che sarà scelto per la suddivisione dei dati.

L'SVM è stato usato in Dual Trauma in tre diverse configurazioni, utilizzando prima una funzione Radial Basis Function (RBF), una Polinomiale ed infine una Lineare.

Modelli Lineari Lasso e Bayesiano

- **Lasso:** In statistica e machine learning, il lasso (operatore di restringimento e selezione minimo assoluto; anche Lasso o LASSO) è un metodo di analisi di regressione che esegue sia la selezione delle variabili che la regolarizzazione al fine di migliorare l'accuratezza della previsione e l'interpretabilità del modello statistico risultante.
Sebbene originariamente definita per la regressione lineare, la regolarizzazione del lasso è facilmente estesa ad altri modelli statistici inclusi modelli lineari generalizzati, equazioni di stima generalizzate, modelli di rischio proporzionale e stimatori. La capacità di Lasso di eseguire la selezione di sottoinsiemi si basa sulla forma del vincolo e ha una varietà di interpretazioni anche in termini di geometria, statistica bayesiana e analisi convessa.
- **Bayesiano:** Il classificatore bayesiano è una tecnica di machine learning che può essere usata per classificare oggetti in due o più classi.
Il classificatore è allenato analizzando un set di dati di allenamento per i quali sono fornite le classi corrette. Il classificatore bayesiano può essere usato per determinare le probabilità delle classi alla luce di una serie di osservazioni differenti. L'ipotesi del modello è che le variabili indipendenti, in relazione alla classe, diventino condizionatamente indipendenti.

Gradient Boosting Classification

Il Gradient Boosting è una implementazione dell'algoritmo di Boosting, dal quale eredita la logica di costruzione di uno Strong Learner. I modelli che vengono utilizzati sono, generalmente, alberi decisionali e il suo scopo è quello di minimizzare una generica funzione di costo.

Il Boosting è un meta-algoritmo iterativo che detta le linee guida sul come collegare tra loro un insieme di Weak Learner per creare uno Strong Learner. Il segreto del successo di questo paradigma si cela nella

costruzione iterativa degli Strong Learner, nella quale ogni step corrisponde all'inserimento di un Weak Learner che ha lo scopo di "aggiustare il tiro" partendo dai risultati ottenuti dai suoi predecessori. In fase di previsione il risultato dello Strong Learner finale sarà la somma dei risultati dei Weak Learner in esso contenuti.

Il concetto non è dissimile dal costruire la formula di una funzione complessa sommando tra loro un insieme di funzioni elementari.

Essendo i Learner stessi delle funzioni matematiche è possibile operare nella stessa maniera, sommandoli tra di loro in modo da aggiustare l'andamento dello Strong Learner finale e modellarlo nella forma che vogliamo. Gli algoritmi di Boosting sono iterativi, ossia costruiscono lo Strong Learner in un ciclo nel quale, ad ogni iterazione, viene costruito un Weak Learner che viene aggiunto a quelli già presenti per comporre un modello. Da questo punto in poi si definirà come modello la composizione di un certo numero di Weak Learner. Il modello costruito all'ultima iterazione corrisponderà allo Strong Learner finale.

Nearest Neighbors Classification

Il nearest neighbors (k-NN) è un algoritmo utilizzato nel riconoscimento di pattern per la classificazione di oggetti basandosi sulle caratteristiche degli oggetti vicini a quello considerato. L'input è costituito dai k esempi di addestramento più vicini nello spazio delle funzionalità. L'output è un'appartenenza a una classe. Un oggetto è classificato da un voto di pluralità dei suoi vicini, con l'oggetto assegnato alla classe più comune tra i suoi k vicini più vicini (k è un numero intero positivo, tipicamente piccolo). Se $k = 1$, l'oggetto viene semplicemente assegnato alla classe di quel singolo vicino più prossimo.

Gaussian Process Classifier

Il Gaussian Process Classifier è un algoritmo di machine learning di classificazione.

I processi gaussiani sono una generalizzazione della distribuzione di probabilità gaussiana e possono essere utilizzati come base per sofisticati algoritmi di apprendimento automatico non parametrico per la classificazione e la regressione.

Sono un tipo di modello del kernel, come gli SVM, ma a differenza degli SVM, sono in grado di prevedere probabilità di appartenenza a classi altamente calibrate, sebbene la scelta e la configurazione del kernel utilizzato al centro del metodo possano essere difficili.

I gaussian process, sono una generalizzazione della distribuzione di probabilità gaussiana (ad esempio la funzione a campana).

Le funzioni di distribuzione di probabilità gaussiane riassumono la distribuzione di variabili casuali, mentre i processi gaussiani riassumono le proprietà delle funzioni, ad es. i parametri delle funzioni. In quanto tale, si può pensare ai processi gaussiani come un livello di astrazione indiretto al di sopra delle funzioni gaussiane.

AdaBoos

Adaboost è un algoritmo di ensemble learning che utilizza il metodo boosting. È utilizzato nel machine learning. AdaBoost formula H ipotesi tramite l'algoritmo di ensemble boosting a partire da un training set di N esempi. A ciascuna ipotesi assegna un peso z per misurare la sua efficacia. Infine, l'algoritmo elabora l'ipotesi finale a maggioranza pesata.

Stochastic Gradient Descent

La stochastic gradient descent, SGD, è un metodo iterativo per l'ottimizzazione di funzioni differenziabili, approssimazione stocastica del metodo di discesa del gradiente (GD) quando la funzione costo ha la forma di una somma. SGD opera similmente a GD ma, ad ogni iterazione, sostituisce il valore esatto del gradiente della funzione costo con una stima ottenuta valutando il gradiente solo su un sottinsieme degli addendi. È ampiamente usato per l'allenamento di una varietà di modelli probabilistici e modelli di apprendimento automatico, come SVM, regressione logistica e modelli grafici. In combinazione con il metodo di retropropagazione dell'errore, è lo standard de facto per l'allenamento delle reti neurali.