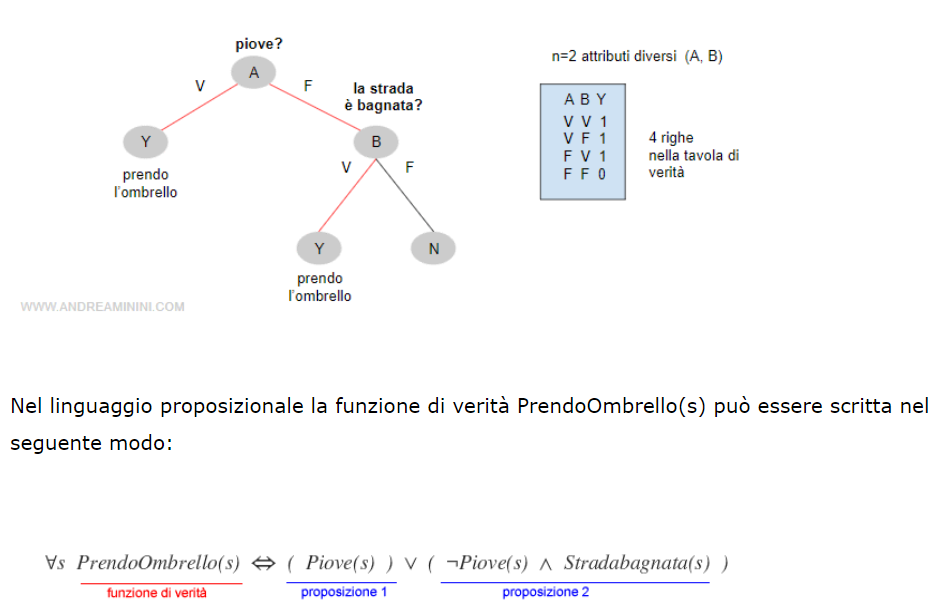
Nel [machine learning](https://it.wikipedia.org/wiki/Machine_learning) un albero di decisione è un modello predittivo, dove ogni nodo interno rappresenta una variabile, un arco verso un nodo figlio rappresenta un possibile valore per quella proprietà e una foglia il valore predetto per la variabile obiettivo a partire dai valori delle altre proprietà, che nell'albero è rappresentato dal cammino (*path*) dal nodo radice (*root*) al nodo foglia. E’ quindi 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. 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.  
In un albero decisionale le variabili possono essere:   
- Variabili discrete: Le variabili hanno valori numerici interi. In questi casi si parla di classificazione. Nota. La forma discreta più semplice è la classificazione booleana dove le variabili hanno soltanto due valori: zero ( falso ) o uno ( vero ).   
- Variabili continue: Le variabili hanno valori numerici reali. Dati due numeri reali qualsiasi, c'è sempre un altro numero intermedio tra questi. In questi casi si parla di regressione.  
 Normalmente un albero di decisione viene costruito utilizzando tecniche di apprendimento a partire dall'insieme dei dati iniziali (*data set*), il quale può essere diviso in due sottoinsiemi: il *training set* sulla base del quale si crea la struttura dell'albero e il *test set* che viene utilizzato per testare l'accuratezza del modello predittivo così creato.

**Pro e contro**

-I vantaggi: Gli alberi logici hanno l'indiscusso vantaggio della semplicità. Sono facili da capire e da eseguire. Rispetto alle reti neurali l'albero decisionale è facilmente comprensibile dagli esseri umani. Pertanto, l'uomo può verificare come la macchina giunge alla decisione. Eventualmente dissentire.  
Inoltre, gli alberi decisionali booleani sono facilmente sviluppabili sotto forma di codice di programmazione, perché possono essere rappresentati con qualsiasi linguaggio proposizionale  
- Lo svantaggio principale è l’alto rischio di overfitting, visto che l’albero procede in una direzione o in un’altra a seconda del valore dell’attributo che sceglie volta volta, dunque rischia di escludere un sottoalbero che invece sarebbe potuto essere corretto. Vediamo un esempio:

****