

# Data Mining & Analytics

## Domande di teoria

Si dia la definizione di apprendimento induttivo da esempi

Si dia la definizione di apprendimento induttivo descrittivo

Si dia la definizione di completezza e consistenza di una ipotesi induttiva

Si parli dei linguaggi di rappresentazione per il machine learning

Si dia la definizione di matrice di confusione, accuratezza, error rate, precision e recall.

Si descrivano la regola della somma, del prodotto e della catena della teoria della probabilità e il teorema di Bayes

Si dia la definizione di ipotesi più generale o uguale ad un'altra

Si dia la definizione di spazio delle versioni e dei confini generale e specifico

Si dia la definizione di bias induttivo

Si descriva l'algoritmo per l'apprendimento degli alberi di decisione

Si parli delle euristiche per l'apprendimento degli alberi di decisione

Come si scelgono i test sugli attributi continui nell'apprendimento degli alberi di decisione?

Come si gestiscono i valori sconosciuti nell'apprendimento degli alberi di decisione?

Come si classificano le istanze che hanno attributi sconosciuti con gli alberi di decisione?

Qual è il bias induttivo di c4.5?

Cosa dice il rasoio di Occam?

Cos'è l'overfitting?

Come si riduce l'overfitting?

Si dia la definizione di indipendenza condizionale

Si dia la definizione di rete bayesiana

Che cos'è il Markov blanket di una variabile in una rete bayesiana?

Cosa significa fare inferenza con una rete bayesiana?

Che cos'è un factor graph?

Si descriva a grandi linee l'algoritmo di belief propagation e si indichi per quali reti è esatto

Si descriva a grandi linee l'algoritmo variable elimination

Si descriva a grandi linee l'algoritmo di inferenza approssimata stochastic simulation

Che cos'è la loopy belief propagation?

Come si fa ad apprendere i parametri di una rete bayesiana da dati completi?

Che cos'è il modello naive Bayes e cosa significa apprendere tale modello?

Si descriva a grandi linee l'algoritmo HGC per apprendere la struttura delle reti bayesiane da dati completi

Come si fa ad apprendere i parametri di una rete bayesiana da dati incompleti?

Si dia la definizione di rete di Markov

Come si fa a capire quando due variabili di una rete di Markov sono indipendenti?

Come si fa l'apprendimento e la classificazione con l'algoritmo k-nearest neighbour?

Si descrivano possibili funzioni di distanza e di similarità tra istanze che sono punti di  $R^n$

Si descriva l'algoritmo nearest neighbour pesato sulla base della distanza

Che cos'è la curse of dimensionality?

Cosa sono le regole di produzione e quali sono gli approcci per apprenderle?

Si descriva a grandi linee l'algoritmo di copertura sequenziale per l'apprendimento di regole di produzione

Che euristiche si possono usare per valutare le regole di produzione?

Che cos'è la beam search?

Come si generano regole di produzione da un albero decisionale?

Che differenza c'è tra copertura sequenziale e simultanea?

Si dia la definizione di learning from entailment in ILP

Si dia la definizione di generalità nel learning from entailment e della theta sussunzione e se ne discutano le proprietà

Si descriva a grandi linee l'algoritmo per il calcolo della least general generalization di due clausole

Che cos'è la relativa subsumption e come si può calcolare la relative least general generalization?

Si descrivano a grandi linee gli algoritmi di ILP bottom-up

Si descrivano a grandi linee gli algoritmi di ILP top-down

Come si può usare un insieme di clausole per classificare una interpretazione?

Si dia la definizione di learning from interpretations in ILP

Si dia la definizione di generalità nel learning from interpretation e della theta sussunzione e se ne discutano le proprietà

Si descriva a grandi linee l'algoritmo ICL

Si descriva a grandi linee l'algoritmo Claudien

Si descriva a grandi linee il percettrone e l'algoritmo per addestrarlo

Come si evita l'overfitting nel percettrone?

Qual è l'interpretazione geometrica dell'apprendimento del perceptrone?

In quali casi l'algoritmo di apprendimento del perceptrone converge?

Si definisca il margine di un dataset

Si descriva a grandi linee il voting perceptron

Si descriva a grandi linee l'averaging perceptron

Quali sono le limitazioni del perceptrone?

Qual è il problema di ottimizzazione per i modelli lineari?

Si descrivano le loss functions e se ne dia qualche esempio

Che cos'è la regolarizzazione?

Cosa sono le convex surrogate loss functions?

Si descrivano alcune possibili funzioni di regolarizzazione per i pesi dei modelli lineari. Che proprietà devono soddisfare?

Si descriva a grandi linee l'algoritmo gradient descent

Qual è il tasso di convergenza dell'algoritmo gradient descent?

Si descriva a grandi linee la closed-form optimization for squared loss nei modelli lineari

Si descrivano i problemi di ottimizzazione hard e soft per le support vector machines

Si descriva il problema di ottimizzazione unconstrained per le support vector machines

Come è fatta una rete neurale a due strati e come calcola il suo output?

Qual è il potere di rappresentazione delle reti neurali a due strati?

Come si sceglie il numero di neuroni nascosti nelle reti a due strati?

Si descriva a grandi linee l'algoritmo di back-propagation per le reti a due strati

Perché si assegnano valori iniziali casuali ai pesi delle reti nell'algoritmo di back-propagation?

Si descriva a grandi linee l'algoritmo di back-propagation per le reti generali

Perché si considerano reti con più di due strati?

Cosa sono le radial basis function networks?

Cosa sono le reti neurali convoluzionali?

A cosa serve il pooling nelle reti neurali convoluzionali?

Si illustri la KDD

Si illustri il Data Mining

Quali sono i compiti del Data Mining?

Si descriva il problema di clustering e si illustri l'algoritmo k-means.

Si descriva il problema di clustering e si illustrino a grandi linee i mixture model e l'algoritmo EM.

Cosa sono le regole associative?

Quali sono le proprietà del vincolo supporto  $\geq$  minsup nelle regole associative?

Si descriva a grandi linee l'algoritmo APRIORI

Come si fanno ad ottenere le regole associative a partire dall'insieme degli itemset grandi?

Si spieghi che cos'è la distribution semantics

Si dia la definizione di Markov Logic Network

Si descriva a grandi linee l'inferenza per programmi logici probabilistici mediante knowledge compilation

Si descriva a grandi linee l'inferenza Monte Carlo per programmi logici probabilistici