

Riconoscimento e recupero dell'informazione per bioinformatica

Dettagli esame

Manuele Bicego

Corso di Laurea in Bioinformatica
Dipartimento di Informatica - Università di Verona

Esempio Esame

⇒ **Esercizio 1 (4 punti).**

Si descriva la regola di decisione di Bayes per la classificazione, evidenziandone vantaggi e svantaggi.

⇒ **Esercizio 2 (4 punti).**

Si descriva la differenza tra metodi di clustering partizionali e metodi di clustering gerarchico, citando un esempio per ognuna delle due categorie.

⇒ **Esercizio 3 (4 punti).**

Si descrivano *in breve* le support vector machines

⇒ **Esercizio 4 (3 punti).**

Dire se il seguente codice matlab implementa una tecnica lineare di riduzione della dimensionalità applicata al vettore X (si prega di motivare la risposta). Qual è la dimensione finale?

```
X = [1; 0.5; 7; 5]; % vettore in input
Y(1) = 2*X(1) + 2*X(2) + X(3)*X(4);
Y(2) = 3*X(3) + sqrt(2)*X(4);
```

Soluzione:

- No, non è una tecnica lineare di riduzione della dimensionalità.
- Motivazione: in una tecnica lineare, le nuove direzioni sono ottenute come combinazione lineare delle direzioni in ingresso: questo non è vero per $Y(1)$, che contiene il termine non lineare $X(3)*X(4)$
- La dimensione finale è 2.

Suggerimenti

⇒ Cercare di fornire risposte precise e complete

Esempio: Si descrivano i vantaggi della regola di classificazione di Bayes

Risposta:

Vantaggi: funziona bene

Risposta non buona!

- imprecisa: occorre specificare che la regola di classificazione di Bayes minimizza la probabilità di errore, cioè rappresenta l'ottimo teorico
- incompleta: occorre anche dire che si basa sulla probabilità a posteriori, che considera sia le informazioni a priori che le osservazioni

Suggerimenti

⇒ Cercare di evitare di scrivere concetti non correlati con la domanda

Esempio: Si descrivano i vantaggi della regola di classificazione di Bayes

Risposta:

La Regola di classificazione di Bayes viene utilizzata per la classificazione, un task specifico della Pattern Recognition. I vantaggi sono [...]

La prima parte non è correlata con la domanda

⇒ **Rischio:** se l'informazione in più è sbagliata (molto sbagliata), si rischia una penalizzazione

Risposta:

La Regola di classificazione di Bayes viene utilizzata per il clustering [...]

L'informazione non richiesta è decisamente sbagliata!