

Reti Neurali

Veronica Piccialli*

Roma 4 Novembre 2014

* Università degli Studi di Roma Tor Vergata

Contesto

Introduzione

- **Contesto**
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

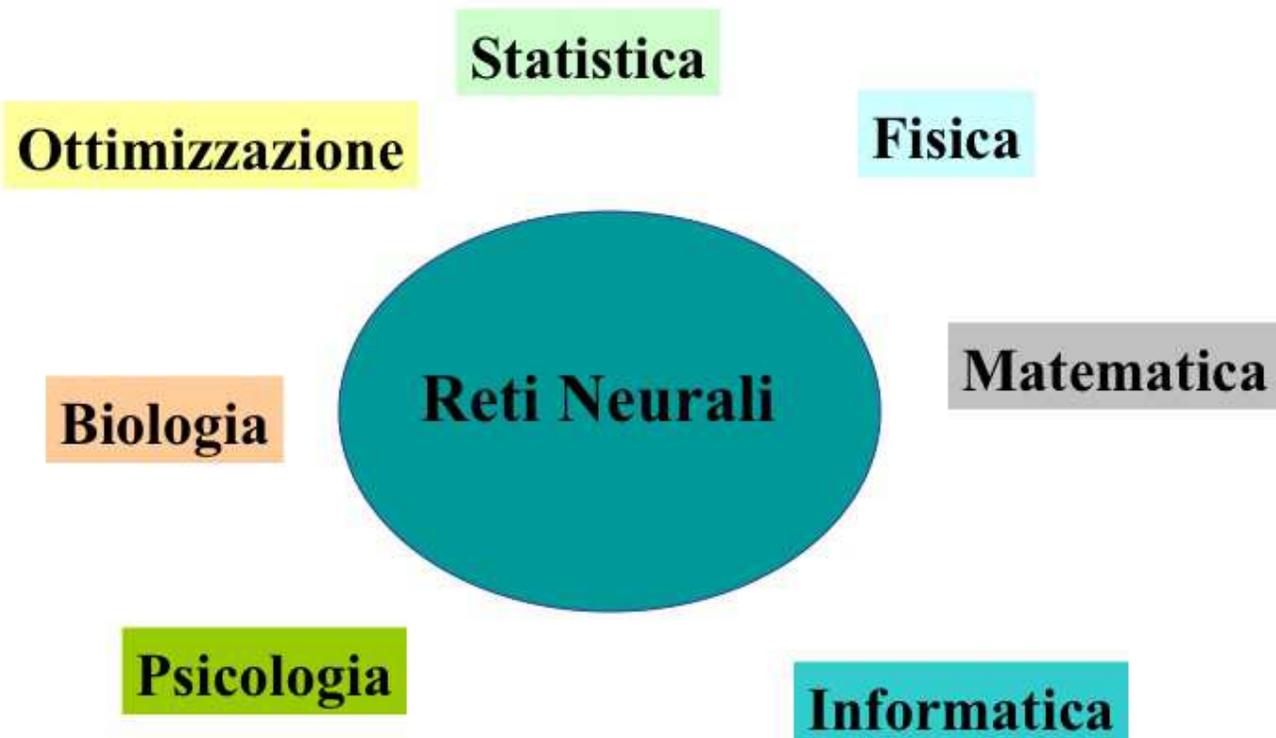
Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF



Idea

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'idea è quella di produrre una macchina o un modello che simuli il comportamento del cervello umano.

Idea

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'idea è quella di produrre una macchina o un modello che simuli il comportamento del cervello umano.

- ◁ Il cervello umano è un calcolatore complesso, non lineare e altamente parallelo: ha la capacità di organizzare le sue unità elementari (neuroni) in modo da effettuare alcuni tipi di calcolo molto più rapidamente di qualunque computer moderno.

Idea

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'idea è quella di produrre una macchina o un modello che simuli il comportamento del cervello umano.

- ◁ Il cervello umano è un calcolatore complesso, non lineare e altamente parallelo: ha la capacità di organizzare le sue unità elementari (neuroni) in modo da effettuare alcuni tipi di calcolo molto più rapidamente di qualunque computer moderno.
- ◁ La vista umana impiega a riconoscere un viso familiare in mezzo a visi sconosciuti pochi microsecondi. Un calcolatore impiega giorni per problemi di riconoscimento di immagini molto più semplici

Idea

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'idea è quella di produrre una macchina o un modello che simuli il comportamento del cervello umano.

- ◁ Il cervello umano è un calcolatore complesso, non lineare e altamente parallelo: ha la capacità di organizzare le sue unità elementari (neuroni) in modo da effettuare alcuni tipi di calcolo molto più rapidamente di qualunque computer moderno.
- ◁ La vista umana impiega a riconoscere un viso familiare in mezzo a visi sconosciuti pochi microsecondi. Un calcolatore impiega giorni per problemi di riconoscimento di immagini molto più semplici
- ◁ Un pipistrello che ha un cervello della dimensione di una prugna ha un radar migliore del radar più avanzato costruito da ingegneri

Il neurone

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

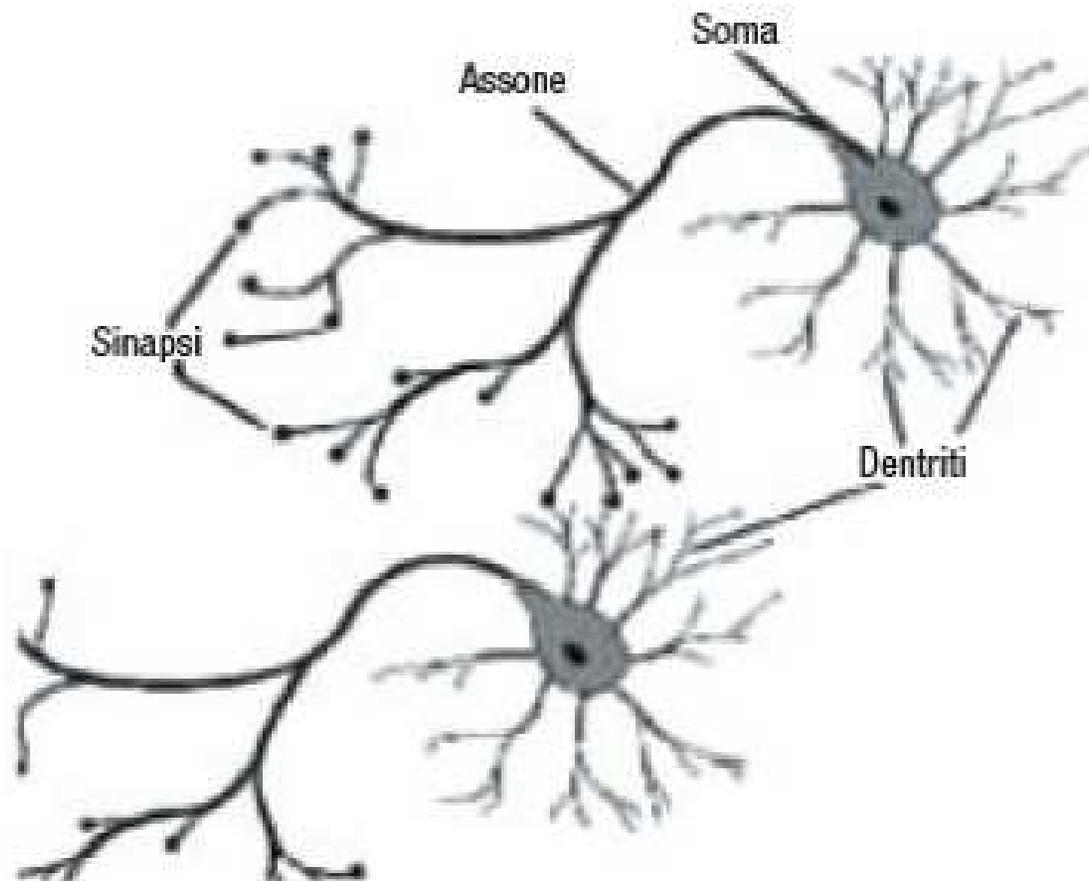
Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF



Neurone e sua struttura cellulare

Rete neurale da un punto di vista fisico

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Rete neurale: processore distribuito ispirato al funzionamento del sistema nervoso negli organismi evoluti, costituito dalla interconnessione di unità computazionali elementari (neuroni) con due caratteristiche fondamentali:

Rete neurale da un punto di vista fisico

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Rete neurale: processore distribuito ispirato al funzionamento del sistema nervoso negli organismi evoluti, costituito dalla interconnessione di unità computazionali elementari (neuroni) con due caratteristiche fondamentali:

- ◁ la **conoscenza** è acquisita dall'ambiente attraverso un processo di **apprendimento** o di **addestramento**.

Rete neurale da un punto di vista fisico

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Rete neurale: processore distribuito ispirato al funzionamento del sistema nervoso negli organismi evoluti, costituito dalla interconnessione di unità computazionali elementari (neuroni) con due caratteristiche fondamentali:

- ◁ la **conoscenza** è acquisita dall'ambiente attraverso un processo di **apprendimento** o di **addestramento**.
- ◁ la conoscenza è immagazzinata nei parametri della rete e, in particolare, nei **pesi** associati alle connessioni

Rete neurale da un punto di vista fisico

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Rete neurale: processore distribuito ispirato al funzionamento del sistema nervoso negli organismi evoluti, costituito dalla interconnessione di unità computazionali elementari (neuroni) con due caratteristiche fondamentali:

- ◁ la **conoscenza** è acquisita dall'ambiente attraverso un processo di **apprendimento** o di **addestramento**.
- ◁ la conoscenza è immagazzinata nei parametri della rete e, in particolare, nei **pesi** associati alle connessioni

Una rete neurale consente di approssimare in un dato contesto applicativo la corrispondenza esistente (o postulata) tra un ingresso e un'uscita di natura opportuna.

A cosa serve

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Data Mining (Wikipedia)

A cosa serve

Introduzione

- Contesto
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- A cosa serve

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Data Mining (Wikipedia)

Ha per oggetto l'estrazione di un sapere o di una conoscenza a partire da grandi quantità di dati (attraverso metodi automatici o semi-automatici) e l'utilizzazione industriale o operativa di questo sapere. Il termine **data mining** (letteralmente: estrazione di dati) è diventato popolare nei tardi anni '90 come versione abbreviata per "estrazione" di informazione utile da insiemi di dati di dimensione cospicua.

Origine DM

Introduzione

Data mining

● **Origine DM**

● DM oggi

● Motivazione

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

1. Crescita notevole degli strumenti e delle tecniche per generare e raccogliere dati (introduzione codici a barre, transazioni economiche tramite carta di credito, dati da satellite o da sensori remoti, servizi on line..)

Origine DM

Introduzione

Data mining

● **Origine DM**

● DM oggi

● Motivazione

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

1. Crescita notevole degli strumenti e delle tecniche per generare e raccogliere dati (introduzione codici a barre, transazioni economiche tramite carta di credito, dati da satellite o da sensori remoti, servizi on line..)
2. Sviluppo delle tecnologie per l'immagazzinamento dei dati, tecniche di gestione di database e data warehouse, supporti più capaci più economici (dischi, CD) hanno consentito l'archiviazione di grosse quantità di dati

Origine DM

Introduzione

Data mining

● Origine DM

● DM oggi

● Motivazione

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

1. Crescita notevole degli strumenti e delle tecniche per generare e raccogliere dati (introduzione codici a barre, transazioni economiche tramite carta di credito, dati da satellite o da sensori remoti, servizi on line..)
2. Sviluppo delle tecnologie per l'immagazzinamento dei dati, tecniche di gestione di database e data warehouse, supporti più capaci più economici (dischi, CD) hanno consentito l'archiviazione di grosse quantità di dati

Questi volumi di dati superano di molto la capacità di analisi dei metodi manuali tradizionali, come le query ad hoc. Tali metodi possono creare report informativi sui dati ma non riescono ad analizzare il contenuto dei report per focalizzarsi sulla conoscenza utile.

DM oggi

Introduzione

Data mining

- Origine DM
- **DM oggi**
- Motivazione

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Oggi il termine data mining ha una duplice valenza:

DM oggi

Introduzione

Data mining

- Origine DM
- **DM oggi**
- Motivazione

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Oggi il termine data mining ha una duplice valenza:

1. **Estrazione**, con tecniche analitiche, di informazione **implicita**, nascosta, da dati già strutturati, per renderla disponibile e direttamente utilizzabile;

DM oggi

Introduzione

Data mining

• Origine DM

● **DM oggi**

• Motivazione

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Oggi il termine data mining ha una duplice valenza:

1. **Estrazione**, con tecniche analitiche, di informazione **implicita**, nascosta, da dati già strutturati, per renderla disponibile e direttamente utilizzabile;
2. Esplorazione ed analisi, eseguita in modo automatico o semiautomatico, su grandi quantità di dati allo scopo di **scoprire pattern** (schemi/regole/configurazioni) caratterizzanti i dati e non evidenti.

DM oggi

Introduzione

Data mining

● Origine DM

● DM oggi

● Motivazione

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Oggi il termine data mining ha una duplice valenza:

1. **Estrazione**, con tecniche analitiche, di informazione **implicita**, nascosta, da dati già strutturati, per renderla disponibile e direttamente utilizzabile;
2. Esplorazione ed analisi, eseguita in modo automatico o semiautomatico, su grandi quantità di dati allo scopo di **scoprire pattern** (schemi/regole/configurazioni) caratterizzanti i dati e non evidenti.

In entrambi i casi i concetti di informazione e di significato sono legati strettamente al dominio applicativo in cui si esegue data mining, cioè un dato può essere interessante o trascurabile a seconda del tipo di applicazione in cui si vuole operare. Questo tipo di attività è cruciale in molti ambiti della ricerca scientifica, ma anche in altri settori (ad es. ricerche di mercato).

Motivazione

Introduzione

Data mining

- Origine DM
- DM oggi
- Motivazione

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Il Data Mining si riferisce all'uso di una varietà di tecniche per identificare pepite di informazione e di conoscenza per il supporto alle decisioni. L' estrazione di tale conoscenza avviene in modo che essa possa essere usata in diverse aree come supporto alle decisioni, previsioni e stime. I dati sono spesso voluminosi ma, così come sono, hanno un basso valore e nessun uso diretto può esserne fatto; è l'informazione nascosta nei dati che è utile (Clementine user guide).

Motivazione

Introduzione

Data mining

- Origine DM
- DM oggi
- Motivazione

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Il Data Mining si riferisce all'uso di una varietà di tecniche per identificare pepite di informazione e di conoscenza per il supporto alle decisioni. L' estrazione di tale conoscenza avviene in modo che essa possa essere usata in diverse aree come supporto alle decisioni, previsioni e stime. I dati sono spesso voluminosi ma, così come sono, hanno un basso valore e nessun uso diretto può esserne fatto; è l'informazione nascosta nei dati che è utile (Clementine user guide).

Motivazione:

esigenza di analizzare e comprendere fenomeni complessi descritti in modo esplicito solo parzialmente e informalmente da insiemi di dati.

Problemi di DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- Classificazione supervisionata
- Esempio 1
- Esempio 2
- Classificazione non supervisionata
- Esempio 1
- Esempio 2
- Regressione
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Classificazione:

Problemi di DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- Classificazione supervisionata
- Esempio 1
- Esempio 2
- Classificazione non supervisionata
- Esempio 1
- Esempio 2
- Regressione
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

□ Classificazione:

◁ supervisionata

Problemi di DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- Classificazione supervisionata
- Esempio 1
- Esempio 2
- Classificazione non supervisionata
- Esempio 1
- Esempio 2
- Regressione
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Classificazione:
 - ◁ supervisionata
 - ◁ non supervisionata

Problemi di DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Classificazione:
 - ◁ supervisionata
 - ◁ non supervisionata
- Regressione

Classificazione e Regressione

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- **Classificazione e Regressione**

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

La **classificazione** individua l'appartenenza ad una classe.

Classificazione e Regressione

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- **Classificazione e Regressione**

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

La **classificazione** individua l'appartenenza ad una classe.

Ad es. un modello potrebbe predire che il potenziale cliente X' risponderà ad un'offerta. Con la classificazione l'output predetto (la classe) è categorico ossia può assumere solo pochi possibili valori come Si, No, Alto, Medio, Basso...

Classificazione e Regressione

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- **Classificazione e Regressione**

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

La **classificazione** individua l'appartenenza ad una classe.

Ad es. un modello potrebbe predire che il potenziale cliente X' risponderà ad un'offerta. Con la classificazione l'output predetto (la classe) è categorico ossia può assumere solo pochi possibili valori come Sì, No, Alto, Medio, Basso...

La **regressione** predice un valore numerico specifico. Ad es. un modello potrebbe predire che il cliente X ci porterà un profitto di Y lire nel corso di un determinato periodo di tempo. Le variabili in uscita possono assumere un numero illimitato (o comunque una grande quantità) di valori. Spesso queste variabili in uscita sono indicate come continue anche se talvolta non lo sono nel senso matematico del termine (ad esempio l'età di una persona)

Classificazione supervisionata

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

● **Classificazione supervisionata**

● Esempio 1

● Esempio 2

● Classificazione non supervisionata

● Esempio 1

● Esempio 2

● Regressione

● Esempio

● Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Pattern:

oggetto descritto da un insieme finito di attributi numerici (caratteristiche/features)

Classificazione supervisionata

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- **Classificazione supervisionata**
- Esempio 1
- Esempio 2
- Classificazione non supervisionata
- Esempio 1
- Esempio 2
- Regressione
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Pattern:

oggetto descritto da un insieme finito di attributi numerici (caratteristiche/features)

Classificazione supervisionata: sono noti **a priori** dei pattern rappresentativi di diverse classi, si vuole determinare un modello matematico che, dato un generico pattern appartenente allo spazio delle caratteristiche, definisca la corrispondente classe di appartenenza

Classificazione supervisionata

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- **Classificazione supervisionata**
- Esempio 1
- Esempio 2
- Classificazione non supervisionata
- Esempio 1
- Esempio 2
- Regressione
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

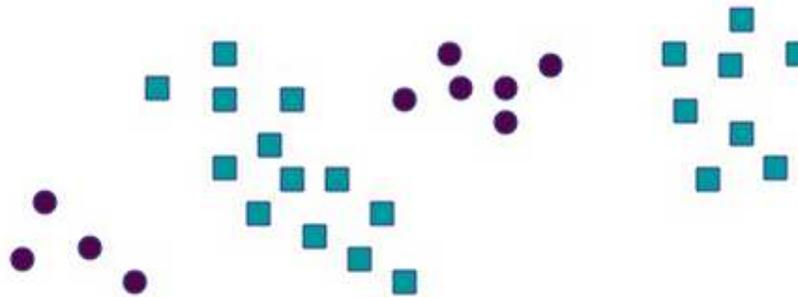
Reti RBF

Addestramento RBF

Pattern:

oggetto descritto da un insieme finito di attributi numerici (caratteristiche/features)

Classificazione supervisionata: sono noti **a priori** dei pattern rappresentativi di diverse classi, si vuole determinare un modello matematico che, dato un generico pattern appartenente allo spazio delle caratteristiche, definisca la corrispondente classe di appartenenza



Esempio 1

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

● **Esempio 1**

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

ESEMPIO DI CLASSIFICAZIONE SUPERVISIONATA

Pattern: cifra manoscritta rappresentata da una matrice di pixel

10 Classi: 0,1,2,3,4,5,6,7,8,9

Esempio 1

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

● Esempio 1

- Esempio 2

● Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

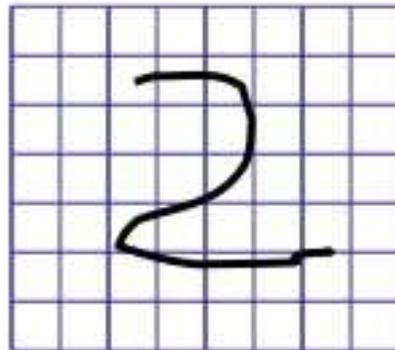
Reti RBF

Addestramento RBF

ESEMPIO DI CLASSIFICAZIONE SUPERVISIONATA

Pattern: cifra manoscritta rappresentata da una matrice di pixel

10 Classi: 0,1,2,3,4,5,6,7,8,9



classificazione



0
1
2
3
4
5
6
7
8
9

Esempio 2

Introduzione

Pattern: elettrocardiogramma di un paziente

Data mining

2 Classi: normale, ischemico

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- **Esempio 2**

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Esempio 2

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- **Esempio 2**

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Pattern: elettrocardiogramma di un paziente

2 Classi: normale, ischemico

Dati a disposizione per un dato paziente:

Esempio 2

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- **Esempio 2**

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Pattern: elettrocardiogramma di un paziente

2 Classi: normale, ischemico

Dati a disposizione per un dato paziente:

1. Tracciato digitale dell'ECG durante una fase normale

Esempio 2

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- Classificazione supervisionata
- Esempio 1
- **Esempio 2**
- Classificazione non supervisionata
- Esempio 1
- Esempio 2
- Regressione
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

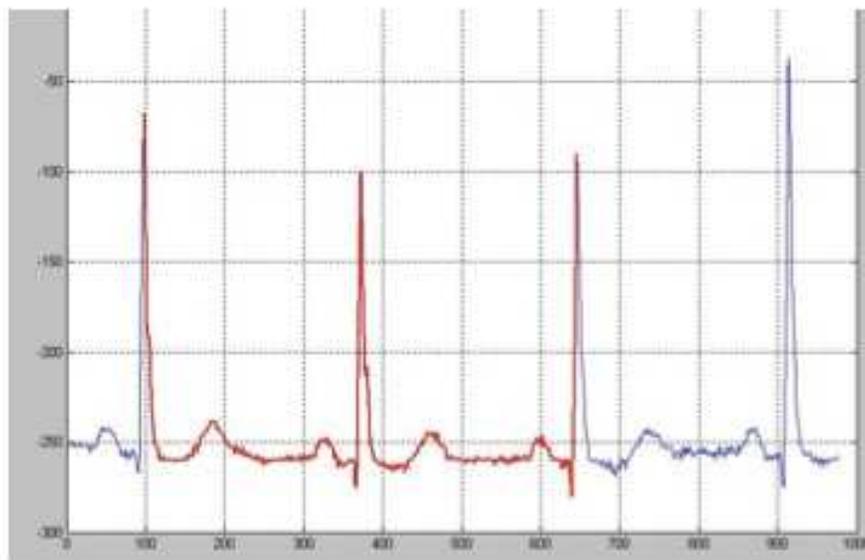
Addestramento RBF

Pattern: elettrocardiogramma di un paziente

2 Classi: normale, ischemico

Dati a disposizione per un dato paziente:

1. Tracciato digitale dell'ECG durante una fase normale
2. Tracciato digitale dell'ECG durante una fase ischemica



Classificazione non supervisionata

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Pattern:

oggetto descritto da un insieme finito di attributi numerici (caratteristiche/features)

Classificazione non supervisionata

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- Classificazione supervisionata
- Esempio 1
- Esempio 2
- **Classificazione non supervisionata**
- Esempio 1
- Esempio 2
- Regressione
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Pattern:

oggetto descritto da un insieme finito di attributi numerici (caratteristiche/features)

Classificazione non supervisionata: non sono noti a priori pattern rappresentativi delle classi; si vuole determinare il numero di classi di “similitudine” e un modello matematico che, dato un generico pattern appartenente allo spazio delle caratteristiche, definisca la corrispondente classe di appartenenza

Classificazione non supervisionata

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- Classificazione supervisionata
- Esempio 1
- Esempio 2
- **Classificazione non supervisionata**
- Esempio 1
- Esempio 2
- Regressione
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

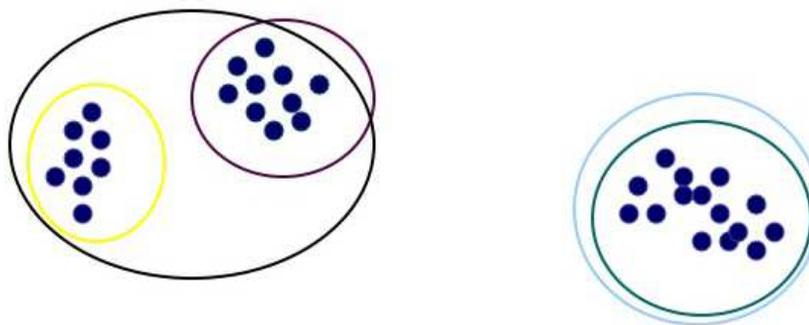
Reti RBF

Addestramento RBF

Pattern:

oggetto descritto da un insieme finito di attributi numerici (caratteristiche/features)

Classificazione non supervisionata: non sono noti a priori pattern rappresentativi delle classi; si vuole determinare il numero di classi di “similitudine” e un modello matematico che, dato un generico pattern appartenente allo spazio delle caratteristiche, definisca la corrispondente classe di appartenenza



Esempio 1

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- Classificazione supervisionata
- Esempio 1
- Esempio 2
- Classificazione non supervisionata
- **Esempio 1**
- Esempio 2
- Regressione
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

ESEMPIO DI CLASSIFICAZIONE NON SUPERVISIONATA

Pattern: cifra manoscritta

Dati disponibili: matrici di pixel di immagini

Obiettivo: raggruppare le immagini in K gruppi i cui elementi presentino caratteristiche “simili”

Esempio 1

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- **Esempio 1**

- Esempio 2

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

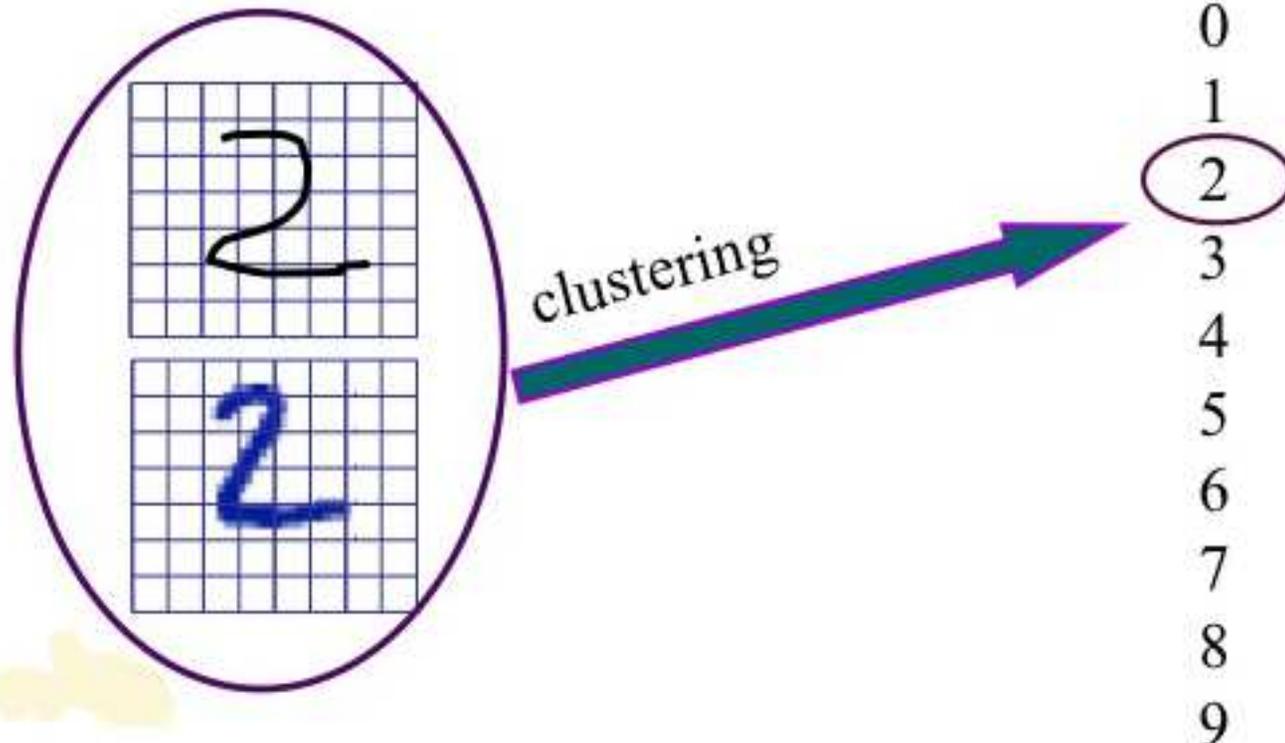
Addestramento RBF

ESEMPIO DI CLASSIFICAZIONE NON SUPERVISIONATA

Pattern: cifra manoscritta

Dati disponibili: matrici di pixel di immagini

Obiettivo: raggruppare le immagini in K gruppi i cui elementi presentino caratteristiche “simili”



Esempio 2

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- **Esempio 2**

- Regressione

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Pattern: paziente afflitto da una determinata patologia e descritto da M fattori clinici (caratteristiche)

Dati disponibili: insieme di N pazienti

Obiettivo: raggruppare i pazienti in K gruppi i cui elementi presentino caratteristiche “simili”

Regressione

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- **Regressione**

- Esempio

- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Pattern:

oggetto descritto da un insieme finito di attributi numerici (caratteristiche/features)

Regressione

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- Classificazione supervisionata
- Esempio 1
- Esempio 2
- Classificazione non supervisionata
- Esempio 1
- Esempio 2
- **Regressione**
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Pattern:

oggetto descritto da un insieme finito di attributi numerici (caratteristiche/features)

Approssimazione o **Regressione**: sono note a priori delle coppie pattern/target rappresentative di una funzione incognita a valori reali; si vuole determinare una funzione analitica che approssimi la funzione incognita

Regressione

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- Classificazione supervisionata
- Esempio 1
- Esempio 2
- Classificazione non supervisionata
- Esempio 1
- Esempio 2
- **Regressione**
- Esempio
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

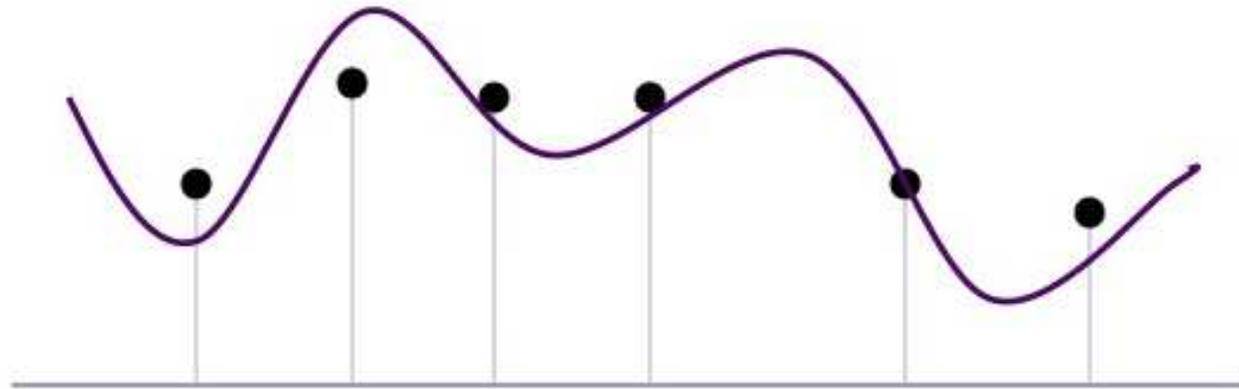
Reti RBF

Addestramento RBF

Pattern:

oggetto descritto da un insieme finito di attributi numerici (caratteristiche/features)

Approssimazione o **Regressione**: sono note a priori delle coppie pattern/target rappresentative di una funzione incognita a valori reali; si vuole determinare una funzione analitica che approssimi la funzione incognita



Esempio

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e Regressione
- Classificazione supervisionata
- Esempio 1
- Esempio 2
- Classificazione non supervisionata
- Esempio 1
- Esempio 2
- Regressione
- **Esempio**
- Metodi per DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

ESEMPIO DI REGRESSIONE

Pattern: vettore di N correnti che circolano in un dispositivo

Dati disponibili: valore del campo magnetico in un determinato punto interno al dispositivo

Obiettivo: determinare una funzione analitica che approssimi il legame funzionale tra il campo magnetico e il valore delle correnti

Metodi per DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- **Metodi per DM**

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Classificazione:

◁ supervisionata

◁ non supervisionata

Regressione

Metodi per DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- **Metodi per DM**

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Classificazione:

◁ supervisionata

◁ non supervisionata

Regressione

Metodi

Metodi per DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- **Metodi per DM**

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Classificazione:

◁ supervisionata

◁ non supervisionata

Regressione

Metodi

1. Clustering

Metodi per DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- **Metodi per DM**

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Classificazione:

◁ supervisionata

◁ non supervisionata

Regressione

Metodi

1. Clustering
2. Reti Neurali

Metodi per DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- **Metodi per DM**

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Classificazione:

◁ supervisionata

◁ non supervisionata

Regressione

Metodi

1. Clustering
2. Reti Neurali
3. Support Vector Machines

Metodi per DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- **Metodi per DM**

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Classificazione:

◁ supervisionata

◁ non supervisionata

Regressione

Metodi

1. Clustering
2. Reti Neurali
3. Support Vector Machines
4. Programmazione logica

Metodi per DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- **Metodi per DM**

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Classificazione:

◁ supervisionata

◁ non supervisionata

Regressione

Metodi

1. Clustering
2. Reti Neurali
3. Support Vector Machines
4. Programmazione logica
5. Alberi di decisione

Metodi per DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- **Metodi per DM**

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Classificazione:

◁ supervisionata

◁ non supervisionata

Regressione

Metodi

1. Clustering
2. Reti Neurali
3. Support Vector Machines
4. Programmazione logica
5. Alberi di decisione



Problemi “difficili”

Metodi per DM

Introduzione

Data mining

Problemi di DM

- Problemi di DM
- Classificazione e

Regressione

- Classificazione supervisionata

- Esempio 1

- Esempio 2

- Classificazione non supervisionata

- Esempio 1

- Esempio 2

- Regressione

- Esempio

- **Metodi per DM**

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Classificazione:

◁ supervisionata

◁ non supervisionata

Regressione

Metodi

1. Clustering
2. Reti Neurali
3. Support Vector Machines
4. Programmazione logica
5. Alberi di decisione



Problemi "difficili"



Metodi di ottimizzazione efficienti

Rete neurale da un punto di vista fisico

Introduzione

Data mining

Problemi di DM

Reti neurali

● Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete
- Architetture
- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Rete neurale: processore distribuito ispirato al funzionamento del sistema nervoso negli organismi evoluti, costituito dalla interconnessione di unità computazionali elementari (neuroni) con due caratteristiche fondamentali:

Rete neurale da un punto di vista fisico

Introduzione

Data mining

Problemi di DM

Reti neurali

● Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete
- Architetture
- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Rete neurale: processore distribuito ispirato al funzionamento del sistema nervoso negli organismi evoluti, costituito dalla interconnessione di unità computazionali elementari (neuroni) con due caratteristiche fondamentali:

- ◁ la **conoscenza** è acquisita dall'ambiente attraverso un processo di **apprendimento** o di **addestramento**.

Rete neurale da un punto di vista fisico

Introduzione

Data mining

Problemi di DM

Reti neurali

● Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete
- Architetture
- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Rete neurale: processore distribuito ispirato al funzionamento del sistema nervoso negli organismi evoluti, costituito dalla interconnessione di unità computazionali elementari (neuroni) con due caratteristiche fondamentali:

- ◁ la **conoscenza** è acquisita dall'ambiente attraverso un processo di **apprendimento** o di **addestramento**.
- ◁ la conoscenza è immagazzinata nei parametri della rete e, in particolare, nei **pesi** associati alle connessioni

Rete neurale da un punto di vista fisico

Introduzione

Data mining

Problemi di DM

Reti neurali

● Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete
- Architetture
- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Rete neurale: processore distribuito ispirato al funzionamento del sistema nervoso negli organismi evoluti, costituito dalla interconnessione di unità computazionali elementari (neuroni) con due caratteristiche fondamentali:

- ◁ la **conoscenza** è acquisita dall'ambiente attraverso un processo di **apprendimento** o di **addestramento**.
- ◁ la conoscenza è immagazzinata nei parametri della rete e, in particolare, nei **pesi** associati alle connessioni

Una rete neurale consente di approssimare in un dato contesto applicativo la corrispondenza esistente (o postulata) tra un ingresso e un'uscita di natura opportuna.

Neuroni

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- **Neuroni**

- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture

- Esempi reti feedforward

- Addestramento

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Sono nodi di una rete orientata provvisti di capacità di elaborazione.

Neuroni

Introduzione

Sono nodi di una rete orientata provvisti di capacità di elaborazione.

Data mining

Problemi di DM

Ricevono in **ingresso** una combinazione di segnali provenienti dall'esterno o da altri nodi e ne effettuano una trasformazione tramite una funzione non lineare detta **funzione di attivazione**.

Reti neurali

- Rete neurale da un punto di vista fisico

- **Neuroni**

- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture

- Esempi reti feedforward

- Addestramento

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Neuroni

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- **Neuroni**

- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture

- Esempi reti feedforward

- Addestramento

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Sono nodi di una rete orientata provvisti di capacità di elaborazione.

Ricevono in **ingresso** una combinazione di segnali provenienti dall'esterno o da altri nodi e ne effettuano una trasformazione tramite una funzione non lineare detta **funzione di attivazione**.

L'uscita viene inviata agli altri nodi o all'uscita della rete tramite connessioni orientate e pesate.

Rete neurale da un punto di vista matematico

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture
- Esempi reti feedforward

- Addestramento
- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Data una funzione $G : X \rightarrow Y$ nota attraverso un insieme di coppie

$$\{(x^p, G(x^p)) : x^p \in X, p = 1, \dots, P\}$$

una rete neurale è un particolare modello di approssimazione di G :

$$F(\cdot, w) : X \rightarrow Y$$

dipendente (in generale in modo **non lineare**) da un vettore di parametri w

Rete neurale da un punto di vista matematico

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico
- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete
- Architetture
- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Data una funzione $G : X \rightarrow Y$ nota attraverso un insieme di coppie

$$\{(x^p, G(x^p)) : x^p \in X, p = 1, \dots, P\}$$

una rete neurale è un particolare modello di approssimazione di G :

$$F(\cdot, w) : X \rightarrow Y$$

dipendente (in generale in modo **non lineare**) da un vettore di parametri w

Da un punto di vista statistico, una rete neurale è un particolare modello di classificazione o di regressione (non lineare), caratterizzato da:

Rete neurale da un punto di vista matematico

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture
- Esempi reti feedforward

- Addestramento
- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Data una funzione $G : X \rightarrow Y$ nota attraverso un insieme di coppie

$$\{(x^p, G(x^p)) : x^p \in X, p = 1, \dots, P\}$$

una rete neurale è un particolare modello di approssimazione di G :

$$F(\cdot, w) : X \rightarrow Y$$

dipendente (in generale in modo **non lineare**) da un vettore di parametri w

Da un punto di vista statistico, una rete neurale è un particolare modello di classificazione o di regressione (non lineare), caratterizzato da:

(i) scelta delle funzioni approssimanti (dipendenza non lineare dai parametri)

Rete neurale da un punto di vista matematico

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture
- Esempi reti feedforward

- Addestramento
- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Data una funzione $G : X \rightarrow Y$ nota attraverso un insieme di coppie

$$\{(x^p, G(x^p)) : x^p \in X, p = 1, \dots, P\}$$

una rete neurale è un particolare modello di approssimazione di G :

$$F(\cdot, w) : X \rightarrow Y$$

dipendente (in generale in modo **non lineare**) da un vettore di parametri w

Da un punto di vista statistico, una rete neurale è un particolare modello di classificazione o di regressione (non lineare), caratterizzato da:

- (i) scelta delle funzioni approssimanti (dipendenza non lineare dai parametri)
- (i) i dati (**training set**) sono discreti cioè la funzione si suppone nota solo su un insieme discreto di punti.

Caratteristiche della rete

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni

- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture

- Esempi reti feedforward

- Addestramento

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Il legame ingresso-uscita realizzato dalla rete dipende da:

Caratteristiche della rete

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni

- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture

- Esempi reti feedforward

- Addestramento

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Il legame ingresso-uscita realizzato dalla rete dipende da:

- tipo di unità elementari (struttura interna più o meno complessa, funzioni di attivazione diverse)

Caratteristiche della rete

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni

- Rete neurale da un punto di vista matematico

- **Caratteristiche della rete**

- Architetture

- Esempi reti feedforward

- Addestramento

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Il legame ingresso-uscita realizzato dalla rete dipende da:

- tipo di unità elementari (struttura interna più o meno complessa, funzioni di attivazione diverse)
- architettura della rete, ossia numero di nodi, struttura e orientamento delle connessioni

Caratteristiche della rete

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni

- Rete neurale da un punto di vista matematico

- **Caratteristiche della rete**

- Architetture

- Esempi reti feedforward

- Addestramento

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Il legame ingresso-uscita realizzato dalla rete dipende da:

- tipo di unità elementari (struttura interna più o meno complessa, funzioni di attivazione diverse)
- architettura della rete, ossia numero di nodi, struttura e orientamento delle connessioni
- valori dei parametri interni associati a unità elementari e connessioni determinati tramite **tecniche di apprendimento**

Architetture

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico
- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete

● **Architetture**

- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Reti **feedforward**: reti acicliche strutturate in diversi **strati**

Architetture

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico
- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete

● Architetture

- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

□ Reti **feedforward**: reti acicliche strutturate in diversi **strati**

1. Perceptron

Architetture

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico
- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete

● Architetture

- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

□ Reti **feedforward**: reti acicliche strutturate in diversi **strati**

1. Perceptron
2. Multilayer Perceptron (MLP)

Architetture

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete

- Architetture

- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

□ Reti **feedforward**: reti acicliche strutturate in diversi **strati**

1. Perceptron
2. Multilayer Perceptron (MLP)
3. Reti di funzioni di base radiali (RBF)

Architetture

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni

- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture

- Esempi reti feedforward

- Addestramento

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

□ Reti **feedforward**: reti acicliche strutturate in diversi **strati**

1. Perceptron
2. Multilayer Perceptron (MLP)
3. Reti di funzioni di base radiali (RBF)

□ Reti **ricorsive**: sono presenti cicli di controreazione; possono essere viste come sistemi dinamici

Esempi reti feedforward

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico
- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete

● Architetture

● Esempi reti feedforward

- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

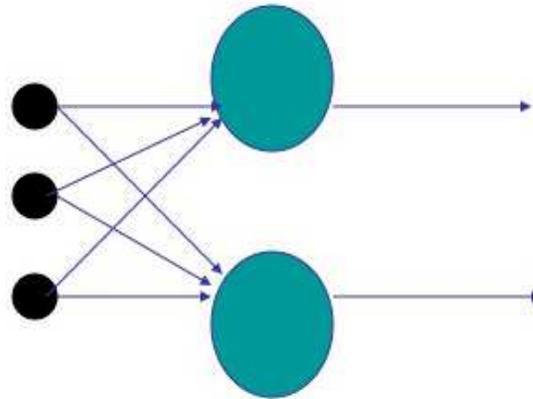
Reti multistrato

Creazione di una rete multistrato

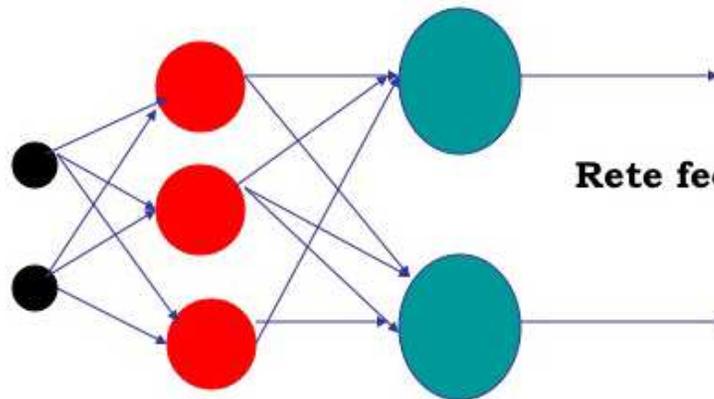
Metodi on-line

Reti RBF

Addestramento RBF



Rete feedforward a 1 strato



Rete feedforward a 2 strati

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture
- Esempi reti feedforward

- **Addestramento**

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'addestramento (apprendimento) è il processo mediante il quale vengono determinati i parametri liberi di una rete neurale. Due paradigmi fondamentali:

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture
- Esempi reti feedforward

- **Addestramento**

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'addestramento (apprendimento) è il processo mediante il quale vengono determinati i parametri liberi di una rete neurale. Due paradigmi fondamentali:

1. addestramento **supervisionato**: i parametri della rete vengono determinati, sulla base di un insieme di addestramento (training set) di esempi, consistenti in coppie pattern/target, **minimizzando una funzione d'errore**

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture
- Esempi reti feedforward

- **Addestramento**

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'addestramento (apprendimento) è il processo mediante il quale vengono determinati i parametri liberi di una rete neurale. Due paradigmi fondamentali:

1. addestramento **supervisionato**: i parametri della rete vengono determinati, sulla base di un insieme di addestramento (training set) di esempi, consistenti in coppie pattern/target, **minimizzando una funzione d'errore**
2. addestramento **non supervisionato**: la rete è dotata di capacità di auto-organizzazione

Overfitting e underfitting

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico

- Neuroni
- Rete neurale da un punto di vista matematico

- Caratteristiche della rete

- Architetture

- Esempi reti feedforward

- Addestramento

- Overfitting e underfitting

- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

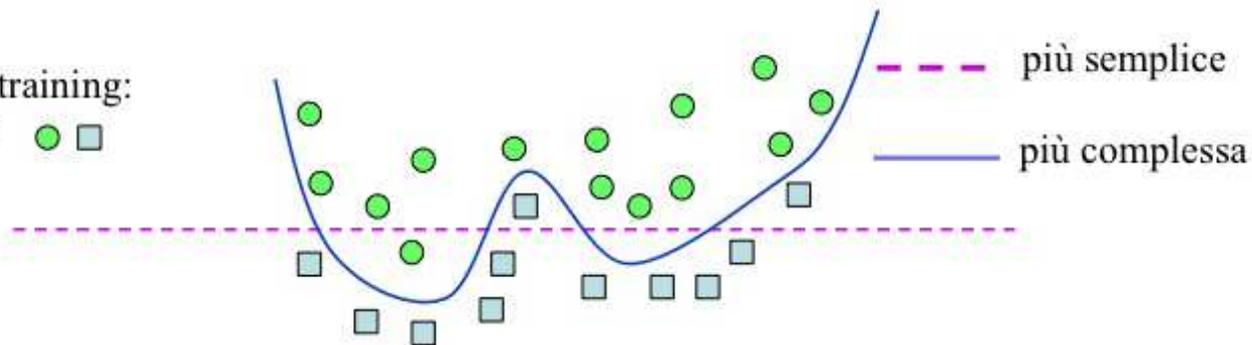
Creazione di una rete multistrato

Metodi on-line

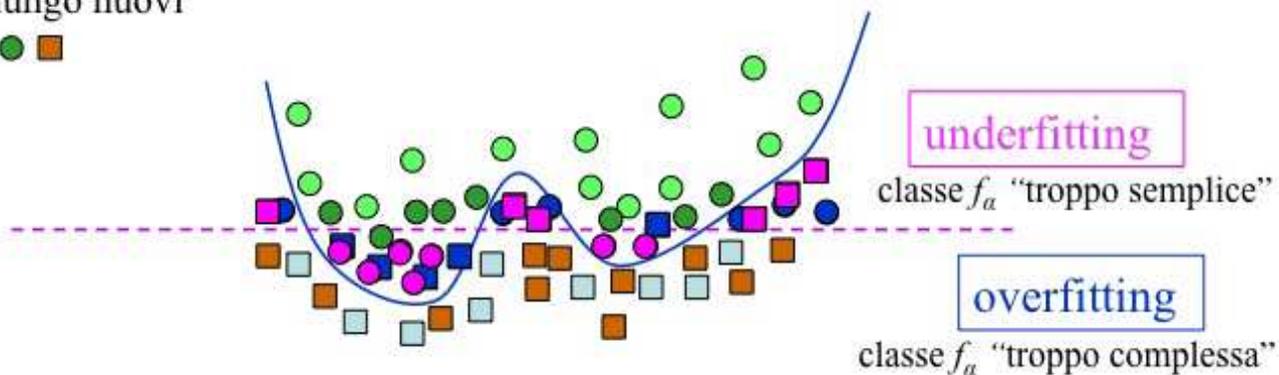
Reti RBF

Addestramento RBF

Dati di training:
2 classi ● □



Aggiungo nuovi
dati ● ■



Capacità di generalizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico
- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete
- Architetture
- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

La capacità di **generalizzazione** di una rete addestrata è la capacità di fornire una risposta corretta a nuovi ingressi (non presentati nella fase di addestramento)

Capacità di generalizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico
- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete
- Architetture
- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

La capacità di **generalizzazione** di una rete addestrata è la capacità di fornire una risposta corretta a nuovi ingressi (non presentati nella fase di addestramento)

Lo scopo ultimo dell'addestramento è quello di costruire un modello del processo che genera i dati e non di interpolare i dati di training

Capacità di generalizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

- Rete neurale da un punto di vista fisico
- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete
- Architetture
- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

La capacità di **generalizzazione** di una rete addestrata è la capacità di fornire una risposta corretta a nuovi ingressi (non presentati nella fase di addestramento)

Lo scopo ultimo dell'addestramento è quello di costruire un modello del processo che genera i dati e non di interpolare i dati di training

La teoria dell'apprendimento studia l'apprendimento e la caratterizzazione delle capacità di generalizzazione delle reti neurali. Il problema è definire in modo ottimo la complessità della rete e il procedimento di identificazione dei parametri w a partire dai dati in modo da ottenere la miglior capacità di generalizzazione possibile.

Neurone formale

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

● Neurone formale

● Perceptron

● Limiti del perceptron

● La crisi delle reti neurali

● Backpropagation

● RN e ottimizzazione

● ma soprattutto...

● e ancora..

Perceptron

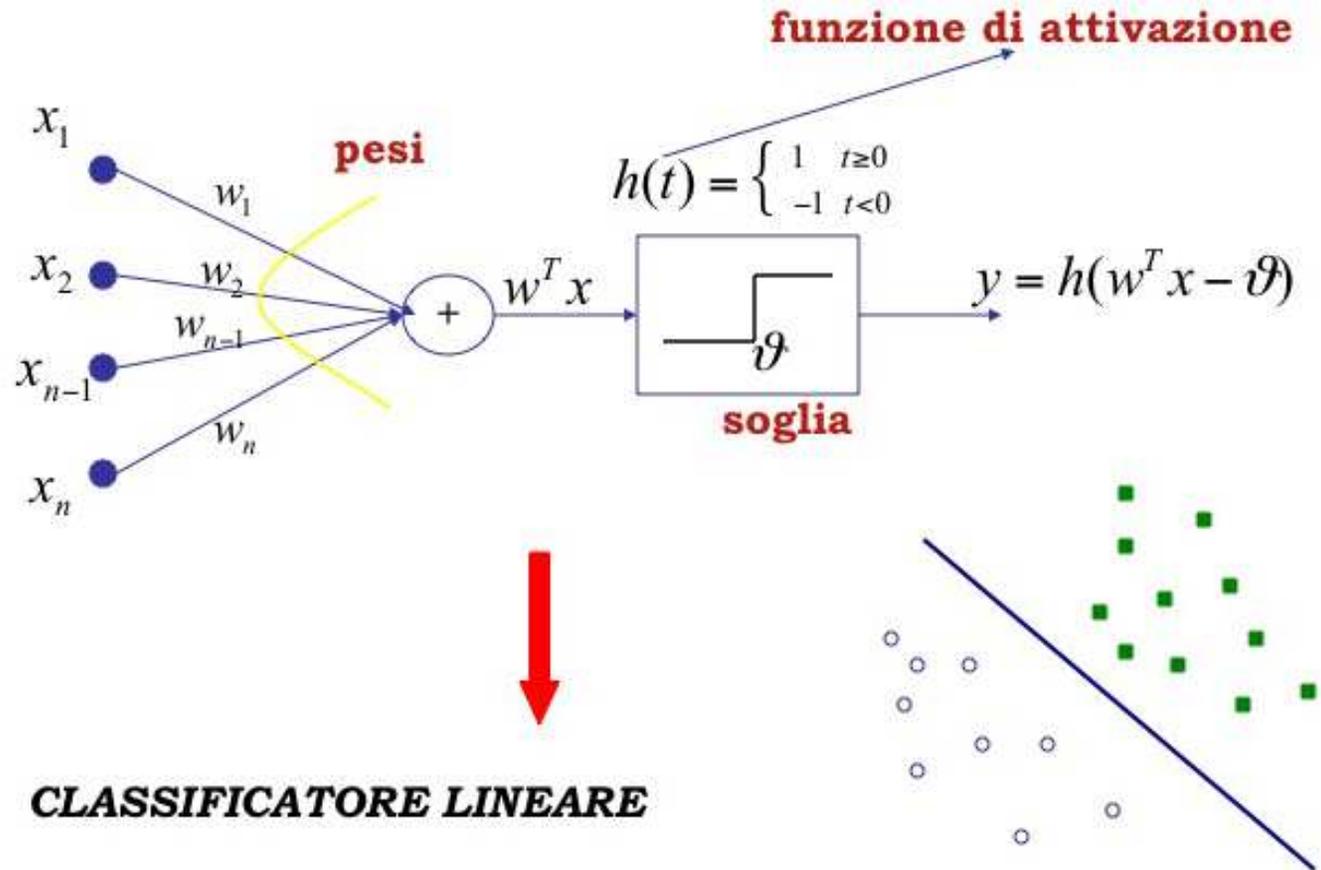
Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF



Perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- **Perceptron**
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Una rete costituita da un singolo strato di neuroni formali è stata denominata **PERCEPTRON** (Rosenblatt, 1962)

Perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- **Perceptron**
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Una rete costituita da un singolo strato di neuroni formali è stata denominata **PERCEPTRON** (Rosenblatt, 1962)

Il **Perceptron** esegue una trasformazione non lineare degli ingressi x e fornisce un'uscita scalare (nella versione di Mc Cullock e Pitts 43)

$$y(x) = \text{sgn}(w^T x - \theta)$$

Perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- **Perceptron**
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Una rete costituita da un singolo strato di neuroni formali è stata denominata **PERCEPTRON** (Rosenblatt, 1962)

Il **Perceptron** esegue una trasformazione non lineare degli ingressi x e fornisce un'uscita scalare (nella versione di Mc Cullock e Pitts 43)

$$y(x) = \text{sgn}(w^T x - \theta)$$

Scegliendo i pesi in modo appropriato è in grado di realizzare le funzioni and, or, not.

Perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- **Perceptron**
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Una rete costituita da un singolo strato di neuroni formali è stata denominata **PERCEPTRON** (Rosenblatt, 1962)

Il **Perceptron** esegue una trasformazione non lineare degli ingressi x e fornisce un'uscita scalare (nella versione di Mc Cullock e Pitts 43)

$$y(x) = \text{sgn}(w^T x - \theta)$$

Scegliendo i pesi in modo appropriato è in grado di realizzare le funzioni and, or, not.

Nel '62 Rosenblatt lo reinterpreta come classificatore lineare: la funzione discriminante è la funzione lineare $w^T x - \theta$

Perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- **Perceptron**
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Una rete costituita da un singolo strato di neuroni formali è stata denominata **PERCEPTRON** (Rosenblatt, 1962)

Il **Perceptron** esegue una trasformazione non lineare degli ingressi x e fornisce un'uscita scalare (nella versione di Mc Cullock e Pitts 43)

$$y(x) = \text{sgn}(w^T x - \theta)$$

Scegliendo i pesi in modo appropriato è in grado di realizzare le funzioni and, or, not.

Nel '62 Rosenblatt lo reinterpreta come classificatore lineare: la funzione discriminante è la funzione lineare $w^T x - \theta$

è stato proposto un algoritmo per il calcolo dei parametri (pesi e soglie) che fornisce una soluzione in un numero finito di iterazioni nell'ipotesi in cui i pattern di ingresso siano **linearmente separabili**

Limiti del perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- **Limiti del perceptron**
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

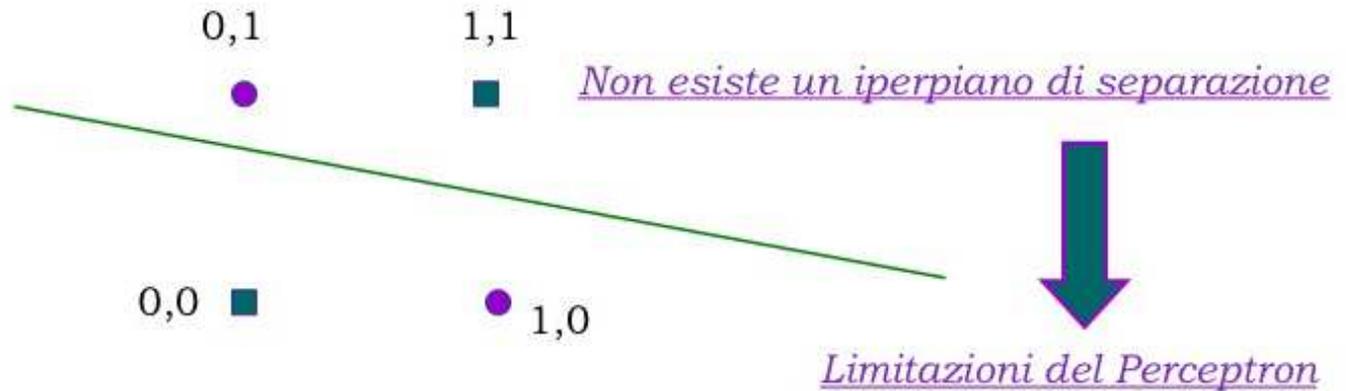
Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Problema XOR



La crisi delle reti neurali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Le limitazioni del Perceptron sono state messe in luce da Minsky e Papert (1969)

La crisi delle reti neurali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Le limitazioni del Perceptron sono state messe in luce da Minsky e Papert (1969)

L'effetto del libro di Minsky e Papert è stato quello di far decadere l'interesse iniziale verso le reti neurali

La crisi delle reti neurali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Le limitazioni del Perceptron sono state messe in luce da Minsky e Papert (1969)

L'effetto del libro di Minsky e Papert è stato quello di far decadere l'interesse iniziale verso le reti neurali

Era noto che le limitazioni del Perceptron potevano essere superate, in linea di principio, collegando fra loro in modo opportuno dei neuroni formali o effettuando delle trasformazioni non lineari degli ingressi

La crisi delle reti neurali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Le limitazioni del Perceptron sono state messe in luce da Minsky e Papert (1969)

L'effetto del libro di Minsky e Papert è stato quello di far decadere l'interesse iniziale verso le reti neurali

Era noto che le limitazioni del Perceptron potevano essere superate, in linea di principio, collegando fra loro in modo opportuno dei neuroni formali o effettuando delle trasformazioni non lineari degli ingressi

Non erano tuttavia disponibili algoritmi di addestramento per il calcolo dei parametri

Backpropagation

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- **Backpropagation**
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Una rinascita dell'interesse verso le reti neurali è stata in gran parte determinata dal lavoro di Rumelhart, Hinton e Williams (1986), che hanno proposto un algoritmo di addestramento per reti di neuroni formali, noto come metodo della **backpropagation**, essenzialmente basato sul metodo di ottimizzazione del gradiente

Backpropagation

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- **Backpropagation**
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Una rinascita dell'interesse verso le reti neurali è stata in gran parte determinata dal lavoro di Rumelhart, Hinton e Williams (1986), che hanno proposto un algoritmo di addestramento per reti di neuroni formali, noto come metodo della **backpropagation**, essenzialmente basato sul metodo di ottimizzazione del gradiente

Gli sviluppi futuri hanno portato allo sviluppo di un'area di ricerca interdisciplinare, in cui sono stati integrati contributi di vari settori

RN e ottimizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- **RN e ottimizzazione**
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

OTTIMIZZAZIONE



RETI NEURALI



“**riscoperta**” di vari metodi di Ottimizzazione nell’ ambito delle reti neurali



motivazioni per la definizione di nuovi metodi di Ottimizzazione (**algoritmi incrementali**, **algoritmi di decomposizione**)

ma soprattutto...

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- **ma soprattutto...**
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Carriere e persone

Ruoli Il suo compito? Elaborare la gran mole di dati per migliorare le prestazioni dell'impresa

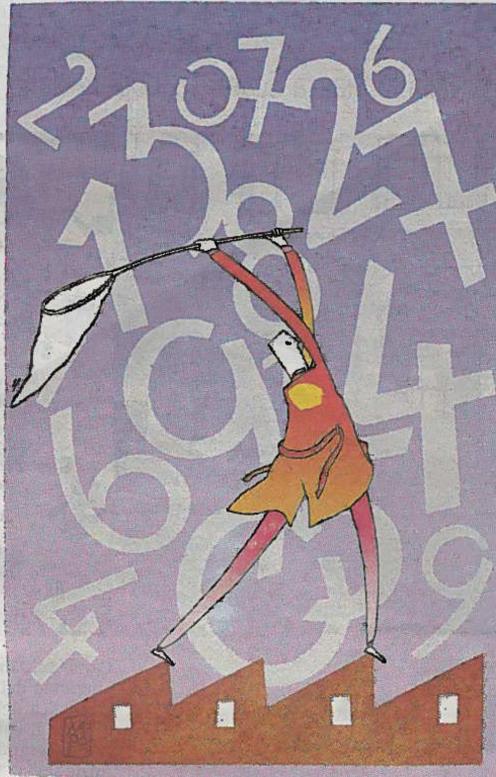
Mr Numeri entra in azienda

Nasce il «data scientist», tra matematica e gestione

Overdose o flusso vitale di informazioni? È il dilemma delle più grandi imprese inondate quotidianamente da una mole spropositata di dati provenienti dai più svariati canali digitali: i server aziendali, i siti web, i social network, gli smartphone, le transazioni bancarie, lo shopping virtuale e quant'altro. Oggi non è più possibile lasciarsi travolgere e chi non è in grado di gestire questi «Big Data» diventerà a breve non più concorrenziale.

Il problema contingente, però, è la carenza di professionisti capaci di organizzare, governare, e volgere a proprio favore i dati, per dare suggerimenti vincenti di business. Secondo un'indagine del McKinsey global institute, infatti, solamente negli Stati Uniti nei prossimi anni occorreranno 190 mila esperti della gestione dei Big Data e più di un milione e mezzo di tecnici e manager a loro agio con la lettura e interpretazione dei dati digitali.

Fatte le dovute proporzioni anche in Italia ci si aspetta una domanda molto vivace. Anche se spesso i capi azienda non sono ancora sensibili al problema. Secondo un'indagine di



Vanson Bourne condotta su un campione di 685 aziende per conto della multinazionale del software Ca Technologies, il 59% dei Cio (i responsabili delle It aziendali) lamenta che il top management considera la gestione dei Big Data come un costo di esercizio e non come un indispensabile strumento per prendere decisioni di business. Eppure, già molto prima dell'era Internet Peter Drucker, il massimo guru del management, aveva detto: «Non si può gestire ciò che non si può misurare». La questione è così calda che l'ultimo numero in versione italiana della Harvard business review dedica un lungo servizio di competenza proprio ai Big Data. «Sono una rivoluzione — sostengono i docenti del MIt Andrew McAfee ed Erik Brynjolfsson — sono molto più efficaci rispetto agli strumenti analitici usati finora. Permettono di programmare interventi efficaci in aree fino ad oggi dominate dall'istinto». Una vera sentenza di morte, cioè, per uno stile di gestione manageriale che si basa sull'intuizione e sull'esperienza, perché, sostengono gli autori, «le decisioni supportate dai dati sono più ef-

ficaci per il business». E chi gestisce basandosi sui Big Data ha risultati migliori: una ricerca del MIt basata su 330 aziende quotate in Borsa conclude che le imprese che decidono sulla base dei dati sono state del 6% più produttive e del 5% più profittevoli delle concorrenti.

«Insomma — spiega Vincenzo Aloisio, responsabile analytics della multinazionale della consulenza Accenture — sta nascendo un nuovo professionista, il «data scientist», capace di elaborare la gran mole di dati per migliorare le prestazioni aziendali. Deve saper usare i modelli matematico-statistici e, contemporaneamente, avere conoscenze di business per mettere a frutto l'interpretazione dei dati». Consigli per un giovane attratto dai Big Data? «Prima una laurea in matematica statistica o ingegneria — conclude Aloisio — poi, prima di entrare in un'azienda che stia sviluppando queste tematiche, seguire corsi post universitari specializzati». Una direzione in cui già si muovono, per esempio, i politecnici di Milano e Torino.

Enzo Riboni

© RIPRODUZIONE RISERVATA

e ancora..

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation
- RN e ottimizzazione
- ma soprattutto...
- e ancora..

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

<http://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the>
"Articolo su Harvard Business Review"

Perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

● Perceptron

● Addestramento I

● Addestramento II

● Limiti del perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Il **Perceptron** esegue una trasformazione non lineare degli ingressi x e fornisce un'uscita scalare (nella versione di Mc Cullock e Pitts 43)

$$y(x) = \text{sgn}(w^T x - \theta)$$

Perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

● Perceptron

● Addestramento I

● Addestramento II

● Limiti del perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Il **Perceptron** esegue una trasformazione non lineare degli ingressi x e fornisce un'uscita scalare (nella versione di Mc Cullock e Pitts 43)

$$y(x) = \text{sgn}(w^T x - \theta)$$

Scegliendo i pesi in modo appropriato è in grado di realizzare le funzioni and, or, not.

Perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

● Perceptron

● Addestramento I

● Addestramento II

● Limiti del perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Il **Perceptron** esegue una trasformazione non lineare degli ingressi x e fornisce un'uscita scalare (nella versione di Mc Cullock e Pitts 43)

$$y(x) = \text{sgn}(w^T x - \theta)$$

Scegliendo i pesi in modo appropriato è in grado di realizzare le funzioni and, or, not.

Nel '62 Rosenblatt lo reinterpreta come classificatore lineare: la funzione discriminante è la funzione lineare $w^T x - \theta$

Addestramento I

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

- Perceptron
- **Addestramento I**
- Addestramento II
- Limiti del perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Sia dato un **training set**

$$\mathcal{T} = \{(x^p, y^p) : x \in \mathbb{R}^n, y^p \in \{-1, 1\}, p = 1, \dots, P\},$$

diviso in due insiemi

$$\mathcal{A} = \{x^p : y^p = 1\} \quad \mathcal{B} = \{x^p : y^p = -1\}$$

Addestramento I

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

● Perceptron

● **Addestramento I**

● Addestramento II

● Limiti del perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Sia dato un **training set**

$$\mathcal{T} = \{(x^p, y^p) : x \in \mathbb{R}^n, y^p \in \{-1, 1\}, p = 1, \dots, P\},$$

diviso in due insiemi

$$\mathcal{A} = \{x^p : y^p = 1\} \quad \mathcal{B} = \{x^p : y^p = -1\}$$

Addestrare il perceptron vuol dire individuare i pesi w e la soglia θ tali che:

$$w^T x^p - \theta > 0 \quad \forall x^p \in \mathcal{A}$$

$$w^T x^p - \theta < 0 \quad \forall x^p \in \mathcal{B}$$

Addestramento I

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

● Perceptron

● Addestramento I

● Addestramento II

● Limiti del perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Sia dato un **training set**

$$\mathcal{T} = \{(x^p, y^p) : x \in \mathbb{R}^n, y^p \in \{-1, 1\}, p = 1, \dots, P\},$$

diviso in due insiemi

$$\mathcal{A} = \{x^p : y^p = 1\} \quad \mathcal{B} = \{x^p : y^p = -1\}$$

Addestrare il perceptron vuol dire individuare i pesi w e la soglia θ tali che:

$$\begin{aligned} w^T x^p - \theta &> 0 \quad \forall x^p \in \mathcal{A} \\ w^T x^p - \theta &< 0 \quad \forall x^p \in \mathcal{B} \end{aligned}$$

Si introduce un ingresso fittizio $x_0^p = -1, p = 1, \dots, P$ e si ridefinisce $w = (\theta \ w)$, il sistema diventa

$$\begin{aligned} w^T x^p &> 0 \quad \forall x^p \in \mathcal{A} \\ w^T x^p &< 0 \quad \forall x^p \in \mathcal{B} \end{aligned}$$

Addestramento II

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

- Perceptron
- Addestramento I
- Addestramento II
- Limiti del perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Dati $(x^p, y^p) \in \mathcal{T}, p = 1, \dots, P$

Iniz $w(0) = 0, k = 0, \text{nclass} = 0$

While $\text{nclass} < P$

For $p = 1, \dots, P,$

If $\text{sgn}(w^T x^p) = y^p$ **then**
 $\text{nclass} = \text{nclass} + 1$

else

$w(k + 1) = w(k) + y^p x^p$
 $k = k + 1$

End if

End for

If $\text{nclass} < P, \text{nclass} = 0$

End do

Limiti del perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

- Perceptron
- Addestramento I
- Addestramento II
- Limiti del perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Se i due insiemi \mathcal{A} e \mathcal{B} sono **linearmente separabili**, l'algoritmo converge.

Limiti del perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

- Perceptron
- Addestramento I
- Addestramento II
- Limiti del perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Se i due insiemi \mathcal{A} e \mathcal{B} sono **linearmente separabili**, l'algoritmo converge.

Si potrebbe addestrare risolvendo il sistema di disequazioni con metodi ai punti interni.

Oppure sostituendo la funzione sgn con una qualunque funzione ρ monotona crescente tale che:

$$\lim_{t \rightarrow -\infty} \rho(t) = -1, \quad \lim_{t \rightarrow \infty} \rho(t) = 1$$

e minimizzare poi la funzione di errore quadratico sul training:

$$E(w) = \frac{1}{2} \sum_{p=1}^P (y^p - \rho(w^T x^p))^2$$

Limiti del perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

- Perceptron
- Addestramento I
- Addestramento II
- Limiti del perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Se i due insiemi \mathcal{A} e \mathcal{B} sono **linearmente separabili**, l'algoritmo converge.

Si potrebbe addestrare risolvendo il sistema di disequazioni con metodi ai punti interni.

Oppure sostituendo la funzione sgn con una qualunque funzione ρ monotona crescente tale che:

$$\lim_{t \rightarrow -\infty} \rho(t) = -1, \quad \lim_{t \rightarrow \infty} \rho(t) = 1$$

e minimizzare poi la funzione di errore quadratico sul training:

$$E(w) = \frac{1}{2} \sum_{p=1}^P (y^p - \rho(w^T x^p))^2$$

Non è in grado di rappresentare funzioni anche semplici come lo XOR (OR esclusivo o somma modulo 2, restituisce 1 se e solo se uno solo dei due operandi 1, mentre restituisce 0 in tutti gli altri casi).

Limiti del perceptron

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

- Perceptron
- Addestramento I
- Addestramento II
- Limiti del perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Se i due insiemi \mathcal{A} e \mathcal{B} sono **linearmente separabili**, l'algoritmo converge.

Si potrebbe addestrare risolvendo il sistema di disequazioni con metodi ai punti interni.

Oppure sostituendo la funzione sgn con una qualunque funzione ρ monotona crescente tale che:

$$\lim_{t \rightarrow -\infty} \rho(t) = -1, \quad \lim_{t \rightarrow \infty} \rho(t) = 1$$

e minimizzare poi la funzione di errore quadratico sul training:

$$E(w) = \frac{1}{2} \sum_{p=1}^P (y^p - \rho(w^T x^p))^2$$

Non è in grado di rappresentare funzioni anche semplici come lo XOR (OR esclusivo o somma modulo 2, restituisce 1 se e solo se uno solo dei due operandi 1, mentre restituisce 0 in tutti gli altri casi).

Si può solo garantire errore nullo sul training set, non ci sono garanzie sulla capacità di generalizzazione.

Definizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

● Definizione

- Struttura
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Sono architetture costituite da più strati di neuroni connessi in cascata:

Definizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

● Definizione

● Struttura

● Esempio

● Funzioni di attivazione

● Rete a due strati

● Proprietà

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Sono architetture costituite da più strati di neuroni connessi in cascata:

- consentono di approssimare qualunque funzione continua su un insieme compatto (sotto opportune ipotesi sulle funzioni di attivazione);

Definizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

● Definizione

● Struttura

● Esempio

● Funzioni di attivazione

● Rete a due strati

● Proprietà

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Sono architetture costituite da più strati di neuroni connessi in cascata:

- consentono di approssimare qualunque funzione continua su un insieme compatto (sotto opportune ipotesi sulle funzioni di attivazione);
- risolvono problemi di classificazione anche di insieme non linearmente separabili;

Definizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

● Definizione

● Struttura

● Esempio

● Funzioni di attivazione

● Rete a due strati

● Proprietà

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Sono architetture costituite da più strati di neuroni connessi in cascata:

- consentono di approssimare qualunque funzione continua su un insieme compatto (sotto opportune ipotesi sulle funzioni di attivazione);
- risolvono problemi di classificazione anche di insieme non linearmente separabili;
- l'addestramento è più complesso.

Struttura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- **Struttura**
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'architettura si definisce tramite:

Struttura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- **Struttura**
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'architettura si definisce tramite:

◁ insieme di n nodi in ingresso associati agli n ingressi della rete

Struttura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- **Struttura**
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'architettura si definisce tramite:

- ◁ insieme di n nodi in ingresso associati agli n ingressi della rete
- ◁ un insieme di neuroni formali organizzati in $L \geq 2$ strati di cui:

Struttura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

● Definizione

● **Struttura**

● Esempio

● Funzioni di attivazione

● Rete a due strati

● Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'architettura si definisce tramite:

- ◁ insieme di n nodi in ingresso associati agli n ingressi della rete
- ◁ un insieme di neuroni formali organizzati in $L \geq 2$ strati di cui:
 - $L - 1$ strati nascosti: le uscite fanno da ingresso ad altri neuroni della rete

Struttura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

● Definizione

● **Struttura**

● Esempio

● Funzioni di attivazione

● Rete a due strati

● Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'architettura si definisce tramite:

- ◁ insieme di n nodi in ingresso associati agli n ingressi della rete
- ◁ un insieme di neuroni formali organizzati in $L \geq 2$ strati di cui:
 - $L - 1$ strati nascosti: le uscite fanno da ingresso ad altri neuroni della rete
 - uno strato di uscita costituito da $K \geq 1$ neuroni le cui uscite sono le uscite della rete

Struttura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

● Definizione

● **Struttura**

● Esempio

● Funzioni di attivazione

● Rete a due strati

● Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'architettura si definisce tramite:

- ◁ insieme di n nodi in ingresso associati agli n ingressi della rete
- ◁ un insieme di neuroni formali organizzati in $L \geq 2$ strati di cui:
 - $L - 1$ strati nascosti: le uscite fanno da ingresso ad altri neuroni della rete
 - uno strato di uscita costituito da $K \geq 1$ neuroni le cui uscite sono le uscite della rete
- ◁ Un insieme di archi orientati e pesati che costituiscono le connessioni interneuroni e quelle con i nodi di ingresso.

Struttura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

● Definizione

● **Struttura**

● Esempio

● Funzioni di attivazione

● Rete a due strati

● Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

L'architettura si definisce tramite:

- ◁ insieme di n nodi in ingresso associati agli n ingressi della rete
- ◁ un insieme di neuroni formali organizzati in $L \geq 2$ strati di cui:
 - $L - 1$ strati nascosti: le uscite fanno da ingresso ad altri neuroni della rete
 - uno strato di uscita costituito da $K \geq 1$ neuroni le cui uscite sono le uscite della rete
- ◁ Un insieme di archi orientati e pesati che costituiscono le connessioni interneuroni e quelle con i nodi di ingresso.
- ◁ NON esistono connessioni tra neuroni dello stesso strato nè archi all'indietro

Esempio

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

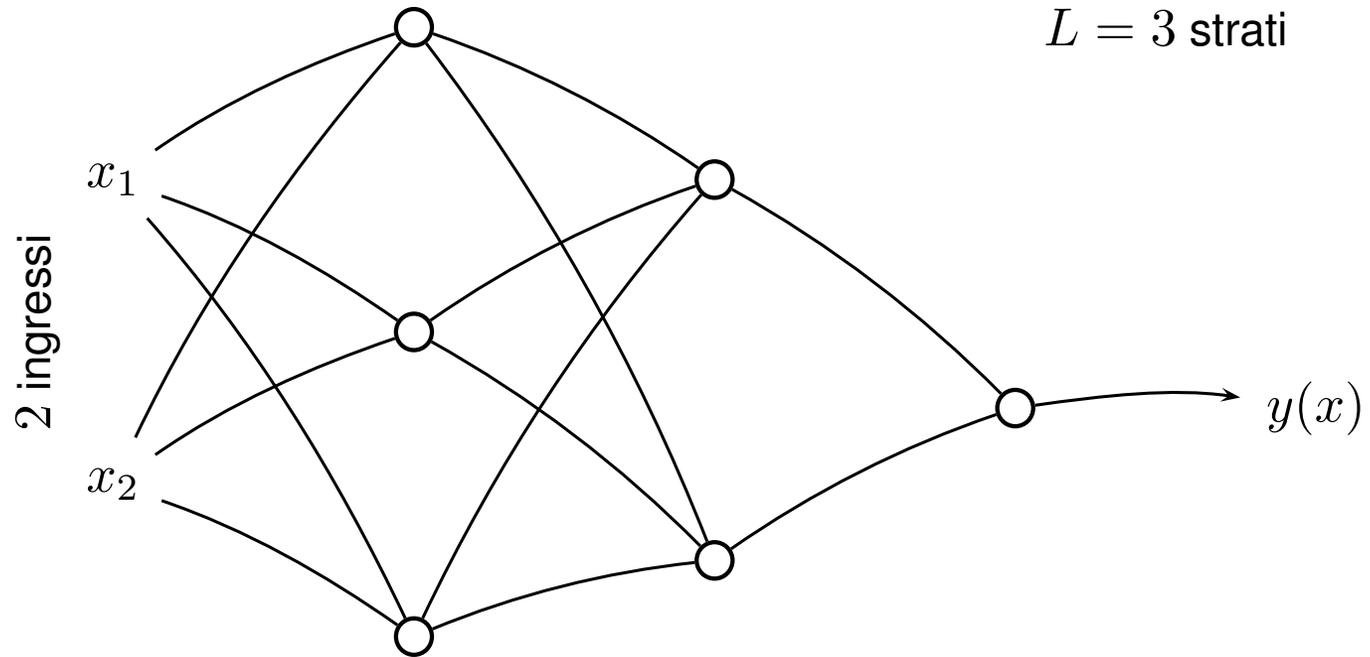
- Definizione
- Struttura
- **Esempio**
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

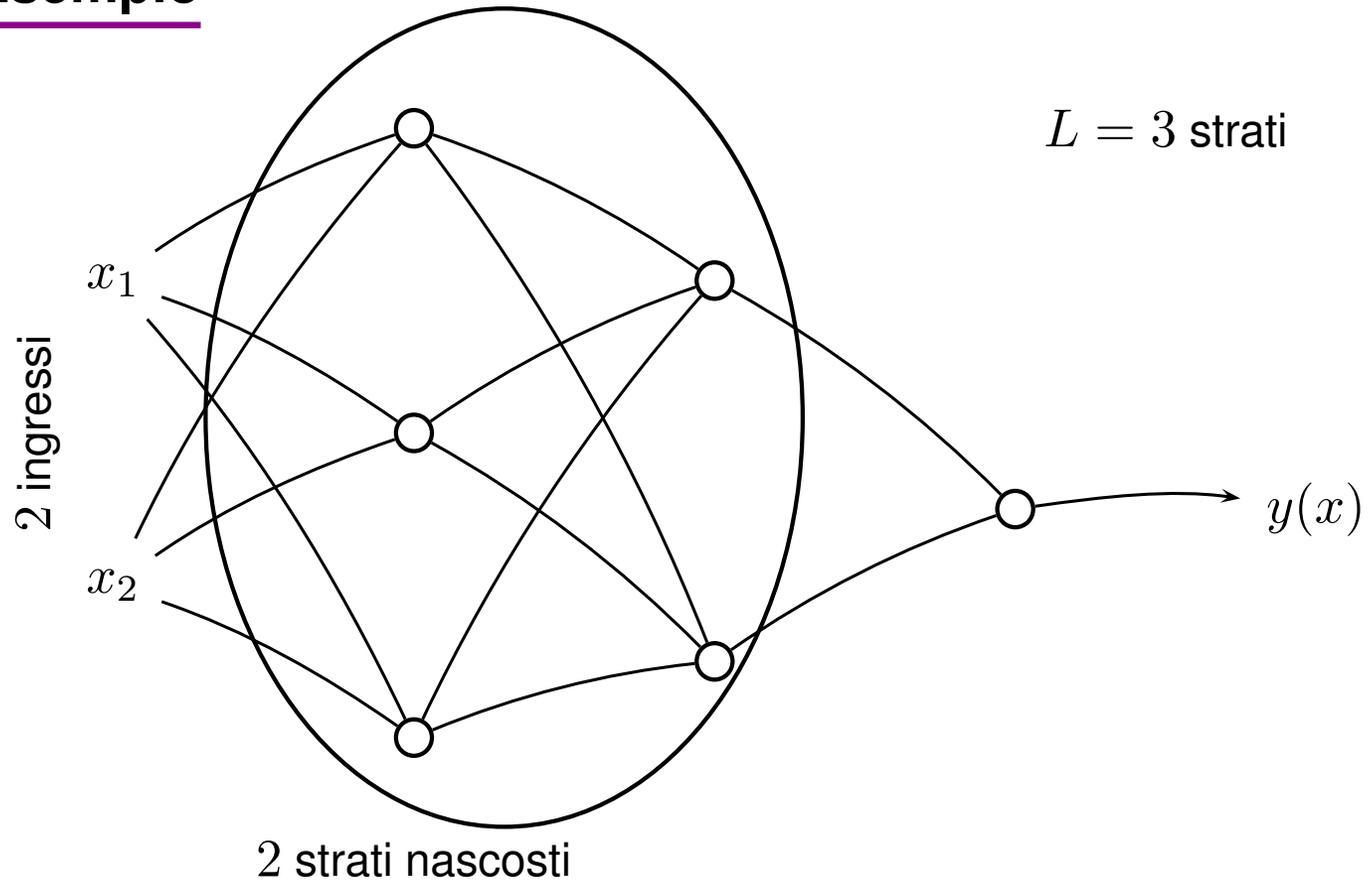
Metodi on-line

Reti RBF

Addestramento RBF



Esempio



Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- **Esempio**
- Funzioni di attivazione
- Rete a due strati
- Proprietà

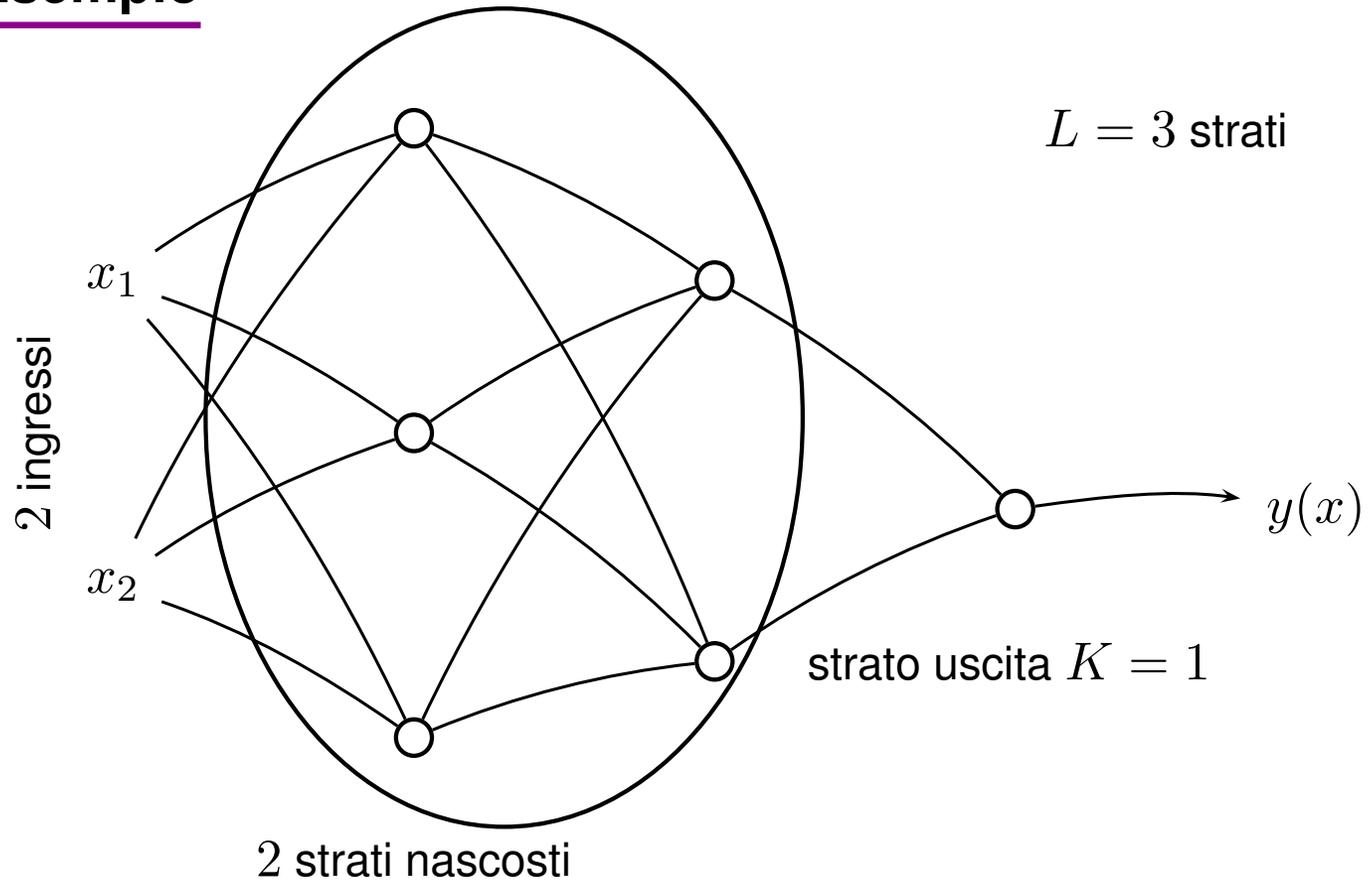
Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Esempio



Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- **Esempio**
- Funzioni di attivazione
- Rete a due strati
- Proprietà

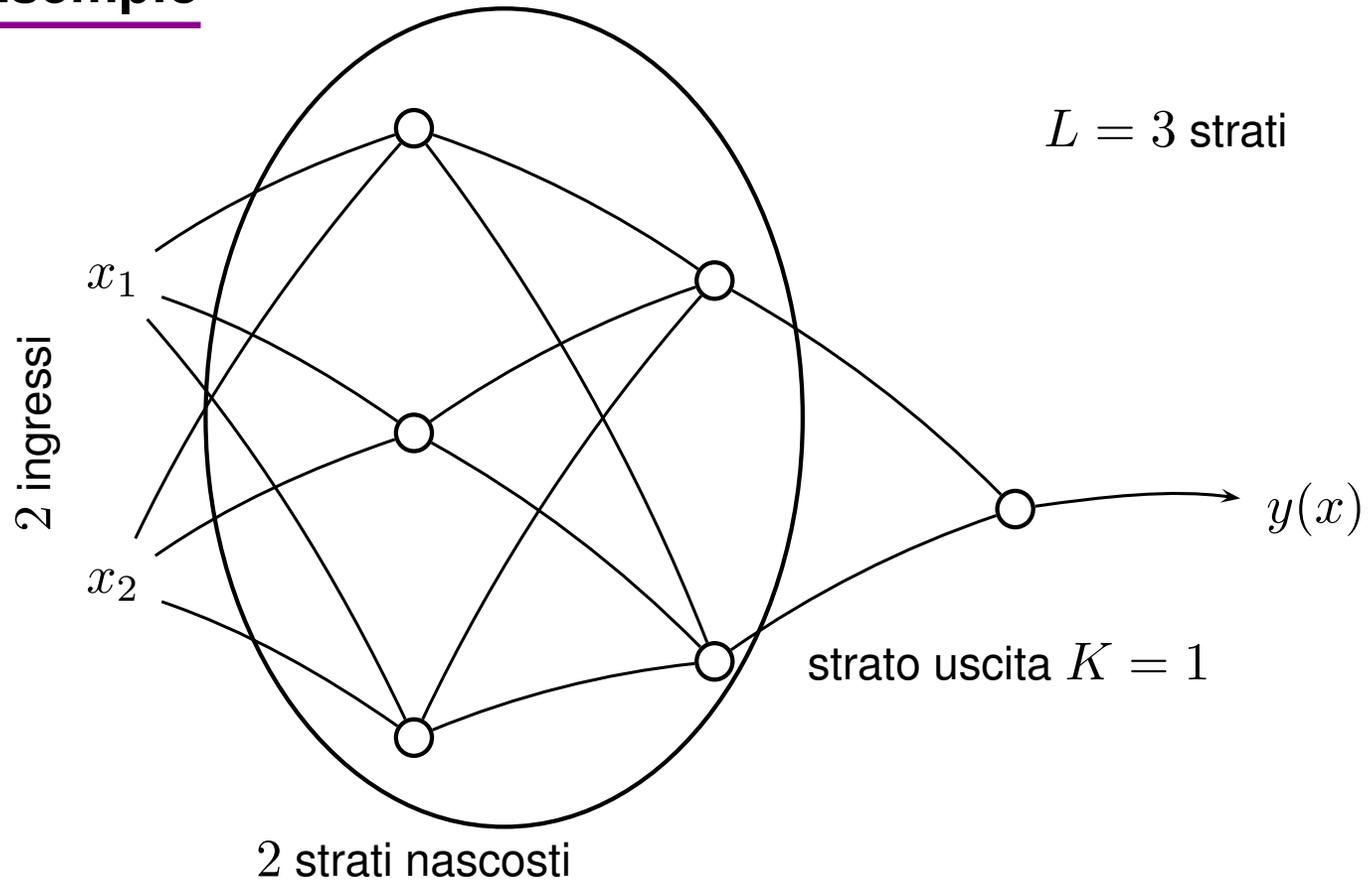
Creazione di una rete multistrato

Metodi on-line

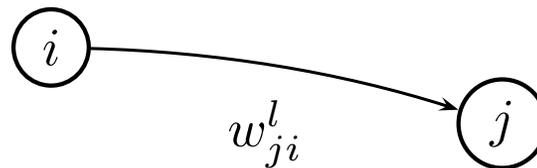
Reti RBF

Addestramento RBF

Esempio



Dati due neuroni i dello strato $l - 1$ e j dello strato l si ha



Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- **Esempio**
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Funzioni di attivazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Ogni neurone j ha una funzione di attivazione $g_j^{(l)} : \mathbb{R} \Rightarrow \mathbb{R}$ che prende in ingresso la combinazione pesata degli ingressi a_j , mentre w_{j0} è la soglia.

Funzioni di attivazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Ogni neurone j ha una funzione di attivazione $g_j^{(l)} : \mathbb{R} \Rightarrow \mathbb{R}$ che prende in ingresso la combinazione pesata degli ingressi a_j , mentre w_{j0} è la soglia.

Per il primo strato si ha:

$$a_j^{(1)} = \sum_{k=1}^n w_{jk} x_k - w_{j0} \quad z_j^1 = g_j(a_j^{(1)}),$$

Funzioni di attivazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

● Definizione

● Struttura

● Esempio

● Funzioni di attivazione

● Rete a due strati

● Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Ogni neurone j ha una funzione di attivazione $g_j^{(l)} : \mathbb{R} \Rightarrow \mathbb{R}$ che prende in ingresso la combinazione pesata degli ingressi a_j , mentre w_{j0} è la soglia.

Per il primo strato si ha:

$$a_j^{(1)} = \sum_{k=1}^n w_{jk} x_k - w_{j0} \quad z_j^1 = g_j(a_j^{(1)}),$$

Per lo strato l generico si ha:

$$a_j^{(l)} = \sum_{k=1}^{N^{l-1}} w_{jk}^{l-1} z_k^{l-1} - w_{j0}^{l-1} \quad z_j^l = g_j^{(l)}(a_j^{(l)}),$$

dove N^l è il numero di neuroni dello strato nascosto l .

Funzioni di attivazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Ogni neurone j ha una funzione di attivazione $g_j^{(l)} : \mathbb{R} \Rightarrow \mathbb{R}$ che prende in ingresso la combinazione pesata degli ingressi a_j , mentre w_{j0} è la soglia.

Per il primo strato si ha:

$$a_j^{(1)} = \sum_{k=1}^n w_{jk} x_k - w_{j0} \quad z_j^1 = g_j(a_j^{(1)}),$$

Per lo strato l generico si ha:

$$a_j^{(l)} = \sum_{k=1}^{N^{l-1}} w_{jk}^{l-1} z_k^{l-1} - w_{j0}^{l-1} \quad z_j^l = g_j^{(l)}(a_j^{(l)}),$$

dove N^l è il numero di neuroni dello strato nascosto l .

Tipici esempi di funzioni di attivazione (che sono di solito differenziabili e sigmoidali) sono:

$$g(t) = \frac{1}{1 + e^{-ct}}, \quad c > 0 \quad g(t) = \frac{1 - e^{-t}}{1 + e^{-t}}$$

Rete a due strati

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Consideriamo la rete a 2 strati con uno strato nascosto e uscita lineare:

Rete a due strati

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

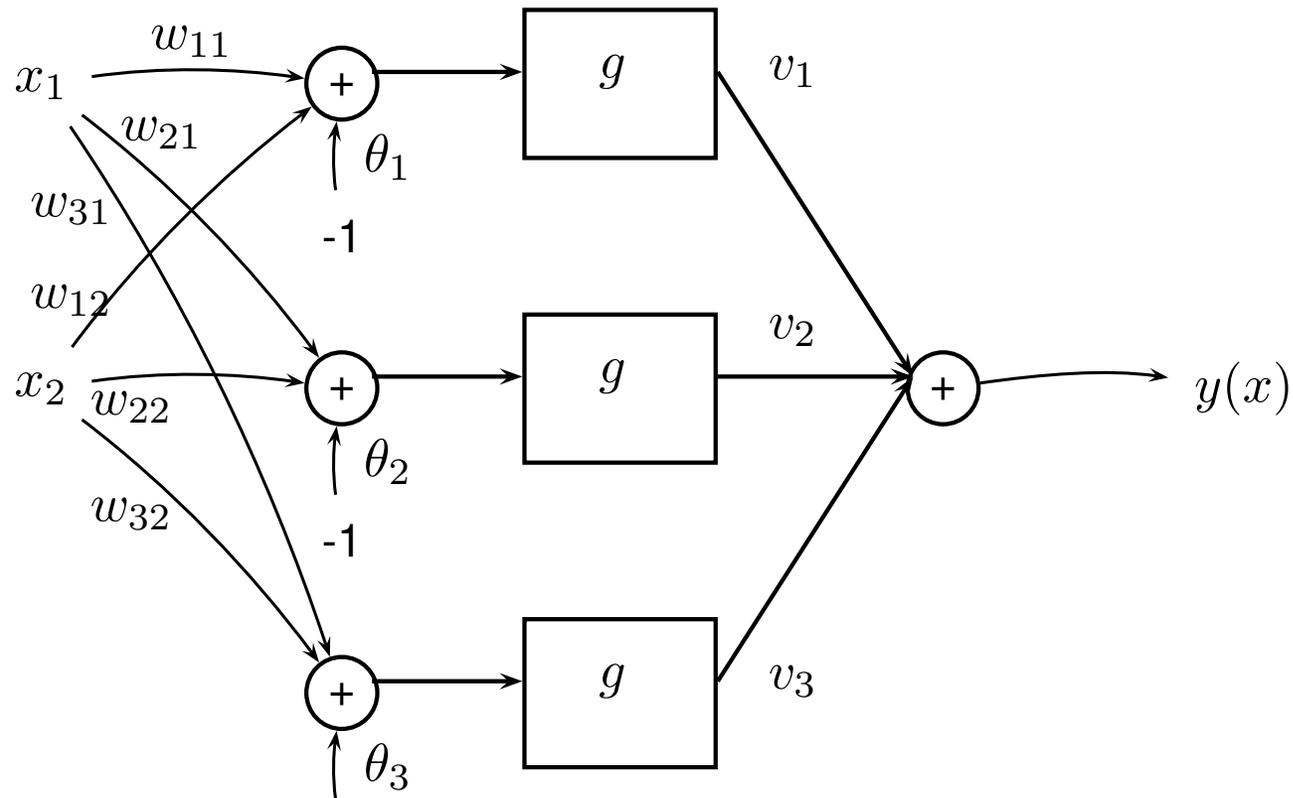
Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Consideriamo la rete a 2 strati con uno strato nascosto e uscita lineare:



$$y(x) = \sum_{k=1}^3 v_k \sum_{j=1}^2 g(w_{kj}x_j - \theta_j)$$

Proprietà

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Una rete multistrato a 2 strati con uno strato nascosto e uscita lineare è un approssimatore universale: approssima qualunque funzione continua su un insieme compatto (sotto l'ipotesi che g sia continua e non sia un polinomio)

Proprietà

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- Esempio
- Funzioni di attivazione
- Rete a due strati
- **Proprietà**

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Una rete multistrato a 2 strati con uno strato nascosto e uscita lineare è un approssimatore universale: approssima qualunque funzione continua su un insieme compatto (sotto l'ipotesi che g sia continua e non sia un polinomio) Approssimare una funzione $f(x)$ su un insieme compatto Ω vuol dire che, per un qualunque $\epsilon > 0$,

$$\max_{x \in \Omega} |f(x) - y(x)| < \epsilon$$

Proprietà

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

- Definizione
- Struttura
- Esempio
- Funzioni di attivazione
- Rete a due strati
- Proprietà

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

Una rete multistrato a 2 strati con uno strato nascosto e uscita lineare è un approssimatore universale: approssima qualunque funzione continua su un insieme compatto (sotto l'ipotesi che g sia continua e non sia un polinomio) Approssimare una funzione $f(x)$ su un insieme compatto Ω vuol dire che, per un qualunque $\epsilon > 0$,

$$\max_{x \in \Omega} |f(x) - y(x)| < \epsilon$$

Le reti multistrato possono essere usate sia per problemi di regressione che per problemi di classificazione.

Struttura e addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

● **Struttura e
addestramento**

● Scelta dell'architettura
● Stabilizzazione
strutturale

● Tecnica di
Regolarizzazione

● Early stopping

● Addestramento

● Classi di algoritmi

● Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Determinare una rete multistrato con n ingressi e K uscite implica due passi:

Struttura e addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

● **Struttura e addestramento**

● Scelta dell'architettura
● Stabilizzazione strutturale

● Tecnica di Regolarizzazione

● Early stopping

● Addestramento

● Classi di algoritmi

● Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Determinare una rete multistrato con n ingressi e K uscite implica due passi:

1. Scegliere l'architettura della rete: numero di strati, numero di neuroni per ogni strato

Struttura e addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

● **Struttura e
addestramento**

● Scelta dell'architettura
● Stabilizzazione
strutturale

● Tecnica di
Regolarizzazione

● Early stopping

● Addestramento

● Classi di algoritmi

● Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Determinare una rete multistrato con n ingressi e K uscite implica due passi:

1. Scegliere l'architettura della rete: numero di strati, numero di neuroni per ogni strato
2. addestrare la rete: scegliere w e θ per ogni strato.

Struttura e addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

● **Struttura e addestramento**

● Scelta dell'architettura
● Stabilizzazione strutturale

● Tecnica di

Regularizzazione

● Early stopping

● Addestramento

● Classi di algoritmi

● Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Determinare una rete multistrato con n ingressi e K uscite implica due passi:

1. Scegliere l'architettura della rete: numero di strati, numero di neuroni per ogni strato
2. addestrare la rete: scegliere w e θ per ogni strato.

Addestrare la rete vuol dire risolvere il problema:

$$\min_{w \in \mathfrak{R}^m} E(w) = \sum_{p=1}^P E_p(w)$$

dove $E_p(w)$ è il termine di errore relativo al campione p del training set, che misura la distanza tra l'uscita desiderata y^p e l'uscita prodotta dalla rete, ad es.:

$$E_p = \frac{1}{2} \|y(x^p, w) - y^p\|^2$$

Struttura e addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

● **Struttura e addestramento**

● Scelta dell'architettura
● Stabilizzazione strutturale

● Tecnica di

Regularizzazione

● Early stopping

● Addestramento

● Classi di algoritmi

● Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Determinare una rete multistrato con n ingressi e K uscite implica due passi:

1. Scegliere l'architettura della rete: numero di strati, numero di neuroni per ogni strato
2. addestrare la rete: scegliere w e θ per ogni strato.

Addestrare la rete vuol dire risolvere il problema:

$$\min_{w \in \mathbb{R}^m} E(w) = \sum_{p=1}^P E_p(w)$$

dove $E_p(w)$ è il termine di errore relativo al campione p del training set, che misura la distanza tra l'uscita desiderata y^p e l'uscita prodotta dalla rete, ad es.:

$$E_p = \frac{1}{2} \|y(x^p, w) - y^p\|^2$$

Supponiamo $E_p \in \mathcal{C}^1$

Scelta dell'architettura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e
addestramento

- Scelta dell'architettura

- Stabilizzazione
strutturale

- Tecnica di
Regolarizzazione

- Early stopping

- Addestramento

- Classi di algoritmi

- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Scopo dell'addestramento non è interpolare i dati del training set, ma ottenere una buona capacità di generalizzazione.

Scelta dell'architettura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento

- Scelta dell'architettura

- Stabilizzazione strutturale

- Tecnica di Regolarizzazione

- Early stopping

- Addestramento

- Classi di algoritmi

- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Scopo dell'addestramento non è interpolare i dati del training set, ma ottenere una buona capacità di generalizzazione.

Uno dei problemi più difficili è quello di stabilire la **complessità** del modello, e di conseguenza il numero di parametri liberi, sulla base dei dati disponibili per l'addestramento.

Scelta dell'architettura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento

- Scelta dell'architettura

- Stabilizzazione strutturale

- Tecnica di

Regolarizzazione

- Early stopping

- Addestramento

- Classi di algoritmi

- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Scopo dell'addestramento non è interpolare i dati del training set, ma ottenere una buona capacità di generalizzazione.

Uno dei problemi più difficili è quello di stabilire la **complessità** del modello, e di conseguenza il numero di parametri liberi, sulla base dei dati disponibili per l'addestramento.

Per le reti multistrato esistono delle stime teoriche del numero di campioni del training set necessari per garantire una buona capacità di generalizzazione. Ma sono spesso inadeguate, si usano quindi tecniche euristiche per scegliere la struttura della rete e dimensionare il training set.

Scelta dell'architettura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento

- Scelta dell'architettura

- Stabilizzazione strutturale

- Tecnica di

Regolarizzazione

- Early stopping

- Addestramento

- Classi di algoritmi

- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Scopo dell'addestramento non è interpolare i dati del training set, ma ottenere una buona capacità di generalizzazione.

Uno dei problemi più difficili è quello di stabilire la **complessità** del modello, e di conseguenza il numero di parametri liberi, sulla base dei dati disponibili per l'addestramento.

Per le reti multistrato esistono delle stime teoriche del numero di campioni del training set necessari per garantire una buona capacità di generalizzazione. Ma sono spesso inadeguate, si usano quindi tecniche euristiche per scegliere la struttura della rete e dimensionare il training set.

Esistono due strategie fondamentali:

Scelta dell'architettura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento

- Scelta dell'architettura

- Stabilizzazione strutturale

- Tecnica di

Regolarizzazione

- Early stopping

- Addestramento

- Classi di algoritmi

- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Scopo dell'addestramento non è interpolare i dati del training set, ma ottenere una buona capacità di generalizzazione.

Uno dei problemi più difficili è quello di stabilire la **complessità** del modello, e di conseguenza il numero di parametri liberi, sulla base dei dati disponibili per l'addestramento.

Per le reti multistrato esistono delle stime teoriche del numero di campioni del training set necessari per garantire una buona capacità di generalizzazione. Ma sono spesso inadeguate, si usano quindi tecniche euristiche per scegliere la struttura della rete e dimensionare il training set.

Esistono due strategie fondamentali:

1. Stabilizzazione strutturale

Scelta dell'architettura

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento

- Scelta dell'architettura

- Stabilizzazione strutturale

- Tecnica di

Regolarizzazione

- Early stopping

- Addestramento

- Classi di algoritmi

- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Scopo dell'addestramento non è interpolare i dati del training set, ma ottenere una buona capacità di generalizzazione.

Uno dei problemi più difficili è quello di stabilire la **complessità** del modello, e di conseguenza il numero di parametri liberi, sulla base dei dati disponibili per l'addestramento.

Per le reti multistrato esistono delle stime teoriche del numero di campioni del training set necessari per garantire una buona capacità di generalizzazione. Ma sono spesso inadeguate, si usano quindi tecniche euristiche per scegliere la struttura della rete e dimensionare il training set.

Esistono due strategie fondamentali:

1. **Stabilizzazione strutturale**
2. **Tecnica di regolarizzazione**

Stabilizzazione strutturale

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

Stabilizzazione strutturale

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

Stabilizzazione strutturale

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

Stabilizzazione strutturale

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

test set

Stabilizzazione strutturale

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

test set

2. Si sceglie il numero di unità facendo crescere/diminuire il numero di neuroni

Stabilizzazione strutturale

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

test set

2. Si sceglie il numero di unità facendo crescere/diminuire il numero di neuroni

3. Per ogni numero di neuroni, si addestra la rete corrispondente usando sempre il training set.

Stabilizzazione strutturale

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

test set

2. Si sceglie il numero di unità facendo crescere/diminuire il numero di neuroni
3. Per ogni numero di neuroni, si addestra la rete corrispondente usando sempre il training set.
4. Si valuta l'architettura scelta tramite il validation set (**cross validation**) e si sceglie quella che dà l'errore più basso.

Stabilizzazione strutturale

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:
training set
validation set
test set
2. Si sceglie il numero di unità facendo crescere/diminuire il numero di neuroni
3. Per ogni numero di neuroni, si addestra la rete corrispondente usando sempre il training set.
4. Si valuta l'architettura scelta tramite il validation set (**cross validation**) e si sceglie quella che dà l'errore più basso.
5. La migliore architettura viene valutata tramite il test set.

Tecnica di Regolarizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
 - **Tecnica di Regolarizzazione**
 - Early stopping
 - Addestramento
 - Classi di algoritmi
 - Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

Tecnica di Regularizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- **Tecnica di Regularizzazione**
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

Tecnica di Regularizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- **Tecnica di Regularizzazione**
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

Tecnica di Regularizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- **Tecnica di Regularizzazione**
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

test set

Tecnica di Regolarizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- **Tecnica di Regolarizzazione**
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

test set

2. Si aggiunge alla funzione di errore un termine di penalità nella norma di w (equivale a fare ipotesi di regolarità sulle funzioni realizzate dalla rete):

$$E(w) = \sum_{p=1}^P E_p(w) + \gamma \|w\|^2$$

Tecnica di Regolarizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- **Tecnica di Regolarizzazione**
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

test set

2. Si aggiunge alla funzione di errore un termine di penalità nella norma di w (equivale a fare ipotesi di regolarità sulle funzioni realizzate dalla rete):

$$E(w) = \sum_{p=1}^P E_p(w) + \gamma \|w\|^2$$

3. Per diversi valori di γ , si addestra la rete usando sempre il training set.

Tecnica di Regolarizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- **Tecnica di Regolarizzazione**
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

test set

2. Si aggiunge alla funzione di errore un termine di penalità nella norma di w (equivale a fare ipotesi di regolarità sulle funzioni realizzate dalla rete):

$$E(w) = \sum_{p=1}^P E_p(w) + \gamma \|w\|^2$$

3. Per diversi valori di γ , si addestra la rete usando sempre il training set.
4. Si sceglie il valore migliore di γ tramite **cross validation**.

Tecnica di Regolarizzazione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- **Tecnica di Regolarizzazione**
- Early stopping
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

1. Si dividono i dati in 3 insiemi:

training set

validation set

test set

2. Si aggiunge alla funzione di errore un termine di penalità nella norma di w (equivale a fare ipotesi di regolarità sulle funzioni realizzate dalla rete):

$$E(w) = \sum_{p=1}^P E_p(w) + \gamma \|w\|^2$$

3. Per diversi valori di γ , si addestra la rete usando sempre il training set.
4. Si sceglie il valore migliore di γ tramite **cross validation**.
5. La rete risultante viene valutata tramite il test set.

Early stopping

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- **Early stopping**
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Un'alternativa alla regolarizzazione è l'**early stopping**:

Early stopping

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- **Early stopping**
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Un'alternativa alla regolarizzazione è l'**early stopping**:

1. Si interrompe prematuramente la minimizzazione della funzione di errore.

Early stopping

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- **Early stopping**
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Un'alternativa alla regolarizzazione è l'**early stopping**:

1. Si interrompe prematuramente la minimizzazione della funzione di errore.
2. Si valuta la rete sul validation set.

Early stopping

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- **Early stopping**
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Un'alternativa alla regolarizzazione è l'**early stopping**:

1. Si interrompe prematuramente la minimizzazione della funzione di errore.
2. Si valuta la rete sul validation set.
3. All'inizio l'errore sul validation set diminuisce al migliorare dell'errore sul training.

Early stopping

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- **Early stopping**
- Addestramento
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Un'alternativa alla regolarizzazione è l'**early stopping**:

1. Si interrompe prematuramente la minimizzazione della funzione di errore.
2. Si valuta la rete sul validation set.
3. All'inizio l'errore sul validation set diminuisce al migliorare dell'errore sul training.
4. Se si ha il fenomeno dell'overfitting, l'errore sul validation set sale. A questo punto si interrompe il processo.

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- Early stopping
- **Addestramento**
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Addestrare una rete multistrato richiede in ogni caso di risolvere (più volte) il problema di ottimizzazione non vincolata (con o senza termine di regolarizzazione):

$$\min_{w \in \mathbb{R}} E(w) \quad (1)$$

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- Early stopping
- **Addestramento**
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Addestrare una rete multistrato richiede in ogni caso di risolvere (più volte) il problema di ottimizzazione non vincolata (con o senza termine di regolarizzazione):

$$\min_{w \in \mathbb{R}} E(w) \quad (1)$$

Un algoritmo efficiente è quindi uno strumento essenziale, e deve tenere conto delle difficoltà intrinseche del problema (1):

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- Early stopping
- **Addestramento**
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Addestrare una rete multistrato richiede in ogni caso di risolvere (più volte) il problema di ottimizzazione non vincolata (con o senza termine di regolarizzazione):

$$\min_{w \in \mathbb{R}} E(w) \quad (1)$$

Un algoritmo efficiente è quindi uno strumento essenziale, e deve tenere conto delle difficoltà intrinseche del problema (1):

1. forti non linearità di $E(w)$

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- Early stopping
- **Addestramento**
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Addestrare una rete multistrato richiede in ogni caso di risolvere (più volte) il problema di ottimizzazione non vincolata (con o senza termine di regolarizzazione):

$$\min_{w \in \mathbb{R}} E(w) \quad (1)$$

Un algoritmo efficiente è quindi uno strumento essenziale, e deve tenere conto delle difficoltà intrinseche del problema (1):

1. forti non linearità di $E(w)$
2. dimensioni di w elevate e numero di campioni P molto grande

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- Early stopping
- **Addestramento**
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Addestrare una rete multistrato richiede in ogni caso di risolvere (più volte) il problema di ottimizzazione non vincolata (con o senza termine di regolarizzazione):

$$\min_{w \in \mathbb{R}} E(w) \quad (1)$$

Un algoritmo efficiente è quindi uno strumento essenziale, e deve tenere conto delle difficoltà intrinseche del problema (1):

1. forti non linearità di $E(w)$
2. dimensioni di w elevate e numero di campioni P molto grande
3. possibile mal condizionamento dell'hessiana di $E(w)$

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

● Struttura e addestramento

● Scelta dell'architettura
● Stabilizzazione strutturale

● Tecnica di Regularizzazione

● Early stopping

● **Addestramento**

● Classi di algoritmi

● Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Addestrare una rete multistrato richiede in ogni caso di risolvere (più volte) il problema di ottimizzazione non vincolata (con o senza termine di regolarizzazione):

$$\min_{w \in \mathbb{R}} E(w) \quad (1)$$

Un algoritmo efficiente è quindi uno strumento essenziale, e deve tenere conto delle difficoltà intrinseche del problema (1):

1. forti non linearità di $E(w)$
2. dimensioni di w elevate e numero di campioni P molto grande
3. possibile mal condizionamento dell'hessiana di $E(w)$
4. presenza di minimi locali non globali

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regularizzazione
- Early stopping
- **Addestramento**
- Classi di algoritmi
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Addestrare una rete multistrato richiede in ogni caso di risolvere (più volte) il problema di ottimizzazione non vincolata (con o senza termine di regolarizzazione):

$$\min_{w \in \mathbb{R}} E(w) \quad (1)$$

Un algoritmo efficiente è quindi uno strumento essenziale, e deve tenere conto delle difficoltà intrinseche del problema (1):

1. forti non linearità di $E(w)$
2. dimensioni di w elevate e numero di campioni P molto grande
3. possibile mal condizionamento dell'hessiana di $E(w)$
4. presenza di minimi locali non globali
5. insiemi di livello non compatti (salvo caso in cui si ha il termine di regolarizzazione)

Classi di algoritmi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- **Classi di algoritmi**
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Esistono due classi di algoritmi di addestramento:

Classi di algoritmi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- **Classi di algoritmi**
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Esistono due classi di algoritmi di addestramento:

1. metodi **batch**: ad ogni passo i pesi vengono aggiornati utilizzando tutti i campioni del training set (metodi di ottimizzazione non vincolata)

Classi di algoritmi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- **Classi di algoritmi**
- Metodi batch

Metodi on-line

Reti RBF

Addestramento RBF

Esistono due classi di algoritmi di addestramento:

1. metodi **batch**: ad ogni passo i pesi vengono aggiornati utilizzando tutti i campioni del training set (metodi di ottimizzazione non vincolata)
2. metodi **on-line**: ad ogni passo i pesi vengono aggiornati tenendo conto di un solo campione del training set (corrispondono ai metodi incrementali, che utilizzano cioè solo un'informazione parziale sulla funzione obiettivo e le sue derivate).

Metodi batch

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- **Metodi batch**

Metodi on-line

Reti RBF

Addestramento RBF

◁ Sono riconducibili a tecniche di ottimizzazione non vincolata, garantiscono convergenza a punti stazionari (buona riduzione dell'errore rispetto alla stima iniziale)

Metodi batch

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- **Metodi batch**

Metodi on-line

Reti RBF

Addestramento RBF

- ◁ Sono riconducibili a tecniche di ottimizzazione non vincolata, garantiscono convergenza a punti stazionari (buona riduzione dell'errore rispetto alla stima iniziale)
- ◁ Tipicamente utilizzano soltanto le derivate prime (o troppo dispendiosi)

Metodi batch

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

- Struttura e addestramento
- Scelta dell'architettura
- Stabilizzazione strutturale
- Tecnica di Regolarizzazione
- Early stopping
- Addestramento
- Classi di algoritmi
- **Metodi batch**

Metodi on-line

Reti RBF

Addestramento RBF

- ◁ Sono riconducibili a tecniche di ottimizzazione non vincolata, garantiscono convergenza a punti stazionari (buona riduzione dell'errore rispetto alla stima iniziale)
- ◁ Tipicamente utilizzano soltanto le derivate prime (o troppo dispendiosi)
- ◁ Sono utilizzabili solo fuori linea, e si deve avere a disposizione l'intero training set

Metodi on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- **Metodi on-line**
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

◁ Il vettore w viene aggiornato iterativamente in corrispondenza a ogni singolo termine $E_p(w)$ della funzione di errore senza formare la funzione complessiva $E(w)$.

Metodi on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- **Metodi on-line**
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

- ◁ Il vettore w viene aggiornato iterativamente in corrispondenza a ogni singolo termine $E_p(w)$ della funzione di errore senza formare la funzione complessiva $E(w)$.
- ◁ Per problemi di apprendimento in tempo reale sono l'unica opzione possibile (on-line learning).

Metodi on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- **Metodi on-line**
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

- ◁ Il vettore w viene aggiornato iterativamente in corrispondenza a ogni singolo termine $E_p(w)$ della funzione di errore senza formare la funzione complessiva $E(w)$.
- ◁ Per problemi di apprendimento in tempo reale sono l'unica opzione possibile (on-line learning).
- ◁ La convergenza può essere caratterizzata solo in termini probabilistici.

Metodi on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

- ◁ Il vettore w viene aggiornato iterativamente in corrispondenza a ogni singolo termine $E_p(w)$ della funzione di errore senza formare la funzione complessiva $E(w)$.
- ◁ Per problemi di apprendimento in tempo reale sono l'unica opzione possibile (on-line learning).
- ◁ La convergenza può essere caratterizzata solo in termini probabilistici.
- ◁ Presentano alcuni vantaggi anche nell'addestramento fuori linea:

Metodi on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

- ◁ Il vettore w viene aggiornato iterativamente in corrispondenza a ogni singolo termine $E_p(w)$ della funzione di errore senza formare la funzione complessiva $E(w)$.
- ◁ Per problemi di apprendimento in tempo reale sono l'unica opzione possibile (on-line learning).
- ◁ La convergenza può essere caratterizzata solo in termini probabilistici.
- ◁ Presentano alcuni vantaggi anche nell'addestramento fuori linea:
 1. a grande distanza dalla soluzione può non essere conveniente calcolare tutta la funzione e le sue derivate: conviene ridurre rapidamente un singolo termine di errore

Metodi on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

- ◁ Il vettore w viene aggiornato iterativamente in corrispondenza a ogni singolo termine $E_p(w)$ della funzione di errore senza formare la funzione complessiva $E(w)$.
- ◁ Per problemi di apprendimento in tempo reale sono l'unica opzione possibile (on-line learning).
- ◁ La convergenza può essere caratterizzata solo in termini probabilistici.
- ◁ Presentano alcuni vantaggi anche nell'addestramento fuori linea:
 1. a grande distanza dalla soluzione può non essere conveniente calcolare tutta la funzione e le sue derivate: conviene ridurre rapidamente un singolo termine di errore
 2. se il training set e il numero di campioni sono molto elevati il calcolo di E e ∇E può essere troppo dispendioso

Metodi on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

- ◁ Il vettore w viene aggiornato iterativamente in corrispondenza a ogni singolo termine $E_p(w)$ della funzione di errore senza formare la funzione complessiva $E(w)$.
- ◁ Per problemi di apprendimento in tempo reale sono l'unica opzione possibile (on-line learning).
- ◁ La convergenza può essere caratterizzata solo in termini probabilistici.
- ◁ Presentano alcuni vantaggi anche nell'addestramento fuori linea:
 1. a grande distanza dalla soluzione può non essere conveniente calcolare tutta la funzione e le sue derivate: conviene ridurre rapidamente un singolo termine di errore
 2. se il training set e il numero di campioni sono molto elevati il calcolo di E e ∇E può essere troppo dispendioso
 3. il training set potrebbe essere ridondante

Metodi on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

- ◁ Il vettore w viene aggiornato iterativamente in corrispondenza a ogni singolo termine $E_p(w)$ della funzione di errore senza formare la funzione complessiva $E(w)$.
- ◁ Per problemi di apprendimento in tempo reale sono l'unica opzione possibile (on-line learning).
- ◁ La convergenza può essere caratterizzata solo in termini probabilistici.
- ◁ Presentano alcuni vantaggi anche nell'addestramento fuori linea:
 1. a grande distanza dalla soluzione può non essere conveniente calcolare tutta la funzione e le sue derivate: conviene ridurre rapidamente un singolo termine di errore
 2. se il training set e il numero di campioni sono molto elevati il calcolo di E e ∇E può essere troppo dispendioso
 3. il training set potrebbe essere ridondante
 4. la componente di randomicità di un algoritmo on-line può permettere di evitare minimi locali

Backpropagation on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- **Backpropagation
on-line**
- Backpropagation
on-line
- Metodi globali

Reti RBF

Addestramento RBF

Il primo esempio di metodo on-line è l'addestramento del perceptron

Backpropagation on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- **Backpropagation on-line**
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

Il primo esempio di metodo on-line è l'addestramento del perceptron

Per le reti multistrato il metodo più noto è la versione on-line della backpropagation, ed è descritto da un'iterazione del tipo

$$w_{k+1} = w_k - \eta_k \nabla E_{p(k)}(w_k)$$

Backpropagation on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

- Metodi on-line
- **Backpropagation on-line**
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

Il primo esempio di metodo on-line è l'addestramento del perceptron

Per le reti multistrato il metodo più noto è la versione on-line della backpropagation, ed è descritto da un'iterazione del tipo

$$w_{k+1} = w_k - \eta_k \nabla E_{p(k)}(w_k)$$

Per ottenere la convergenza, tutti i termini E_p devono essere considerati ciclicamente, ovvero $p(k) = k(\text{mod}P) + 1$, e il passo η_k deve essere scelto in modo da soddisfare:

$$\sum_{k=0}^{\infty} \eta_k = \infty, \quad \sum_{k=0}^{\infty} (\eta_k)^2 < \infty$$

Backpropagation on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

- Metodi on-line
- **Backpropagation on-line**
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

Il primo esempio di metodo on-line è l'addestramento del perceptron

Per le reti multistrato il metodo più noto è la versione on-line della backpropagation, ed è descritto da un'iterazione del tipo

$$w_{k+1} = w_k - \eta_k \nabla E_{p(k)}(w_k)$$

Per ottenere la convergenza, tutti i termini E_p devono essere considerati ciclicamente, ovvero $p(k) = k(\text{mod}P) + 1$, e il passo η_k deve essere scelto in modo da soddisfare:

$$\sum_{k=0}^{\infty} \eta_k = \infty, \quad \sum_{k=0}^{\infty} (\eta_k)^2 < \infty$$

Un esempio di scelta di η_k è $\eta_k = c/k$ con $c > 0$.

Backpropagation on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

Il primo esempio di metodo on-line è l'addestramento del perceptron

Per le reti multistrato il metodo più noto è la versione on-line della backpropagation, ed è descritto da un'iterazione del tipo

$$w_{k+1} = w_k - \eta_k \nabla E_{p(k)}(w_k)$$

Per ottenere la convergenza, tutti i termini E_p devono essere considerati ciclicamente, ovvero $p(k) = k(\text{mod}P) + 1$, e il passo η_k deve essere scelto in modo da soddisfare:

$$\sum_{k=0}^{\infty} \eta_k = \infty, \quad \sum_{k=0}^{\infty} (\eta_k)^2 < \infty$$

Un esempio di scelta di η_k è $\eta_k = c/k$ con $c > 0$.

Esiste anche la versione on-line della momentum updating rule (con proprietà di convergenza analoghe):

$$w_{k+1} = w_k - \eta_k \nabla E_{p(k)}(w_k) + \beta(w_k - w_{k-1})$$

Backpropagation on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- **Backpropagation on-line**
- Metodi globali

Reti RBF

Addestramento RBF

La principale limitazione è la lentezza (dovuta alla scelta del passo che deve tendere a zero con una legge prefissata)

Backpropagation on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- **Backpropagation on-line**
- Metodi globali

Reti RBF

Addestramento RBF

La principale limitazione è la lentezza (dovuta alla scelta del passo che deve tendere a zero con una legge prefissata)

Nell'addestramento fuori-linea un compromesso è l'uso di tecniche ibride on-line-batch chiamate **bold driver methods**: la backpropagation on-line è usata per uno o più cicli di iterazioni e il passo è ricalcolato ciclicamente usando la funzione di errore complessiva (esistono risultati di convergenza). Queste tecniche però richiedono il calcolo di E e ∇E .

Backpropagation on-line

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- **Backpropagation on-line**
- Metodi globali

Reti RBF

Addestramento RBF

La principale limitazione è la lentezza (dovuta alla scelta del passo che deve tendere a zero con una legge prefissata)

Nell'addestramento fuori-linea un compromesso è l'uso di tecniche ibride on-line-batch chiamate **bold driver methods**: la backpropagation on-line è usata per uno o più cicli di iterazioni e il passo è ricalcolato ciclicamente usando la funzione di errore complessiva (esistono risultati di convergenza). Queste tecniche però richiedono il calcolo di E e ∇E .

Un altro tipo di compromesso si ottiene passando gradualmente dalla backpropagation on-line a quella batch, per accelerare la convergenza.

Metodi globali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

La presenza di minimi locali non globali della funzione di errore è stata provata per alcuni specifici problemi test e rappresenta una delle difficoltà dei problemi di addestramento

Metodi globali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- Backpropagation on-line
- Metodi globali

Reti RBF

Addestramento RBF

La presenza di minimi locali non globali della funzione di errore è stata provata per alcuni specifici problemi test e rappresenta una delle difficoltà dei problemi di addestramento

NB: non esistono minimi locali non globali nel caso di problemi di classificazione di insiemi linearmente separabili.

Metodi globali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- Backpropagation
on-line
- Backpropagation
on-line

● Metodi globali

Reti RBF

Addestramento RBF

La presenza di minimi locali non globali della funzione di errore è stata provata per alcuni specifici problemi test e rappresenta una delle difficoltà dei problemi di addestramento

NB: non esistono minimi locali non globali nel caso di problemi di classificazione di insiemi linearmente separabili.

Spesso la difficoltà maggiore è la presenza di plateau.

Metodi globali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- Backpropagation on-line

● Metodi globali

Reti RBF

Addestramento RBF

La presenza di minimi locali non globali della funzione di errore è stata provata per alcuni specifici problemi test e rappresenta una delle difficoltà dei problemi di addestramento

NB: non esistono minimi locali non globali nel caso di problemi di classificazione di insiemi linearmente separabili.

Spesso la difficoltà maggiore è la presenza di plateau.

In realtà quello che si vuole è un errore ragionevolmente basso sul training set che garantisca una buona capacità di generalizzazione.

Metodi globali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

- Metodi on-line
- Backpropagation on-line
- Backpropagation on-line

● **Metodi globali**

Reti RBF

Addestramento RBF

La presenza di minimi locali non globali della funzione di errore è stata provata per alcuni specifici problemi test e rappresenta una delle difficoltà dei problemi di addestramento

NB: non esistono minimi locali non globali nel caso di problemi di classificazione di insiemi linearmente separabili.

Spesso la difficoltà maggiore è la presenza di plateau.

In realtà quello che si vuole è un errore ragionevolmente basso sul training set che garantisca una buona capacità di generalizzazione.

Per questo i metodi globali (troppo onerosi a grandi dimensioni) non sono molto usati. A volte si utilizza una prima fase stocastica per eliminare dei minimi non interessanti e poi si usano metodi locali.

Funzioni di base radiali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

● Funzioni di base
radiali

● Funzioni di base
radiali

● Interpolazione II

● Reti RBF

regolarizzate

● RBF regolarizzate

● RBF regolarizzate

● RBF generalizzate

● Esempio

Addestramento RBF

Problema di interpolazione: dati P punti $\{x^p \in \mathbb{R}^n, p = 1, \dots, P\}$ e un corrispondente insieme di numeri reali $\{y^p \in \mathbb{R}, p = 1, \dots, P\}$ determinare una funzione $y : \mathbb{R}^n \Rightarrow \mathbb{R}$ in una classe di funzioni assegnate tale che

$$y(x^p) = y^p, \quad p = 1, \dots, P$$

Funzioni di base radiali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

● Funzioni di base radiali

● Funzioni di base radiali

● Interpolazione II

● Reti RBF regolarizzate

● RBF regolarizzate

● RBF regolarizzate

● RBF generalizzate

● Esempio

Addestramento RBF

Problema di interpolazione: dati P punti $\{x^p \in \mathbb{R}^n, p = 1, \dots, P\}$ e un corrispondente insieme di numeri reali $\{y^p \in \mathbb{R}, p = 1, \dots, P\}$ determinare una funzione $y : \mathbb{R}^n \Rightarrow \mathbb{R}$ in una classe di funzioni assegnate tale che

$$y(x^p) = y^p, \quad p = 1, \dots, P$$

Una particolare tecnica di interpolazione consiste nella scelta di y della forma:

$$y(x) = \sum_{p=1}^P w_p \phi(\|x - x^p\|)$$

dove $\phi : \mathbb{R}^+ \Rightarrow \mathbb{R}^+$ è una funzione continua, chiamata **funzione di base radiale**. I coefficienti w_p sono detti **pesi** e i punti di interpolazione **centri**.

Funzioni di base radiali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base
radiali

- Funzioni di base
radiali

- Interpolazione II

- Reti RBF
regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Esempi di funzioni di base radiali sono:

Funzioni di base radiali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF
regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Esempi di funzioni di base radiali sono:

1. $\phi(r) = r$ lineare

Funzioni di base radiali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF
regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Esempi di funzioni di base radiali sono:

1. $\phi(r) = r$ lineare
2. $\phi(r) = e^{-r^2}$ gaussiana

Funzioni di base radiali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF
regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Esempi di funzioni di base radiali sono:

1. $\phi(r) = r$ lineare
2. $\phi(r) = e^{-r^2}$ gaussiana
3. $\phi(r) = (r^2 + \sigma^2)^{-1/2}$ multiquadrica inversa

Funzioni di base radiali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF

regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Esempi di funzioni di base radiali sono:

1. $\phi(r) = r$ lineare
2. $\phi(r) = e^{-r^2}$ gaussiana
3. $\phi(r) = (r^2 + \sigma^2)^{-1/2}$ multiquadrica inversa
4. $\phi(r) = (r^2 + \sigma^2)^{1/2}$ multiquadrica diretta

Funzioni di base radiali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF

regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Esempi di funzioni di base radiali sono:

1. $\phi(r) = r$ lineare
2. $\phi(r) = e^{-r^2}$ gaussiana
3. $\phi(r) = (r^2 + \sigma^2)^{-1/2}$ multiquadrica inversa
4. $\phi(r) = (r^2 + \sigma^2)^{1/2}$ multiquadrica diretta

dove $r \geq 0$ e σ è una costante positiva.

Funzioni di base radiali

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Esempi di funzioni di base radiali sono:

1. $\phi(r) = r$ lineare
2. $\phi(r) = e^{-r^2}$ gaussiana
3. $\phi(r) = (r^2 + \sigma^2)^{-1/2}$ multiquadrica inversa
4. $\phi(r) = (r^2 + \sigma^2)^{1/2}$ multiquadrica diretta

dove $r \geq 0$ e σ è una costante positiva.

Scelta la funzione ϕ , usando l'espressione di $y(x)$ si ottiene la condizione di interpolazione:

$$\Phi w = y$$

dove $y = (y^1, \dots, y^P)^T$, Φ (**matrice di interpolazione**) è la matrice $P \times P$ il cui generico elemento è

$$\Phi_{ij} = \phi(\|x^i - x^j\|)$$

Interpolazione II

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base
radiali

- Funzioni di base
radiali

- Interpolazione II

- Reti RBF
regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Sotto opportune ipotesi sui punti di interpolazione (che devono essere tutti distinti), si dimostra che la matrice Φ è non singolare.

Interpolazione II

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base
radiali

- Funzioni di base
radiali

- Interpolazione II

- Reti RBF
regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Sotto opportune ipotesi sui punti di interpolazione (che devono essere tutti distinti), si dimostra che la matrice Φ è non singolare.

In questo caso si dimostra l'invertibilità della matrice Φ per tutte le scelte di ϕ riportate, e per la gaussiana e per la multiquadrica inversa si dimostra che le corrispondenti matrici di interpolazione sono definite positive.

Reti RBF regolarizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base
radiali

- Funzioni di base
radiali

- Interpolazione II

- Reti RBF
regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Le tecniche di **regolarizzazione** sono studiate nell'ambito dei problemi di approssimazione con insieme di campioni disponibili finito.

Reti RBF regolarizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Le tecniche di **regolarizzazione** sono studiate nell'ambito dei problemi di approssimazione con insieme di campioni disponibili finito.

Supponiamo di voler approssimare una funzione $f : \mathbb{R}^n \Rightarrow \mathbb{R}$, e sia dato il training set

$$\mathcal{T} = \{(x^p, y^p), p = 1, \dots, P, y^p = f(x^p)\}$$

Esistono infinite funzioni che lo risolvono e ci possono essere errori di misura nel training set.

Reti RBF regolarizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Le tecniche di **regolarizzazione** sono studiate nell'ambito dei problemi di approssimazione con insieme di campioni disponibili finito.

Supponiamo di voler approssimare una funzione $f : \mathbb{R}^n \Rightarrow \mathbb{R}$, e sia dato il training set

$$\mathcal{T} = \{(x^p, y^p), p = 1, \dots, P, y^p = f(x^p)\}$$

Esistono infinite funzioni che lo risolvono e ci possono essere errori di misura nel training set.

Le tecniche di regolarizzazione determinano la funzione approssimante minimizzando un funzionale costituito da due termini: il primo misura la distanza da \mathcal{T} , il secondo penalizza la violazione delle condizioni di regolarità sull'approssimante (sfrutta le informazioni note):

$$\mathcal{E}(y) = \mathcal{E}_1(y) + \mathcal{E}_2(y) = \frac{1}{2} \sum_{p=1}^P [y^p - y(x^p)]^2 + \frac{1}{2} \lambda \|\mathcal{P}y\|^2$$

dove λ è il parametro di regolarizzazione e \mathcal{P} è un operatore differenziale che contiene le informazioni che si hanno sulla funzione da approssimare.

RBF regolarizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base
radiali

- Funzioni di base
radiali

- Interpolazione II

- Reti RBF

regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Sotto opportune ipotesi su \mathcal{P} , si dimostra che la $y(x)$ che minimizza $\mathcal{E}(y)$ è

$$y(x) = \sum_{p=1}^P w_p \phi(\|x - x^p\|)$$

con ϕ funzione di base radiale e w è soluzione del sistema lineare:

$$(\Phi + \lambda I)w = y$$

RBF regolarizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base
radiali

- Funzioni di base
radiali

- Interpolazione II

- Reti RBF
regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

Sotto opportune ipotesi su \mathcal{P} , si dimostra che la $y(x)$ che minimizza $\mathcal{E}(y)$ è

$$y(x) = \sum_{p=1}^P w_p \phi(\|x - x^p\|)$$

con ϕ funzione di base radiale e w è soluzione del sistema lineare:

$$(\Phi + \lambda I)w = y$$

Questa funzione y può essere interpretata come l'uscita di una rete feedforward in cui la funzione di attivazione dei neuroni dello strato nascosto è la funzione ϕ .

RBF regolarizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali
- Funzioni di base radiali
- Interpolazione II
- Reti RBF regolarizzate
- RBF regolarizzate
- RBF regolarizzate
- RBF generalizzate
- Esempio

Addestramento RBF

1. hanno un solo strato nascosto

RBF regolarizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali
- Funzioni di base radiali
- Interpolazione II
- Reti RBF regolarizzate
- RBF regolarizzate
- RBF regolarizzate
- RBF generalizzate
- Esempio

Addestramento RBF

1. hanno un solo strato nascosto
2. i neuroni dello strato nascosto hanno come funzione di attivazione una funzione di base radiale e l'argomento della funzione è data dalla distanza tra il vettore di ingresso e il centro dell'unità

RBF regolarizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali
- Funzioni di base radiali
- Interpolazione II
- Reti RBF regolarizzate
- RBF regolarizzate
- RBF regolarizzate
- RBF generalizzate
- Esempio

Addestramento RBF

1. hanno un solo strato nascosto
2. i neuroni dello strato nascosto hanno come funzione di attivazione una funzione di base radiale e l'argomento della funzione è data dalla distanza tra il vettore di ingresso e il centro dell'unità
3. il numero di neuroni dello strato nascosto è pari al numero di elementi del training set P

RBF regolarizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali
- Funzioni di base radiali
- Interpolazione II
- Reti RBF regolarizzate
- RBF regolarizzate
- RBF regolarizzate
- RBF generalizzate
- Esempio

Addestramento RBF

1. hanno un solo strato nascosto
2. i neuroni dello strato nascosto hanno come funzione di attivazione una funzione di base radiale e l'argomento della funzione è data dalla distanza tra il vettore di ingresso e il centro dell'unità
3. il numero di neuroni dello strato nascosto è pari al numero di elementi del training set P
4. lo strato di uscita ha un unico neurone che effettua una combinazione lineare delle uscite dei neuroni dello strato nascosto.

RBF regolarizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- RBF generalizzate

- Esempio

Addestramento RBF

1. hanno un solo strato nascosto
2. i neuroni dello strato nascosto hanno come funzione di attivazione una funzione di base radiale e l'argomento della funzione è data dalla distanza tra il vettore di ingresso e il centro dell'unità
3. il numero di neuroni dello strato nascosto è pari al numero di elementi del training set P
4. lo strato di uscita ha un unico neurone che effettua una combinazione lineare delle uscite dei neuroni dello strato nascosto.

Per ogni funzione continua f definita su un sottoinsieme compatto S di \mathbb{R}^n esiste una rete RBF regolarizzata ovvero una funzione $y(x) = \sum_{p=1}^P w_p \phi(\|x - x^p\|)$ tale che per ogni $x \in S$ e per ogni $\epsilon > 0$ risulta

$$|f(x) - y(x)| < \epsilon$$

RBF generalizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali
- Funzioni di base radiali
- Interpolazione II
- Reti RBF regolarizzate
- RBF regolarizzate
- RBF regolarizzate
- **RBF generalizzate**
- Esempio

Addestramento RBF

La corrispondenza uno a uno tra elementi del training set e i termini della funzione $y(x)$ è troppo onerosa perchè determinare w richiede la soluzione di un sistema $P \times P$ che potrebbe essere proibitivo.

RBF generalizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali
- Funzioni di base radiali
- Interpolazione II
- Reti RBF regolarizzate
- RBF regolarizzate
- RBF regolarizzate
- **RBF generalizzate**
- Esempio

Addestramento RBF

La corrispondenza uno a uno tra elementi del training set e i termini della funzione $y(x)$ è troppo onerosa perchè determinare w richiede la soluzione di un sistema $P \times P$ che potrebbe essere proibitivo.

Per questo motivo sono state introdotte le reti RBF generalizzate in cui il numero N dei neuroni dello strato nascosto è di molto minore di P e i centri c_i non coincidono necessariamente con i vettori x^p del training set. La funzione approssimante diventa quindi:

$$y(x) = \sum_{i=1}^N w_i \phi(\|x - c_i\|)$$

RBF generalizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF

regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- **RBF generalizzate**

- Esempio

Addestramento RBF

La corrispondenza uno a uno tra elementi del training set e i termini della funzione $y(x)$ è troppo onerosa perchè determinare w richiede la soluzione di un sistema $P \times P$ che potrebbe essere proibitivo.

Per questo motivo sono state introdotte le reti RBF generalizzate in cui il numero N dei neuroni dello strato nascosto è di molto minore di P e i centri c_i non coincidono necessariamente con i vettori x^p del training set. La funzione approssimante diventa quindi:

$$y(x) = \sum_{i=1}^N w_i \phi(\|x - c_i\|)$$

Devono essere determinati sia w che i centri, quindi il legame ingresso-uscita dipende in modo nonlineare dai parametri come nelle reti multistrato

RBF generalizzate

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali

- Funzioni di base radiali

- Interpolazione II

- Reti RBF

regolarizzate

- RBF regolarizzate

- RBF regolarizzate

- **RBF generalizzate**

- Esempio

Addestramento RBF

La corrispondenza uno a uno tra elementi del training set e i termini della funzione $y(x)$ è troppo onerosa perchè determinare w richiede la soluzione di un sistema $P \times P$ che potrebbe essere proibitivo.

Per questo motivo sono state introdotte le reti RBF generalizzate in cui il numero N dei neuroni dello strato nascosto è di molto minore di P e i centri c_i non coincidono necessariamente con i vettori x^p del training set. La funzione approssimante diventa quindi:

$$y(x) = \sum_{i=1}^N w_i \phi(\|x - c_i\|)$$

Devono essere determinati sia w che i centri, quindi il legame ingresso-uscita dipende in modo nonlineare dai parametri come nelle reti multistrato

Le RBF generalizzate includono le RBF regolarizzate, quindi sono approssimatori universali

Esempio

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

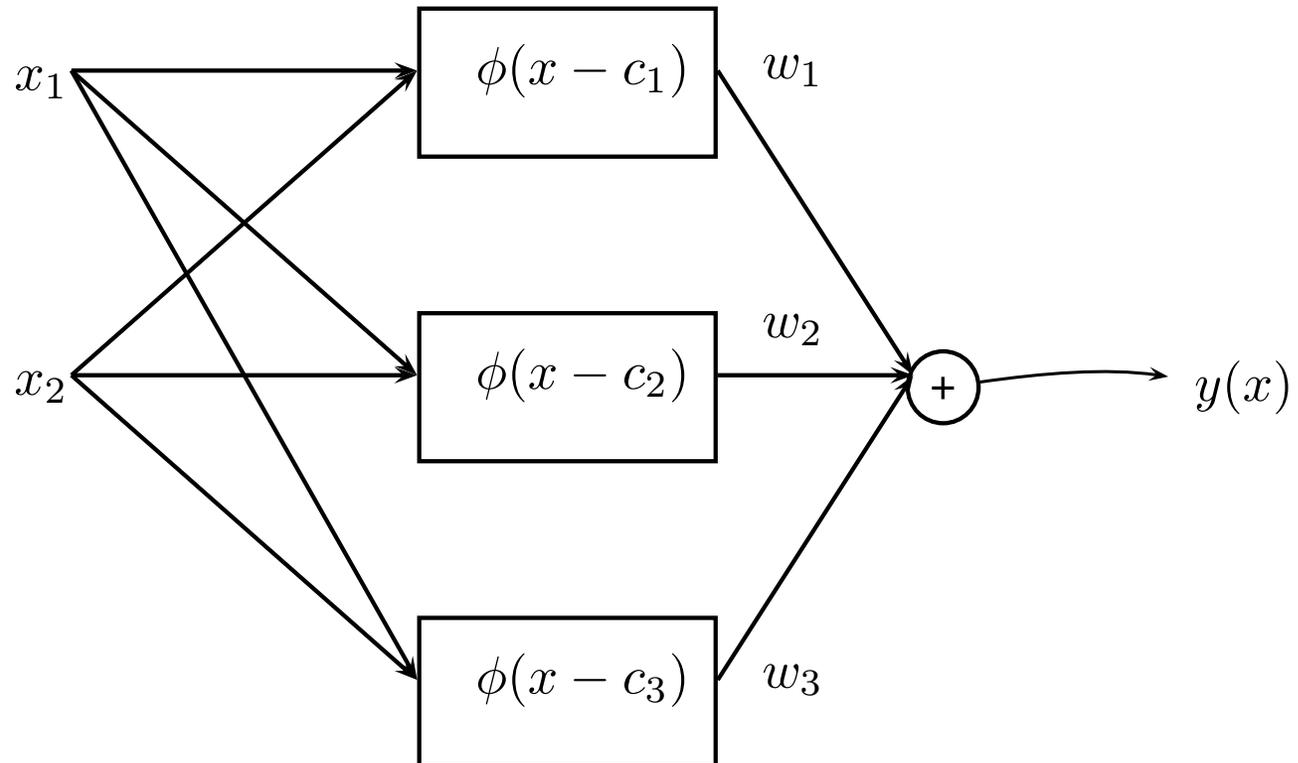
Creazione di una rete multistrato

Metodi on-line

Reti RBF

- Funzioni di base radiali
- Funzioni di base radiali
- Interpolazione II
- Reti RBF regolarizzate
- RBF regolarizzate
- RBF regolarizzate
- RBF generalizzate
- Esempio

Addestramento RBF



Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento

- Scelta non supervisionata dei centri

- Scelta supervisionata

- Tecniche di decomposizione

- Decomposizione per RBF

- Decomposizione in 2 blocchi

- Decomposizione in $N + 1$ blocchi

- Decomposizione in $N + 1$ blocchi

- Proprietà

Addestrare una rete RBF generalizzata vuol dire individuare il vettore $w \in \mathbb{R}^N$ e il vettore $c \in \mathbb{R}^{nN}$ sulla base del training set

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento

- Scelta non supervisionata dei centri

- Scelta supervisionata

- Tecniche di decomposizione

- Decomposizione per RBF

- Decomposizione in 2 blocchi

- Decomposizione in $N + 1$ blocchi

- Decomposizione in $N + 1$ blocchi

- Proprietà

Addestrare una rete RBF generalizzata vuol dire individuare il vettore $w \in \mathbb{R}^N$ e il vettore $c \in \mathbb{R}^{nN}$ sulla base del training set

Sono possibili due diverse strategie di addestramento:

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

● Addestramento

● Scelta non
supervisionata dei centri

● Scelta supervisionata

● Tecniche di
decomposizione

● Decomposizione per
RBF

● Decomposizione in 2
blocchi

● Decomposizione in
 $N + 1$ blocchi

● Decomposizione in
 $N + 1$ blocchi

● Proprietà

Addestrare una rete RBF generalizzata vuol dire individuare il vettore $w \in \mathbb{R}^N$ e il vettore $c \in \mathbb{R}^{nN}$ sulla base del training set

Sono possibili due diverse strategie di addestramento:

1. scelta **non supervisionata** dei centri e **supervisionata** dei pesi

Addestramento

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

● Addestramento

● Scelta non
supervisionata dei centri

● Scelta supervisionata

● Tecniche di
decomposizione

● Decomposizione per
RBF

● Decomposizione in 2
blocchi

● Decomposizione in
 $N + 1$ blocchi

● Decomposizione in
 $N + 1$ blocchi

● Proprietà

Addestrare una rete RBF generalizzata vuol dire individuare il vettore $w \in \mathbb{R}^N$ e il vettore $c \in \mathbb{R}^{nN}$ sulla base del training set

Sono possibili due diverse strategie di addestramento:

1. scelta **non supervisionata** dei centri e **supervisionata** dei pesi
2. scelta **supervisionata** sia dei centri che dei pesi

Scelta non supervisionata dei centri

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si fissano prima i centri o scegliendoli **casualmente** tra i vettori di ingresso del training set oppure con tecniche di **clustering**

Scelta non supervisionata dei centri

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si fissano prima i centri o scegliendoli **casualmente** tra i vettori di ingresso del training set oppure con tecniche di **clustering**

Si scelgono poi i pesi w minimizzando la funzione di errore

$$E(w) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho}{2} \|w\|^2, \quad \rho > 0$$

Scelta non supervisionata dei centri

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

● Addestramento

● Scelta non

supervisionata dei centri

● Scelta supervisionata

● Tecniche di decomposizione

● Decomposizione per RBF

● Decomposizione in 2 blocchi

● Decomposizione in

$N + 1$ blocchi

● Decomposizione in

$N + 1$ blocchi

● Proprietà

Si fissano prima i centri o scegliendoli **casualmente** tra i vettori di ingresso del training set oppure con tecniche di **clustering**

Si scelgono poi i pesi w minimizzando la funzione di errore

$$E(w) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho}{2} \|w\|^2, \quad \rho > 0$$

Ponendo $y = (y^1, \dots, y^P)^T$ e definendo la matrice $\Phi_{pi} = \phi(\|x^p - c_i\|)$, la minimizzazione di $E(w)$ si riduce al problema ai minimi quadrati lineari:

$$\min_w E(w) = \left\| \begin{pmatrix} \Phi(c) \\ \sqrt{\rho} I \end{pmatrix} w - \begin{pmatrix} y \\ 0 \end{pmatrix} \right\|^2.$$

Scelta non supervisionata dei centri

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

● Addestramento

● Scelta non

supervisionata dei centri

● Scelta supervisionata

● Tecniche di decomposizione

● Decomposizione per RBF

● Decomposizione in 2 blocchi

● Decomposizione in $N + 1$ blocchi

● Decomposizione in $N + 1$ blocchi

● Proprietà

Si fissano prima i centri o scegliendoli **casualmente** tra i vettori di ingresso del training set oppure con tecniche di **clustering**

Si scelgono poi i pesi w minimizzando la funzione di errore

$$E(w) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho}{2} \|w\|^2, \quad \rho > 0$$

Ponendo $y = (y^1, \dots, y^P)^T$ e definendo la matrice $\Phi_{pi} = \phi(\|x^p - c_i\|)$, la minimizzazione di $E(w)$ si riduce al problema ai minimi quadrati lineari:

$$\min_w E(w) = \left\| \begin{pmatrix} \Phi(c) \\ \sqrt{\rho} I \end{pmatrix} w - \begin{pmatrix} y \\ 0 \end{pmatrix} \right\|^2.$$

Si possono usare metodi diretti ($x = (AA^T)^{-1} A^T b$) se N non è troppo elevato se no metodi iterativi (gradiente coniugato).

Scelta non supervisionata dei centri

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

● Addestramento

● Scelta non

supervisionata dei centri

● Scelta supervisionata

● Tecniche di decomposizione

● Decomposizione per RBF

● Decomposizione in 2 blocchi

● Decomposizione in $N + 1$ blocchi

● Decomposizione in $N + 1$ blocchi

● Proprietà

Si fissano prima i centri o scegliendoli **casualmente** tra i vettori di ingresso del training set oppure con tecniche di **clustering**

Si scelgono poi i pesi w minimizzando la funzione di errore

$$E(w) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho}{2} \|w\|^2, \quad \rho > 0$$

Ponendo $y = (y^1, \dots, y^P)^T$ e definendo la matrice $\Phi_{pi} = \phi(\|x^p - c_i\|)$, la minimizzazione di $E(w)$ si riduce al problema ai minimi quadrati lineari:

$$\min_w E(w) = \left\| \begin{pmatrix} \Phi(c) \\ \sqrt{\rho} I \end{pmatrix} w - \begin{pmatrix} y \\ 0 \end{pmatrix} \right\|^2.$$

Si possono usare metodi diretti ($x = (AA^T)^{-1}A^T b$) se N non è troppo elevato se no metodi iterativi (gradiente coniugato).

Svantaggio: non si tiene conto delle uscite y^p nella scelta dei centri.

Scelta supervisionata

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- **Scelta supervisionata**
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si risolve direttamente il problema non vincolato:

$$\min_{w,c} E(w, c) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho_1}{2} \|w\|^2 + \frac{\rho_2}{2} \|c\|^2,$$

con $\rho_1, \rho_2 > 0$.

Scelta supervisionata

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- **Scelta supervisionata**
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si risolve direttamente il problema non vincolato:

$$\min_{w,c} E(w, c) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho_1}{2} \|w\|^2 + \frac{\rho_2}{2} \|c\|^2,$$

con $\rho_1, \rho_2 > 0$.

Con questa seconda tecnica si ottengono capacità di generalizzazione migliori, rispetto alla scelta non supervisionata dei centri.

Scelta supervisionata

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri

- **Scelta supervisionata**

- Tecniche di decomposizione

- Decomposizione per RBF

- Decomposizione in 2 blocchi

- Decomposizione in $N + 1$ blocchi

- Decomposizione in $N + 1$ blocchi

- Proprietà

Si risolve direttamente il problema non vincolato:

$$\min_{w,c} E(w, c) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho_1}{2} \|w\|^2 + \frac{\rho_2}{2} \|c\|^2,$$

con $\rho_1, \rho_2 > 0$.

Con questa seconda tecnica si ottengono capacità di generalizzazione migliori, rispetto alla scelta non supervisionata dei centri.

Le dimensioni del problema sono notevolmente più elevate: si utilizzano tecniche di decomposizione.

Tecniche di decomposizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- **Tecniche di decomposizione**
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si partizionano le variabili in blocchi e si minimizza rispetto al singolo blocco di variabili: se si sta risolvendo il problema $\min_x f(x)$ si partiziona il vettore x in NB blocchi x_1, \dots, x_{NB} e si definisce

$$x_i^{k+1} = \arg \min_y f(x_1^{k+1}, \dots, x_{i-1}^{k+1}, y, x_{i+1}^k, \dots, x_{NB}^k)$$

Tecniche di decomposizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si partizionano le variabili in blocchi e si minimizza rispetto al singolo blocco di variabili: se si sta risolvendo il problema $\min_x f(x)$ si partiziona il vettore x in NB blocchi x_1, \dots, x_{NB} e si definisce

$$x_i^{k+1} = \arg \min_y f(x_1^{k+1}, \dots, x_{i-1}^{k+1}, y, x_{i+1}^k, \dots, x_{NB}^k)$$

In generale, non converge, esiste un controesempio in 3 variabili per il metodo delle coordinate con ricerca di linea esatta.

Tecniche di decomposizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si partizionano le variabili in blocchi e si minimizza rispetto al singolo blocco di variabili: se si sta risolvendo il problema $\min_x f(x)$ si partiziona il vettore x in NB blocchi x_1, \dots, x_{NB} e si definisce

$$x_i^{k+1} = \arg \min_y f(x_1^{k+1}, \dots, x_{i-1}^{k+1}, y, x_{i+1}^k, \dots, x_{NB}^k)$$

In generale, non converge, esiste un controesempio in 3 variabili per il metodo delle coordinate con ricerca di linea esatta.

Converge in uno dei seguenti casi:

Tecniche di decomposizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si partizionano le variabili in blocchi e si minimizza rispetto al singolo blocco di variabili: se si sta risolvendo il problema $\min_x f(x)$ si partiziona il vettore x in NB blocchi x_1, \dots, x_{NB} e si definisce

$$x_i^{k+1} = \arg \min_y f(x_1^{k+1}, \dots, x_{i-1}^{k+1}, y, x_{i+1}^k, \dots, x_{NB}^k)$$

In generale, non converge, esiste un controesempio in 3 variabili per il metodo delle coordinate con ricerca di linea esatta.

Converge in uno dei seguenti casi:

1. Il vettore delle variabili è decomposto in $NB = 2$ blocchi

Tecniche di decomposizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento supervisionata dei centri
- Scelta non supervisionata
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si partizionano le variabili in blocchi e si minimizza rispetto al singolo blocco di variabili: se si sta risolvendo il problema $\min_x f(x)$ si partiziona il vettore x in NB blocchi x_1, \dots, x_{NB} e si definisce

$$x_i^{k+1} = \arg \min_y f(x_1^{k+1}, \dots, x_{i-1}^{k+1}, y, x_{i+1}^k, \dots, x_{NB}^k)$$

In generale, non converge, esiste un controesempio in 3 variabili per il metodo delle coordinate con ricerca di linea esatta.

Converge in uno dei seguenti casi:

1. Il vettore delle variabili è decomposto in $NB = 2$ blocchi
2. la funzione obiettivo f è pseudo-convessa ($\forall x, y \in \mathbb{R}^n, x \neq y$ si ha che $\nabla f(x)^T (y - x) \geq 0$ implica $f(y) \geq f(x)$) e insiemi di livello sono compatti

Tecniche di decomposizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si partizionano le variabili in blocchi e si minimizza rispetto al singolo blocco di variabili: se si sta risolvendo il problema $\min_x f(x)$ si partiziona il vettore x in NB blocchi x_1, \dots, x_{NB} e si definisce

$$x_i^{k+1} = \arg \min_y f(x_1^{k+1}, \dots, x_{i-1}^{k+1}, y, x_{i+1}^k, \dots, x_{NB}^k)$$

In generale, non converge, esiste un controesempio in 3 variabili per il metodo delle coordinate con ricerca di linea esatta.

Converge in uno dei seguenti casi:

1. Il vettore delle variabili è decomposto in $NB = 2$ blocchi
2. la funzione obiettivo f è pseudo-convessa ($\forall x, y \in \mathbb{R}^n, x \neq y$ si ha che $\nabla f(x)^T (y - x) \geq 0$ implica $f(y) \geq f(x)$) e insiemi di livello sono compatti
3. la funzione obiettivo f è strettamente convessa rispetto a $NB - 2$ blocchi.

Tecniche di decomposizione

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Si partizionano le variabili in blocchi e si minimizza rispetto al singolo blocco di variabili: se si sta risolvendo il problema $\min_x f(x)$ si partiziona il vettore x in NB blocchi x_1, \dots, x_{NB} e si definisce

$$x_i^{k+1} = \arg \min_y f(x_1^{k+1}, \dots, x_{i-1}^{k+1}, y, x_{i+1}^k, \dots, x_{NB}^k)$$

In generale, non converge, esiste un controesempio in 3 variabili per il metodo delle coordinate con ricerca di linea esatta.

Converge in uno dei seguenti casi:

1. Il vettore delle variabili è decomposto in $NB = 2$ blocchi
2. la funzione obiettivo f è pseudo-convessa ($\forall x, y \in \mathbb{R}^n, x \neq y$ si ha che $\nabla f(x)^T (y - x) \geq 0$ implica $f(y) \geq f(x)$) e insiemi di livello sono compatti
3. la funzione obiettivo f è strettamente convessa rispetto a $NB - 2$ blocchi.

In alternativa esistono metodi di discesa a blocchi in cui non è richiesta la minimizzazione globale rispetto a ogni blocco di variabili.

Decomposizione per RBF

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- **Decomposizione per RBF**
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Un primo schema di decomposizione si ottiene decomponendo il problema

$$\min_{w,c} E(w, c) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho_1}{2} \|w\|^2 + \frac{\rho_2}{2} \|c\|^2,$$

rispetto a due blocchi: pesi e centri. (nota: gli insiemi di livello sono compatti)

Decomposizione per RBF

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- **Decomposizione per RBF**
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Un primo schema di decomposizione si ottiene decomponendo il problema

$$\min_{w,c} E(w, c) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho_1}{2} \|w\|^2 + \frac{\rho_2}{2} \|c\|^2,$$

rispetto a due blocchi: pesi e centri. (nota: gli insiemi di livello sono compatti)

Si iterano due fasi:

Decomposizione per RBF

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- **Decomposizione per RBF**
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Un primo schema di decomposizione si ottiene decomponendo il problema

$$\min_{w,c} E(w, c) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho_1}{2} \|w\|^2 + \frac{\rho_2}{2} \|c\|^2,$$

rispetto a due blocchi: pesi e centri. (nota: gli insiemi di livello sono compatti)

Si iterano due fasi:

1. w_{k+1} risolve il problema dei minimi quadrati lineari in corrispondenza di c_k

Decomposizione per RBF

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- **Decomposizione per RBF**
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Un primo schema di decomposizione si ottiene decomponendo il problema

$$\min_{w,c} E(w, c) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho_1}{2} \|w\|^2 + \frac{\rho_2}{2} \|c\|^2,$$

rispetto a due blocchi: pesi e centri. (nota: gli insiemi di livello sono compatti)

Si iterano due fasi:

1. w_{k+1} risolve il problema dei minimi quadrati lineari in corrispondenza di c_k
2. c_{k+1} viene individuato tramite algoritmo di discesa con ricerca unidimensionale (Armijo) lungo l'antigradiente

Decomposizione per RBF

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- **Decomposizione per RBF**
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Un primo schema di decomposizione si ottiene decomponendo il problema

$$\min_{w,c} E(w, c) = \frac{1}{2} \sum_{p=1}^P \left(\sum_{i=1}^N w_i \phi(\|x^p - c_i\|) - y_p \right)^2 + \frac{\rho_1}{2} \|w\|^2 + \frac{\rho_2}{2} \|c\|^2,$$

rispetto a due blocchi: pesi e centri. (nota: gli insiemi di livello sono compatti)

Si iterano due fasi:

1. w_{k+1} risolve il problema dei minimi quadrati lineari in corrispondenza di c_k
2. c_{k+1} viene individuato tramite algoritmo di discesa con ricerca unidimensionale (Armijo) lungo l'antigradiente

Si dimostra la convergenza a punti stazionari e che i pesi di uscita corrispondono a un minimo globale nello spazio di w per la data scelta di c .

Decomposizione in 2 blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- **Decomposizione in 2 blocchi**
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $k = 0$.

Decomposizione in 2 blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- **Decomposizione in 2 blocchi**
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $k = 0$.

While $\nabla E(w_k, c_k) \neq 0$

Decomposizione in 2 blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- **Decomposizione in 2 blocchi**
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $k = 0$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Decomposizione in 2 blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- **Decomposizione in 2 blocchi**
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $k = 0$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

(a) calcola $d_k = -\nabla_c E(w_{k+1}, c_k)$

Decomposizione in 2 blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- **Decomposizione in 2 blocchi**
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $k = 0$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

(a) calcola $d_k = -\nabla_c E(w_{k+1}, c_k)$

(b) calcola η_k con una ricerca unidimensionale opportuna (Armijo)

Decomposizione in 2 blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- **Decomposizione in 2 blocchi**
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $k = 0$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

(a) calcola $d_k = -\nabla_c E(w_{k+1}, c_k)$

(b) calcola η_k con una ricerca unidimensionale opportuna (Armijo)

(c) scegli c_{k+1} tale che $E(w_{k+1}, c_{k+1}) \leq E(w_{k+1}, c_k + \eta_k d_k)$

Decomposizione in 2 blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- **Decomposizione in 2 blocchi**
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $k = 0$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

(a) calcola $d_k = -\nabla_c E(w_{k+1}, c_k)$

(b) calcola η_k con una ricerca unidimensionale opportuna (Armijo)

(c) scegli c_{k+1} tale che $E(w_{k+1}, c_{k+1}) \leq E(w_{k+1}, c_k + \eta_k d_k)$

Passo 3 $k = k + 1$

Decomposizione in 2 blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- **Decomposizione in 2 blocchi**
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $k = 0$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

(a) calcola $d_k = -\nabla_c E(w_{k+1}, c_k)$

(b) calcola η_k con una ricerca unidimensionale opportuna (Armijo)

(c) scegli c_{k+1} tale che $E(w_{k+1}, c_{k+1}) \leq E(w_{k+1}, c_k + \eta_k d_k)$

Passo 3 $k = k + 1$

End

Decomposizione in 2 blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- **Decomposizione in 2 blocchi**
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $k = 0$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

(a) calcola $d_k = -\nabla_c E(w_{k+1}, c_k)$

(b) calcola η_k con una ricerca unidimensionale opportuna (Armijo)

(c) scegli c_{k+1} tale che $E(w_{k+1}, c_{k+1}) \leq E(w_{k+1}, c_k + \eta_k d_k)$

Passo 3 $k = k + 1$

End

La sequenza $\{(w_k, c_k)\}$ generata ha punti di accumulazione, $\{E(w_k, c_k)\}$ converge e ogni punto di accumulazione di $\{(w_k, c_k)\}$ è un punto stazionario.

Decomposizione in $N + 1$ blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- **Decomposizione in $N + 1$ blocchi**
- Decomposizione in $N + 1$ blocchi
- Proprietà

Per problemi di dimensioni molto elevate, la minimizzazione approssimata rispetto a c può essere troppo costosa.

Decomposizione in $N + 1$ blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- **Decomposizione in $N + 1$ blocchi**
- Decomposizione in $N + 1$ blocchi
- Proprietà

Per problemi di dimensioni molto elevate, la minimizzazione approssimata rispetto a c può essere troppo costosa.

Si decompone ulteriormente il passo 2 in una successione di minimizzazioni sequenziali rispetto ai singoli centri

Decomposizione in $N + 1$ blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- **Decomposizione in $N + 1$ blocchi**
- Decomposizione in $N + 1$ blocchi
- Proprietà

Per problemi di dimensioni molto elevate, la minimizzazione approssimata rispetto a c può essere troppo costosa.

Si decompone ulteriormente il passo 2 in una successione di minimizzazioni sequenziali rispetto ai singoli centri

Le variabili vengono partizionate in $NB = N + 1$ blocchi costituiti dal vettore dei pesi w e dai vettori dei centri $c_i, i = 1, \dots, N$.

Decomposizione in $N + 1$ blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- **Decomposizione in $N + 1$ blocchi**
- Decomposizione in $N + 1$ blocchi
- Proprietà

Per problemi di dimensioni molto elevate, la minimizzazione approssimata rispetto a c può essere troppo costosa.

Si decompone ulteriormente il passo 2 in una successione di minimizzazioni sequenziali rispetto ai singoli centri

Le variabili vengono partizionate in $NB = N + 1$ blocchi costituiti dal vettore dei pesi w e dai vettori dei centri $c_i, i = 1, \dots, N$.

Come nell'algoritmo precedente i pesi vengono aggiornati risolvendo il problema dei minimi quadrati lineare, mentre per i centri si possono usare metodi di discesa per minimizzare rispetto ai centri.

Decomposizione in $N + 1$ blocchi

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- **Decomposizione in $N + 1$ blocchi**
- Decomposizione in $N + 1$ blocchi
- Proprietà

Per problemi di dimensioni molto elevate, la minimizzazione approssimata rispetto a c può essere troppo costosa.

Si decompone ulteriormente il passo 2 in una successione di minimizzazioni sequenziali rispetto ai singoli centri

Le variabili vengono partizionate in $NB = N + 1$ blocchi costituiti dal vettore dei pesi w e dai vettori dei centri $c_i, i = 1, \dots, N$.

Come nell'algoritmo precedente i pesi vengono aggiornati risolvendo il problema dei minimi quadrati lineare, mentre per i centri si possono usare metodi di discesa per minimizzare rispetto ai centri.

Per garantire la convergenza la linesearch deve soddisfare ipotesi più forti, in particolare va bene Armijo combinata con una limitazione superiore del passo.

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0, k = 0$, $\xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- **Decomposizione in $N + 1$ blocchi**
- Proprietà

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0, k = 0, \xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- **Decomposizione in $N + 1$ blocchi**
- Proprietà

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0, k = 0, \xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- **Decomposizione in $N + 1$ blocchi**
- Proprietà

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0, k = 0, \xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0, k = 0, \xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

For $i = 1, \dots, N$

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0, k = 0, \xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

For $i = 1, \dots, N$

If $\|\nabla_{c_i} E(w_{k+1}, c_k)\| \leq \xi_i^k$ **then** poni $c_i^{k+1} = c_i^k$

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0$, $k = 0$, $\xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

For $i = 1, \dots, N$

If $\|\nabla_{c_i} E(w_{k+1}, c_k)\| \leq \xi_i^k$ **then** poni $c_i^{k+1} = c_i^k$

else

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- **Decomposizione in $N + 1$ blocchi**
- Proprietà

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0, k = 0, \xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

For $i = 1, \dots, N$

If $\|\nabla_{c_i} E(w_{k+1}, c_k)\| \leq \xi_i^k$ **then** poni $c_i^{k+1} = c_i^k$

else

(a) calcola $d_i^k = -\nabla_{c_i} E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k, \dots, c_N^k)$

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0, k = 0, \xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

For $i = 1, \dots, N$

If $\|\nabla_{c_i} E(w_{k+1}, c_k)\| \leq \xi_i^k$ **then** poni $c_i^{k+1} = c_i^k$

else

(a) calcola $d_i^k = -\nabla_{c_i} E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k, \dots, c_N^k)$

(b) calcola η_i^k con una ricerca unidimensionale opportuna (Armijo con limitazione superiore sul passo)

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- **Decomposizione in $N + 1$ blocchi**
- Proprietà

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0, k = 0, \xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

For $i = 1, \dots, N$

If $\|\nabla_{c_i} E(w_{k+1}, c_k)\| \leq \xi_i^k$ **then** poni $c_i^{k+1} = c_i^k$

else

(a) calcola $d_i^k = -\nabla_{c_i} E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k, \dots, c_N^k)$

(b) calcola η_i^k con una ricerca unidimensionale opportuna (Armijo con limitazione superiore sul passo)

(c) scegli \tilde{c}_i^k tale che $E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, \tilde{c}_i^k, \dots, c_N^k) \leq E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k + \eta_i^k d_i^k, \dots, c_N^k)$

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0$, $k = 0$, $\xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

For $i = 1, \dots, N$

If $\|\nabla_{c_i} E(w_{k+1}, c_k)\| \leq \xi_i^k$ **then** poni $c_i^{k+1} = c_i^k$

else

(a) calcola $d_i^k = -\nabla_{c_i} E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k, \dots, c_N^k)$

(b) calcola η_i^k con una ricerca unidimensionale opportuna (Armijo con limitazione superiore sul passo)

(c) scegli \tilde{c}_i^k tale che $E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, \tilde{c}_i^k, \dots, c_N^k) \leq$

$$E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k + \eta_i^k d_i^k, \dots, c_N^k)$$

If $E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, \tilde{c}_i^k, \dots, c_N^k) \leq$

$$E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k, \dots, c_N^k) - \tau_i \|\tilde{c}_i^k - c_i^k\|^2$$
 then poni $c_i^{k+1} = \tilde{c}_i^k$

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0$, $k = 0$, $\xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

For $i = 1, \dots, N$

If $\|\nabla_{c_i} E(w_{k+1}, c_k)\| \leq \xi_i^k$ **then** poni $c_i^{k+1} = c_i^k$

else

(a) calcola $d_i^k = -\nabla_{c_i} E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k, \dots, c_N^k)$

(b) calcola η_i^k con una ricerca unidimensionale opportuna (Armijo con limitazione superiore sul passo)

(c) scegli \tilde{c}_i^k tale che $E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, \tilde{c}_i^k, \dots, c_N^k) \leq$

$$E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k + \eta_i^k d_i^k, \dots, c_N^k)$$

If $E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, \tilde{c}_i^k, \dots, c_N^k) \leq$

$$E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k, \dots, c_N^k) - \tau_i \|\tilde{c}_i^k - c_i^k\|^2$$

then poni $c_i^{k+1} = \tilde{c}_i^k$

else poni $c_i^{k+1} = c_i^k + \eta_i^k d_i^k$

Decomposizione in $N + 1$ blocchi

Dati. $c_0 \in \mathbb{R}^{nN}$, $w_0 \in \mathbb{R}^N$, $\tau_i > 0$, $k = 0$, $\xi_i^k > 0$ tali che $\xi_i^k \rightarrow 0$ per $i = 1, \dots, N$.

While $\nabla E(w_k, c_k) \neq 0$

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

Passo 1 Minimizzazione rispetto ai pesi:

calcola $w_{k+1} = \arg \min_w E(w, c_k)$ risolvendo il problema ai minimi quadrati in w .

Passo 2 Minimizzazione rispetto ai centri:

For $i = 1, \dots, N$

If $\|\nabla_{c_i} E(w_{k+1}, c_k)\| \leq \xi_i^k$ **then** poni $c_i^{k+1} = c_i^k$

else

(a) calcola $d_i^k = -\nabla_{c_i} E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k, \dots, c_N^k)$

(b) calcola η_i^k con una ricerca unidimensionale opportuna (Armijo con limitazione superiore sul passo)

(c) scegli \tilde{c}_i^k tale che $E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, \tilde{c}_i^k, \dots, c_N^k) \leq$

$$E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k + \eta_i^k d_i^k, \dots, c_N^k)$$

If $E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, \tilde{c}_i^k, \dots, c_N^k) \leq$

$$E(w_{k+1}, c_1^{k+1}, \dots, c_{i-1}^{k+1}, c_i^k, \dots, c_N^k) - \tau_i \|\tilde{c}_i^k - c_i^k\|^2$$

then poni $c_i^{k+1} = \tilde{c}_i^k$

else poni $c_i^{k+1} = c_i^k + \eta_i^k d_i^k$

Passo 3 $k = k + 1$

Proprietà

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

- ◁ Non si possono più effettuare le minimizzazioni parziali con qualunque algoritmo di discesa, poichè i punti tentativo \tilde{c}_i^k possono non essere accettati se producono spostamenti troppo grandi dal punto c_i^k .

Proprietà

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete
multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento
- Scelta non supervisionata dei centri
- Scelta supervisionata
- Tecniche di decomposizione
- Decomposizione per RBF
- Decomposizione in 2 blocchi
- Decomposizione in $N + 1$ blocchi
- Decomposizione in $N + 1$ blocchi
- Proprietà

- ◁ Non si possono più effettuare le minimizzazioni parziali con qualunque algoritmo di discesa, poichè i punti tentativo \tilde{c}_i^k possono non essere accettati se producono spostamenti troppo grandi dal punto c_i^k .
- ◁ Il test al passo 2 fa sì che la minimizzazione parziale rispetto al centro c_i è evitata quando il corrispondente gradiente parziale è piccolo, risparmiando in termini di tempo

Proprietà

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento

- Scelta non

supervisionata dei centri

- Scelta supervisionata

- Tecniche di

decomposizione

- Decomposizione per

RBF

- Decomposizione in 2

blocchi

- Decomposizione in

$N + 1$ blocchi

- Decomposizione in

$N + 1$ blocchi

- Proprietà

- ◁ Non si possono più effettuare le minimizzazioni parziali con qualunque algoritmo di discesa, poichè i punti tentativo \tilde{c}_i^k possono non essere accettati se producono spostamenti troppo grandi dal punto c_i^k .
- ◁ Il test al passo 2 fa sì che la minimizzazione parziale rispetto al centro c_i è evitata quando il corrispondente gradiente parziale è piccolo, risparmiando in termini di tempo
- ◁ Il calcolo della funzione di errore in corrispondenza di un nuovo centro c_i può essere fatto in modo molto efficiente, memorizzando le uscite della rete corrispondenti a c_i^k . Infatti

$$y(x^p; w_{k+1}, c_1^{k+1}, \dots, c_i, \dots, c_N^k) = y(x^p; w_{k+1}, c_1^{k+1}, \dots, c_i^k, \dots, c_N^k) - w_i^k \phi(\|x^p - c_i^k\|) + w_i^k \phi(\|x^p - c_i\|)$$

Proprietà

Introduzione

Data mining

Problemi di DM

Reti neurali

Un po' di storia

Perceptron

Reti multistrato

Creazione di una rete multistrato

Metodi on-line

Reti RBF

Addestramento RBF

- Addestramento

- Scelta non

supervisionata dei centri

- Scelta supervisionata

- Tecniche di decomposizione

- Decomposizione per RBF

- Decomposizione in 2 blocchi

- Decomposizione in $N + 1$ blocchi

- Decomposizione in $N + 1$ blocchi

- Proprietà

- ◁ Non si possono più effettuare le minimizzazioni parziali con qualunque algoritmo di discesa, poichè i punti tentativo \tilde{c}_i^k possono non essere accettati se producono spostamenti troppo grandi dal punto c_i^k .
- ◁ Il test al passo 2 fa sì che la minimizzazione parziale rispetto al centro c_i è evitata quando il corrispondente gradiente parziale è piccolo, risparmiando in termini di tempo
- ◁ Il calcolo della funzione di errore in corrispondenza di un nuovo centro c_i può essere fatto in modo molto efficiente, memorizzando le uscite della rete corrispondenti a c_i^k . Infatti

$$y(x^p; w_{k+1}, c_1^{k+1}, \dots, c_i, \dots, c_N^k) = y(x^p; w_{k+1}, c_1^{k+1}, \dots, c_i^k, \dots, c_N^k) - w_i^k \phi(\|x^p - c_i^k\|) + w_i^k \phi(\|x^p - c_i\|)$$

- ◁ si hanno proprietà di convergenza analoghe a quelle dell'algoritmo a due blocchi