
Introduzione alle Reti Neurali Artificiali*

Corso di MOBD

Roma 23 Ottobre 2017

*Grippo, L., and M. Sciandrone. *Metodi di ottimizzazione per le reti neurali*. Rapporto Tecnico (2003): 09-03.

Contesto

Introduzione

● Contesto

● Definizione

● Caratteristiche

● Interesse Big Data

● DM oggi

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

Ogni giorno moltissime persone utilizzano il web per comunicare, scambiare messaggi e accedere ad informazioni, riversando, più o meno consapevolmente, un'enorme quantità di dati nel web.

Anche la produzione di dati all'interno delle aziende è aumentata considerevolmente, tanto che, oramai, le dimensioni dei database devono essere misurate almeno in terabyte se non in petabyte o exabyte. Questa crescita esponenziale dei dati disponibili in termini di volume, di varietà e di velocità è sicuramente un patrimonio di inestimabile valore.

Definizione

Introduzione

- Contesto
- **Definizione**
- Caratteristiche
- Interesse Big Data
- DM oggi

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

Big Data (Wikipedia)

Si parla di **Big Data** quando si ha un dataset talmente grande da richiedere strumenti non convenzionali per estrapolare, gestire e processare informazioni entro un tempo ragionevole. Non esiste una dimensione di riferimento, ma questa cambia sempre, poichè le macchine sono sempre più veloci e i dataset sono sempre più grandi. Secondo uno studio del 2001, l'analista Doug Laney aveva definito il modello di crescita come tridimensionale (modello delle "3V"): con il passare del tempo aumentano volume (dei dati), velocità e varietà (dei dati). In molti casi questo modello è ancora valido, nonostante nel 2012 il modello sia stato esteso ad una quarta variabile, la veridicità .

Caratteristiche

Introduzione

- Contesto
- Definizione
- **Caratteristiche**
- Interesse Big Data
- DM oggi

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

1. **volume**: rappresenta la dimensione effettiva del dataset;
2. **velocità** : si riferisce alla velocità di generazione dei dati; si tende all'effettuare analisi dei dati in tempo reale o quasi;
3. **varietà** : riferita alle varie tipologie di dati, provenienti da fonti diverse (strutturate e non);

Caratteristiche

Introduzione

- Contesto
- Definizione
- **Caratteristiche**
- Interesse Big Data
- DM oggi

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

1. **volume**: rappresenta la dimensione effettiva del dataset;
2. **velocità**: si riferisce alla velocità di generazione dei dati; si tende all'effettuare analisi dei dati in tempo reale o quasi;
3. **varietà**: riferita alle varie tipologie di dati, provenienti da fonti diverse (strutturate e non);
4. **variabilità**: questa caratteristica può essere un problema; si riferisce alla possibilità di inconsistenza dei dati;
5. **complessità**: maggiore è la dimensione del dataset, maggiore è la complessità dei dati da gestire; il compito più difficile è collegare le informazioni, ed ottenerne di interessanti.

Caratteristiche

Introduzione

- Contesto
- Definizione
- **Caratteristiche**
- Interesse Big Data
- DM oggi

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

1. **volume**: rappresenta la dimensione effettiva del dataset;
2. **velocità** : si riferisce alla velocità di generazione dei dati; si tende all'effettuare analisi dei dati in tempo reale o quasi;
3. **varietà** : riferita alle varie tipologie di dati, provenienti da fonti diverse (strutturate e non);
4. **variabilità** : questa caratteristica può essere un problema; si riferisce alla possibilità di inconsistenza dei dati;
5. **complessità** : maggiore è la dimensione del dataset, maggiore è la complessità dei dati da gestire; il compito più difficile è collegare le informazioni, ed ottenerne di interessanti.
6. Alcune organizzazioni utilizzano una quarta V per indicare la **veridicità** dei dati, ossia la qualità dei dati intesa come il valore informativo che si riesce ad estrarre

Interesse Big Data

Introduzione

- Contesto
- Definizione
- Caratteristiche
- Interesse Big Data
- DM oggi

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

L'interesse è motivato da diversi fattori storici/sociali:

Interesse Big Data

Introduzione

- Contesto
- Definizione
- Caratteristiche
- Interesse Big Data
- DM oggi

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

L'interesse è motivato da diversi fattori storici/sociali:

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..)

Interesse Big Data

Introduzione

- Contesto
- Definizione
- Caratteristiche
- Interesse Big Data
- DM oggi

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

L'interesse è motivato da diversi fattori storici/sociali:

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

Interesse Big Data

Introduzione

- Contesto
- Definizione
- Caratteristiche
- Interesse Big Data
- DM oggi

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

L'interesse è motivato da diversi fattori storici/sociali:

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

- Contesto
- Definizione
- Caratteristiche
- Interesse Big Data
- **DM oggi**

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

Oggi il termine data mining ha una duplice valenza:

DM oggi

Introduzione

- Contesto
- Definizione
- Caratteristiche
- Interesse Big Data
- DM oggi

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

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

- Contesto
- Definizione
- Caratteristiche
- Interesse Big Data
- **DM oggi**

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

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

- Contesto
- Definizione
- Caratteristiche
- Interesse Big Data
- **DM oggi**

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

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).

DM oggi

Introduzione

- Contesto
- Definizione
- Caratteristiche
- Interesse Big Data
- **DM oggi**

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

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:

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

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

● **Tipo di apprendimento**

● Modelli predittivi e descrittivi

● Classificazione

● Formalmente

● Esempio 1: BCI

● Esempio 2 : Credit

Scoring

● Regressione

● Esempio

● Clustering

● Esempio 1

● Esempio 2

● Metodi per DM

Reti neurali

Un po' di storia

□ Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

● **Tipo di apprendimento**

● Modelli predittivi e descrittivi

● Classificazione

● Formalmente

● Esempio 1: BCI

● Esempio 2 : Credit

Scoring

● Regressione

● Esempio

● Clustering

● Esempio 1

● Esempio 2

● Metodi per DM

Reti neurali

Un po' di storia

□ Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:

◁ Support vector machine (SVM);

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

● **Tipo di apprendimento**

● Modelli predittivi e descrittivi

● Classificazione

● Formalmente

● Esempio 1: BCI

● Esempio 2 : Credit

Scoring

● Regressione

● Esempio

● Clustering

● Esempio 1

● Esempio 2

● Metodi per DM

Reti neurali

Un po' di storia

□ Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:

◁ Support vector machine (SVM);

◁ Reti neurali;

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

● **Tipo di apprendimento**

● Modelli predittivi e descrittivi

● Classificazione

● Formalmente

● Esempio 1: BCI

● Esempio 2 : Credit

Scoring

● Regressione

● Esempio

● Clustering

● Esempio 1

● Esempio 2

● Metodi per DM

Reti neurali

Un po' di storia

□ Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:

◁ Support vector machine (SVM);

◁ Reti neurali;

◁ Alberi decisionali con boosting e bagging;

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

- **Tipo di apprendimento**

- Modelli predittivi e descrittivi

- Classificazione

- Formalmente

- Esempio 1: BCI

- Esempio 2 : Credit

Scoring

- Regressione

- Esempio

- Clustering

- Esempio 1

- Esempio 2

- Metodi per DM

Reti neurali

Un po' di storia

- Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:
 - ◁ Support vector machine (SVM);
 - ◁ Reti neurali;
 - ◁ Alberi decisionali con boosting e bagging;
 - ◁ Classificatore Bayesiano Naive;

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

- **Tipo di apprendimento**

- Modelli predittivi e descrittivi

- Classificazione

- Formalmente

- Esempio 1: BCI

- Esempio 2 : Credit

Scoring

- Regressione

- Esempio

- Clustering

- Esempio 1

- Esempio 2

- Metodi per DM

Reti neurali

Un po' di storia

□ Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:

◁ Support vector machine (SVM);

◁ Reti neurali;

◁ Alberi decisionali con boosting e bagging;

◁ Classificatore Bayesiano Naive;

◁ etc.

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

● Tipo di apprendimento

● Modelli predittivi e descrittivi

● Classificazione

● Formalmente

● Esempio 1: BCI

● Esempio 2 : Credit

Scoring

● Regressione

● Esempio

● Clustering

● Esempio 1

● Esempio 2

● Metodi per DM

Reti neurali

Un po' di storia

- Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:
 - ◁ Support vector machine (SVM);
 - ◁ Reti neurali;
 - ◁ Alberi decisionali con boosting e bagging;
 - ◁ Classificatore Bayesiano Naive;
 - ◁ etc.

- Unsupervised Learning: cercano di individuare delle strutture (o informazioni) utili, all'interno di dati non classificati in precedenza.

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

● Tipo di apprendimento

● Modelli predittivi e descrittivi

● Classificazione

● Formalmente

● Esempio 1: BCI

● Esempio 2 : Credit

Scoring

● Regressione

● Esempio

● Clustering

● Esempio 1

● Esempio 2

● Metodi per DM

Reti neurali

Un po' di storia

- Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:
 - ◁ Support vector machine (SVM);
 - ◁ Reti neurali;
 - ◁ Alberi decisionali con boosting e bagging;
 - ◁ Classificatore Bayesiano Naive;
 - ◁ etc.
- Unsupervised Learning: cercano di individuare delle strutture (o informazioni) utili, all'interno di dati non classificati in precedenza.
 - ◁ k-mean e clustering gerarchico;

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento

- Modelli predittivi e descrittivi

- Classificazione

- Formalmente

- Esempio 1: BCI

- Esempio 2 : Credit

Scoring

- Regressione

- Esempio

- Clustering

- Esempio 1

- Esempio 2

- Metodi per DM

Reti neurali

Un po' di storia

- Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:
 - ◁ Support vector machine (SVM);
 - ◁ Reti neurali;
 - ◁ Alberi decisionali con boosting e bagging;
 - ◁ Classificatore Bayesiano Naive;
 - ◁ etc.
- Unsupervised Learning: cercano di individuare delle strutture (o informazioni) utili, all'interno di dati non classificati in precedenza.
 - ◁ k-mean e clustering gerarchico;
 - ◁ Catene di Markov nascoste;

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

● Tipo di apprendimento

● Modelli predittivi e descrittivi

● Classificazione

● Formalmente

● Esempio 1: BCI

● Esempio 2 : Credit

Scoring

● Regressione

● Esempio

● Clustering

● Esempio 1

● Esempio 2

● Metodi per DM

Reti neurali

Un po' di storia

- Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:
 - ◁ Support vector machine (SVM);
 - ◁ Reti neurali;
 - ◁ Alberi decisionali con boosting e bagging;
 - ◁ Classificatore Bayesiano Naive;
 - ◁ etc.

- Unsupervised Learning: cercano di individuare delle strutture (o informazioni) utili, all'interno di dati non classificati in precedenza.
 - ◁ k-mean e clustering gerarchico;
 - ◁ Catene di Markov nascoste;
 - ◁ Mappe auto-organizzanti (SOM);

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento

- Modelli predittivi e descrittivi

- Classificazione

- Formalmente

- Esempio 1: BCI

- Esempio 2 : Credit

Scoring

- Regressione

- Esempio

- Clustering

- Esempio 1

- Esempio 2

- Metodi per DM

Reti neurali

Un po' di storia

- Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:
 - ◁ Support vector machine (SVM);
 - ◁ Reti neurali;
 - ◁ Alberi decisionali con boosting e bagging;
 - ◁ Classificatore Bayesiano Naive;
 - ◁ etc.
- Unsupervised Learning: cercano di individuare delle strutture (o informazioni) utili, all'interno di dati non classificati in precedenza.
 - ◁ k-mean e clustering gerarchico;
 - ◁ Catene di Markov nascoste;
 - ◁ Mappe auto-organizzanti (SOM);
 - ◁ Reti neurali;

Tipo di apprendimento

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento

- Modelli predittivi e descrittivi

- Classificazione

- Formalmente

- Esempio 1: BCI

- Esempio 2 : Credit

Scoring

- Regressione

- Esempio

- Clustering

- Esempio 1

- Esempio 2

- Metodi per DM

Reti neurali

Un po' di storia

- Supervised Learning: sviluppano delle regole di classificazione, utilizzando per l'addestramento dei dati già classificati e i più utilizzati sono:
 - ◁ Support vector machine (SVM);
 - ◁ Reti neurali;
 - ◁ Alberi decisionali con boosting e bagging;
 - ◁ Classificatore Bayesiano Naive;
 - ◁ etc.
- Unsupervised Learning: cercano di individuare delle strutture (o informazioni) utili, all'interno di dati non classificati in precedenza.
 - ◁ k-mean e clustering gerarchico;
 - ◁ Catene di Markov nascoste;
 - ◁ Mappe auto-organizzanti (SOM);
 - ◁ Reti neurali;
 - ◁ etc.

Modelli predittivi e descrittivi

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento

● Modelli predittivi e descrittivi

- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

I modelli **predittivi** utilizzano dati che rappresentano fatti o eventi, per i quali sono noti a priori l'evoluzione finale o i risultati. Tali informazioni vengono poi sfruttate per costruire un modello in grado di predire, esplicitamente, l'evoluzione e i risultati di altri fenomeni di interesse.

Modelli predittivi e descrittivi

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento

- **Modelli predittivi e descrittivi**

- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit Scoring

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

I modelli **predittivi** utilizzano dati che rappresentano fatti o eventi, per i quali sono noti a priori l'evoluzione finale o i risultati. Tali informazioni vengono poi sfruttate per costruire un modello in grado di predire, esplicitamente, l'evoluzione e i risultati di altri fenomeni di interesse.

Nel caso dei modelli **descrittivi**, invece, lo scopo è quello di caratterizzare le proprietà generali dei dati analizzati.

Classificazione

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi

● **Classificazione**

- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

La **classificazione** è un task prettamente predittivo, lo scopo è quello di individuare l'appartenenza di un dato pattern ad una certa classe.

Classificazione

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi

● **Classificazione**

- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

La **classificazione** è un task prettamente predittivo, lo scopo è quello di individuare l'appartenenza di un dato pattern ad una certa classe.

Sono noti a priori dei pattern rappresentativi delle diverse classi, sulla base dei quali si costruisce un modello matematico che, dato un generico elemento appartenente allo spazio delle caratteristiche, definisce la corrispondente classe di appartenenza.

Classificazione

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi

● Classificazione

- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

La **classificazione** è un task prettamente predittivo, lo scopo è quello di individuare l'appartenenza di un dato pattern ad una certa classe.

Sono noti a priori dei pattern rappresentativi delle diverse classi, sulla base dei quali si costruisce un modello matematico che, dato un generico elemento appartenente allo spazio delle caratteristiche, definisce la corrispondente classe di appartenenza.

I problemi di classificazione si possono dividere in due sottocategorie:

Classificazione

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi

● Classificazione

- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

La **classificazione** è un task prettamente predittivo, lo scopo è quello di individuare l'appartenenza di un dato pattern ad una certa classe.

Sono noti a priori dei pattern rappresentativi delle diverse classi, sulla base dei quali si costruisce un modello matematico che, dato un generico elemento appartenente allo spazio delle caratteristiche, definisce la corrispondente classe di appartenenza.

I problemi di classificazione si possono dividere in due sottocategorie:

- ◁ si parla di **classificazione binaria** quando il target può assumere solamente due valori;

Classificazione

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi

● Classificazione

- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

La **classificazione** è un task prettamente predittivo, lo scopo è quello di individuare l'appartenenza di un dato pattern ad una certa classe.

Sono noti a priori dei pattern rappresentativi delle diverse classi, sulla base dei quali si costruisce un modello matematico che, dato un generico elemento appartenente allo spazio delle caratteristiche, definisce la corrispondente classe di appartenenza.

I problemi di classificazione si possono dividere in due sottocategorie:

- ◁ si parla di **classificazione binaria** quando il target può assumere solamente due valori;
- ◁ si parla di **classificazione multiclasse** quando il target può assumere un numero finito di valori interi numerici.

In entrambi i casi l'esito della classificazione è comunque categorico.

Formalmente

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- **Formalmente**
- Esempio 1: BCI
- Esempio 2 : Credit Scoring
- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

□ INPUT: Dato il Training set

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

dove P rappresenta il numero di **pattern** disponibili, ognuna descritta da n attributi. Per ogni osservazione si ha:

Formalmente

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- **Formalmente**
- Esempio 1: BCI
- Esempio 2 : Credit Scoring
- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

□ INPUT: Dato il Training set

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

dove P rappresenta il numero di **pattern** disponibili, ognuna descritta da n attributi. Per ogni osservazione si ha:

◁ un vettore x^i , con $1 \leq i \leq P$ di dimensione n : il generico elemento x_j^i , con $1 \leq j \leq n$, rappresenta il valore della j -esima feature nell'osservazione x^i ;

Formalmente

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- **Formalmente**
- Esempio 1: BCI
- Esempio 2 : Credit Scoring
- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

□ INPUT: Dato il Training set

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

dove P rappresenta il numero di **pattern** disponibili, ognuna descritta da n attributi. Per ogni osservazione si ha:

- ◁ un vettore x^i , con $1 \leq i \leq P$ di dimensione n : il generico elemento x_j^i , con $1 \leq j \leq n$, rappresenta il valore della j -esima feature nell'osservazione x^i ;
- ◁ un vettore y^i , con $1 \leq i \leq P$, che rappresenta la classe dell' i -esima osservazione.

Formalmente

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- **Formalmente**
- Esempio 1: BCI
- Esempio 2 : Credit Scoring
- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

□ INPUT: Dato il Training set

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

dove P rappresenta il numero di **pattern** disponibili, ognuna descritta da n attributi. Per ogni osservazione si ha:

- ◁ un vettore x^i , con $1 \leq i \leq P$ di dimensione n : il generico elemento x_j^i , con $1 \leq j \leq n$, rappresenta il valore della j -esima feature nell'osservazione x^i ;
- ◁ un vettore y^i , con $1 \leq i \leq P$, che rappresenta la classe dell' i -esima osservazione.

□ OUTPUT: Classification Model

$$f : R^n \rightarrow S \subset N$$

in grado di approssimare la relazione $x^i \rightarrow y^i$ e quindi, dato un nuovo elemento x , di predire il corrispondente valore di y .

Esempio 1: BCI

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- **Esempio 1: BCI**
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

Pattern: tracciato elettroencefalografico in seguito a uno stimolo per riconoscere il potenziale evocato della P300

2 Classi: 0 se stimolo non target, 1 se stimolo target

Esempio 1: BCI

Introduzione

Gli algoritmi utilizzati

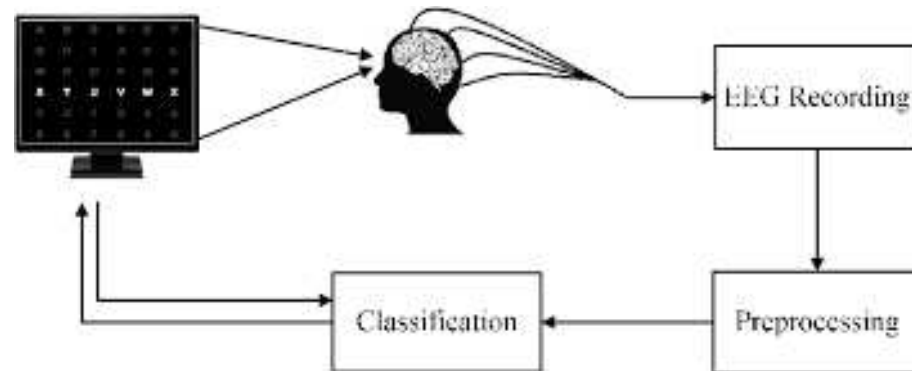
- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- **Esempio 1: BCI**
- Esempio 2 : Credit Scoring
- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

Pattern: tracciato elettroencefalografico in seguito a uno stimolo per riconoscere il potenziale evocato della P300

2 Classi: 0 se stimolo non target, 1 se stimolo target



Esempio 2 : Credit Scoring

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- **Esempio 2 : Credit Scoring**
- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

Pattern: Dati di bilancio negli ultimi anni, indici economici di mercato di una data azienda che vorrebbe essere finanziata

3 Classi: sicura, mediamente sicura, a rischio default

Regressione

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit Scoring

● Regressione

- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

Nei problemi di regressione, le coppie pattern/target sono anch'esse note a priori, come accade nella classificazione ma, in questo caso, sono rappresentative di una funzione incognita a valori reali.

Pattern:

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

Regressione

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit Scoring

● **Regressione**

- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

Nei problemi di regressione, le coppie pattern/target sono anch'esse note a priori, come accade nella classificazione ma, in questo caso, sono rappresentative di una funzione incognita a valori reali.

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

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit Scoring

● Regressione

- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

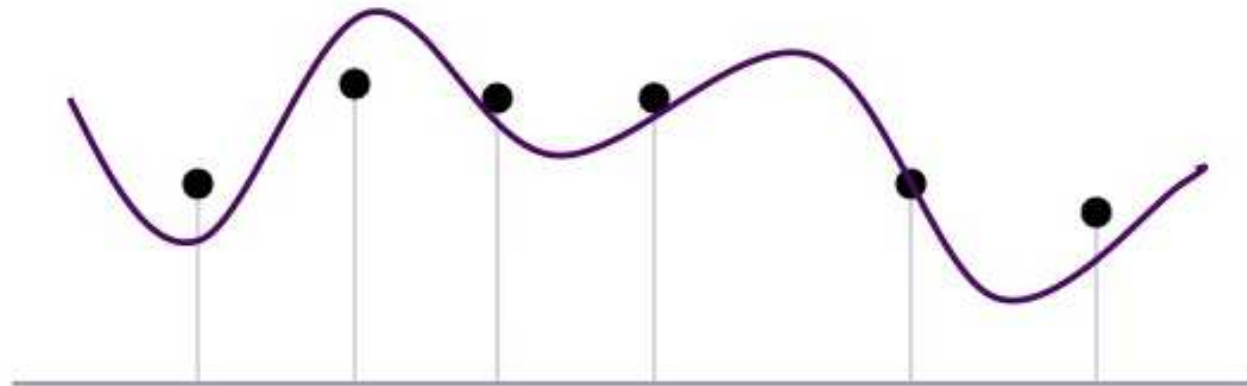
Un po' di storia

Nei problemi di regressione, le coppie pattern/target sono anch'esse note a priori, come accade nella classificazione ma, in questo caso, sono rappresentative di una funzione incognita a valori reali.

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

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- **Esempio**
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

ESEMPIO DI REGRESSIONE: elettrodeposizione di rame e grafene su lamine di alluminio

Il grafene è caratterizzato da inerzia chimica, elevata conducibilità elettrica, straordinaria flessibilità, alta trasparenza ottica e ottima conducibilità e stabilità termica.

Pattern: temperatura e corrente di deposizione, e concentrazione di grafene utilizzata in un dato esperimento

Dati disponibili: valore della diffusività termica del campione ottenuto

Obiettivo: determinare una funzione analitica che approssimi il legame funzionale tra la diffusività termica (da minimizzare) e i parametri che definiscono l'esperimento

Clustering

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

Il clustering è un task a carattere tipicamente descrittivo, attraverso il quale si cerca di identificare un numero finito di categorie (o cluster) per descrivere i dati. L'obiettivo è quello di identificare gruppi che siano il più possibile differenti l'uno dall'altro ma, nel contempo, è fondamentale che gli elementi di ogni categoria siano il più possibile simili tra loro.

Pattern:

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

Clustering

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit Scoring
- Regressione
- Esempio
- **Clustering**
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

Il clustering è un task a carattere tipicamente descrittivo, attraverso il quale si cerca di identificare un numero finito di categorie (o cluster) per descrivere i dati. L'obiettivo è quello di identificare gruppi che siano il più possibile differenti l'uno dall'altro ma, nel contempo, è fondamentale che gli elementi di ogni categoria siano il più possibile simili tra loro.

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

Clustering

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit Scoring
- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

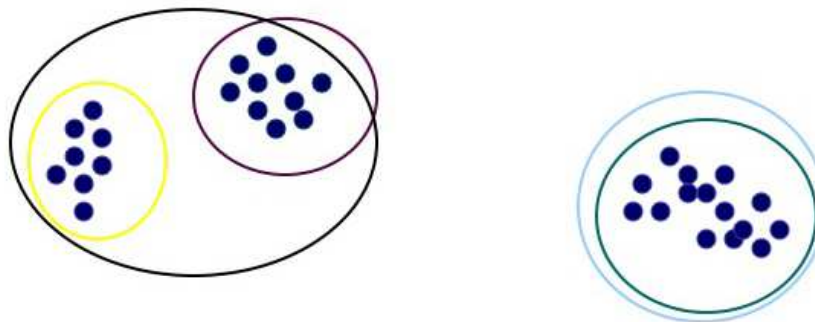
Un po' di storia

Il clustering è un task a carattere tipicamente descrittivo, attraverso il quale si cerca di identificare un numero finito di categorie (o cluster) per descrivere i dati. L'obiettivo è quello di identificare gruppi che siano il più possibile differenti l'uno dall'altro ma, nel contempo, è fondamentale che gli elementi di ogni categoria siano il più possibile simili tra loro.

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

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit Scoring
- Regressione
- Esempio
- Clustering
- **Esempio 1**
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

ESEMPIO

Pattern: comportamento di uno scommettitore in termini di scommesse e gestione del conto sul sito di scommesse

Dati disponibili: storico di scommesse e prelievi e depositi sul conto

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

Esempio 2

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit Scoring
- Regressione
- Esempio
- Clustering
- Esempio 1
- **Esempio 2**
- Metodi per DM

Reti neurali

Un po' di storia

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”

Metodi per DM

Introduzione

In questo corso vedremo:

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

Metodi per DM

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

In questo corso vedremo:

1. Clustering (classificazione non supervisionata)

Metodi per DM

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

In questo corso vedremo:

1. Clustering (classificazione non supervisionata)
2. Reti Neurali (regressione e classificazione supervisionata)

Metodi per DM

Introduzione

Gli algoritmi utilizzati

- Tipo di apprendimento
- Modelli predittivi e descrittivi
- Classificazione
- Formalmente
- Esempio 1: BCI
- Esempio 2 : Credit

Scoring

- Regressione
- Esempio
- Clustering
- Esempio 1
- Esempio 2
- Metodi per DM

Reti neurali

Un po' di storia

In questo corso vedremo:

1. Clustering (classificazione non supervisionata)
2. Reti Neurali (regressione e classificazione supervisionata)
3. Support Vector Machines (regressione e classificazione supervisionata)

Idea

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea

- Il neurone
- 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

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

Idea

Introduzione

Gli algoritmi utilizzati

Reti neurali

● Idea

- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

● Idea

- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

● Idea

- Il neurone
- 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

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

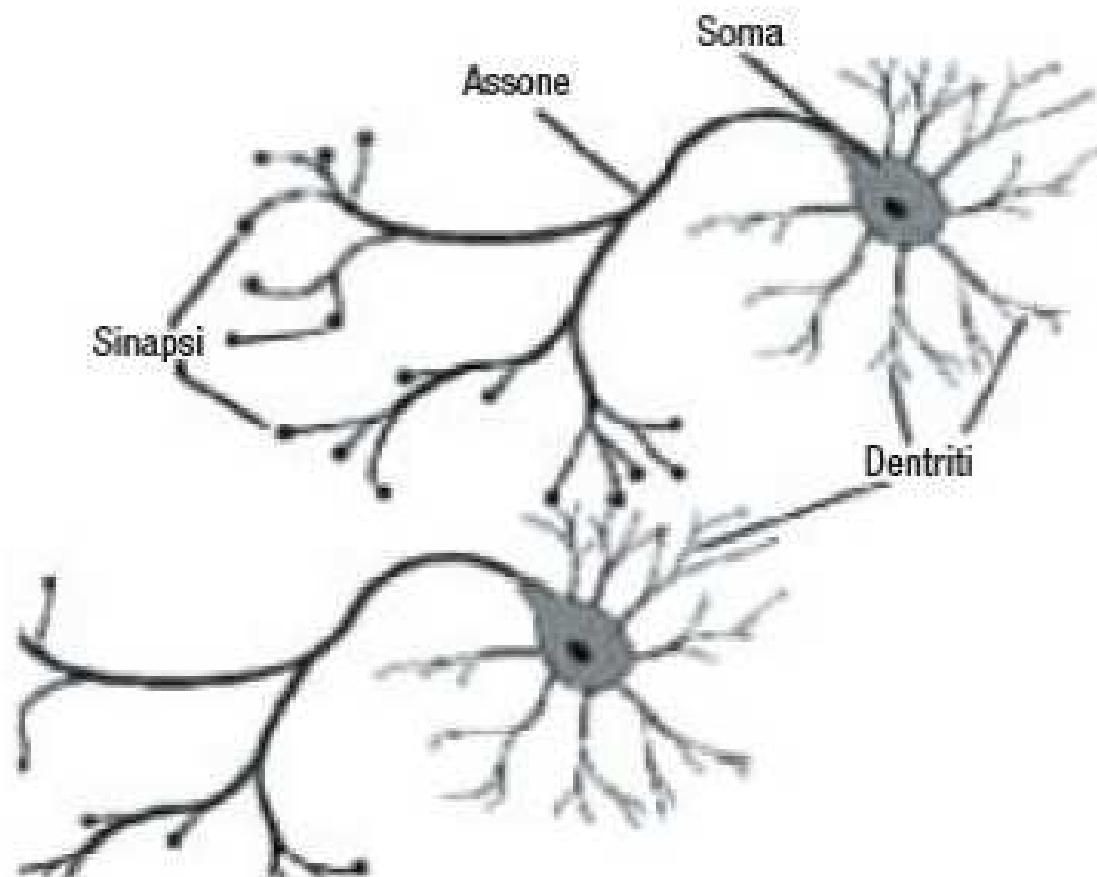
Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- **Il neurone**
- 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



Neurone e sua struttura cellulare

Rete neurale da un punto di vista fisico

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

Neuroni

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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**.

Neuroni

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

Il legame ingresso-uscita realizzato dalla rete dipende da:

Caratteristiche della rete

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Architetture

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

1. Perceptron

Architetture

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

1. Perceptron
2. Multilayer Perceptron (MLP)

Architetture

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

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

Architetture

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

□ 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

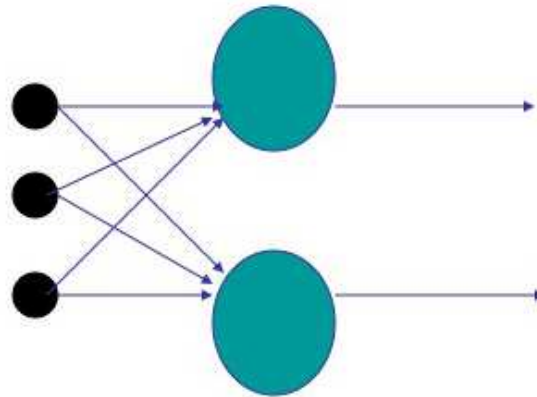
Introduzione

Gli algoritmi utilizzati

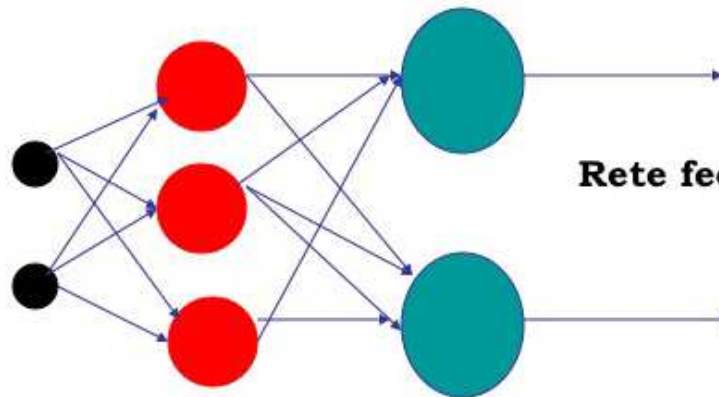
Reti neurali

- Idea
- Il neurone
- 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



Rete feedforward a 1 strato



Rete feedforward a 2 strati

Addestramento

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Addestramento

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

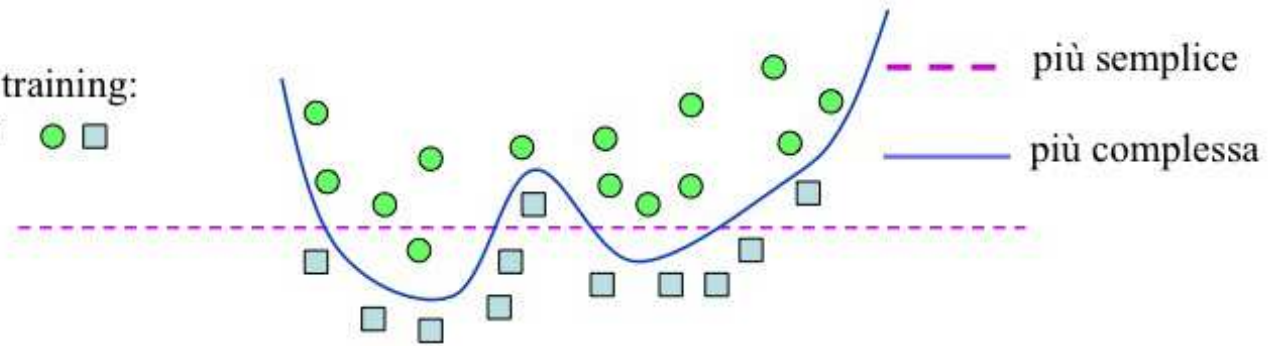
Gli algoritmi utilizzati

Reti neurali

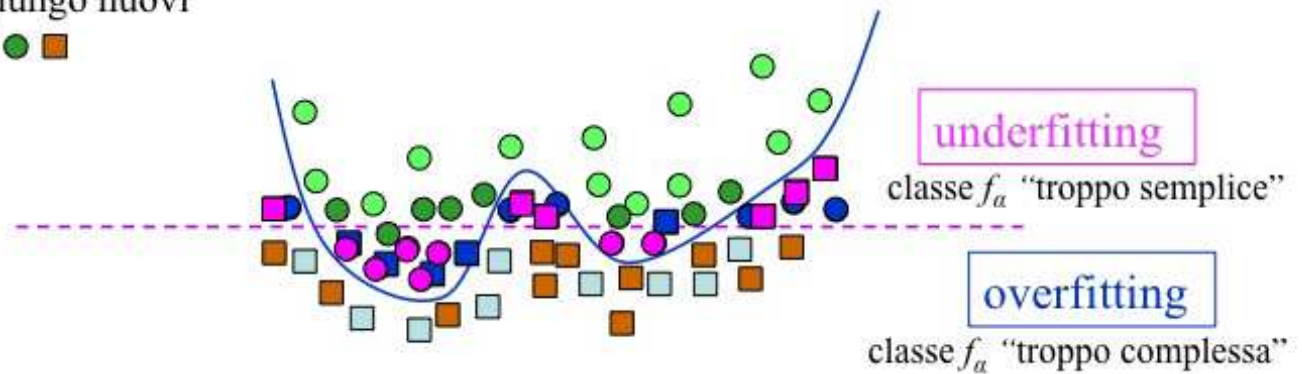
- Idea
- Il neurone
- Rete neurale da un punto di vista fisico
- Neuroni
- Rete neurale da un punto di vista matematico
- Caratteristiche della rete
- Architetture
- Esempi reti feedforward
- Addestramento
- Overfitting e underfitting
- Capacità di generalizzazione

Un po' di storia

Dati di training:
2 classi ● □



Aggiungo nuovi
dati ● ■



Capacità di generalizzazione

Introduzione

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

- Idea
- Il neurone
- 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

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

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

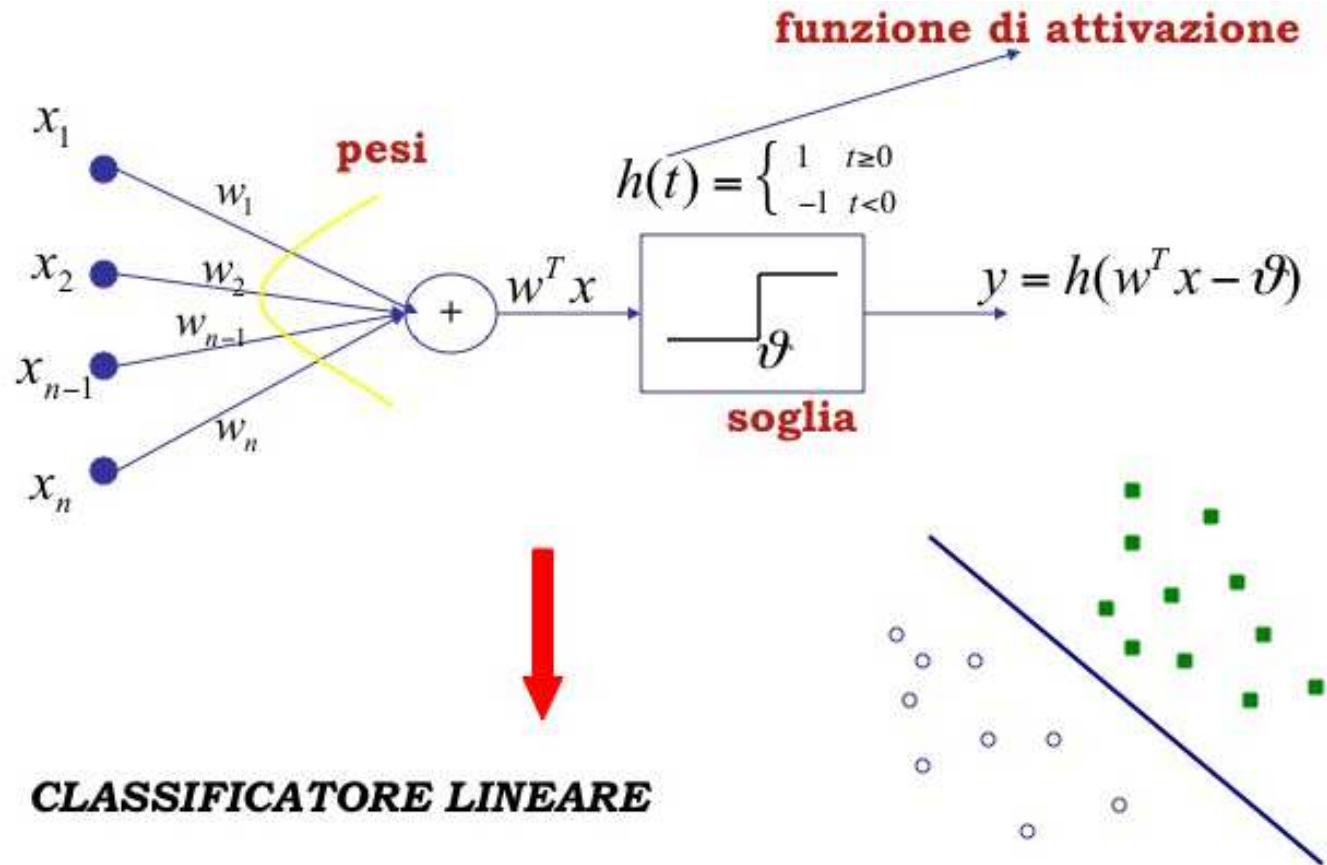
● Neurone formale

● Perceptron

● Limiti del perceptron

● La crisi delle reti neurali

● Backpropagation



Perceptron

Introduzione

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- **Perceptron**
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation

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

Perceptron

Introduzione

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- **Perceptron**
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation

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

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- **Perceptron**
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation

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

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- **Perceptron**
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation

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

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- **Perceptron**
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation

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

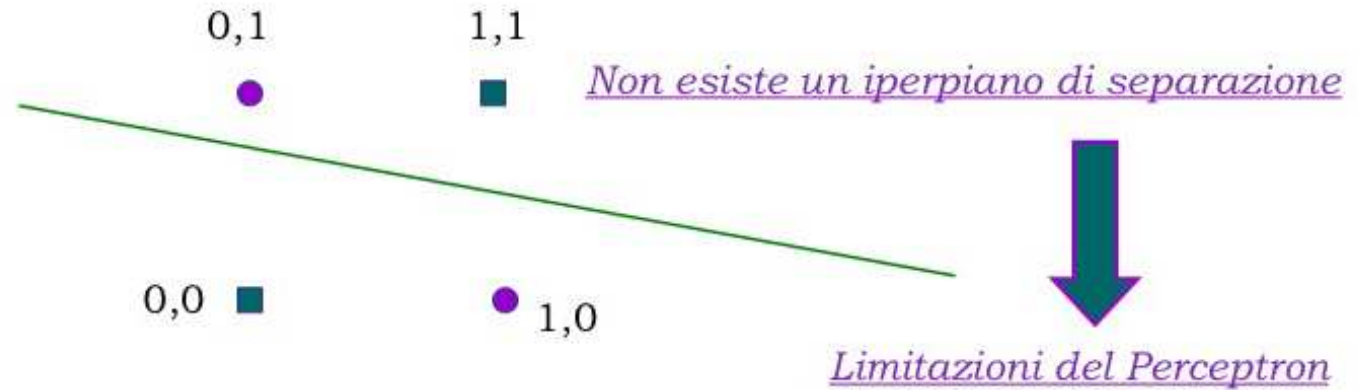
Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- **Limiti del perceptron**
- La crisi delle reti neurali
- Backpropagation

Problema XOR



La crisi delle reti neurali

Introduzione

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation

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

La crisi delle reti neurali

Introduzione

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation

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

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation

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

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- Backpropagation

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

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- **Backpropagation**

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

Gli algoritmi utilizzati

Reti neurali

Un po' di storia

- Neurone formale
- Perceptron
- Limiti del perceptron
- La crisi delle reti neurali
- **Backpropagation**

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