

Lezione introduttiva Algoritmi per il Web

Veronica Piccialli*

Roma 05 Marzo 2019

* Università degli Studi di Roma Tor Vergata



Vero titolo del corso

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- Materiale didattico
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Machine Learning



Vero titolo del corso

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- Materiale didattico
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Machine Learning

Dall'anno prossimo: MOBD 9 cfu, Machine Learning 9 cfu

Sito del corso

Introduzione al corso

- Vero titolo del corso
- **Sito del corso**
- Materiale didattico
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Sito web del docente:

<http://people.uniroma2.it/veronica.piccialli/index.html>

Sito del corso:

<http://people.uniroma2.it/veronica.piccialli/AW.html>

Mi aiuteranno l'Ing. Chiara Liti studentessa di dottorato (chiaraliti@gmail.com), e l'ing. Antonio Sudoso (antoniomaria.sudoso@gmail.com), e l'ing. Andrea Pomente (pomente.andrea@gmail.com) per la parte di deep learning, studenti di dottorato



Materiale didattico

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- **Materiale didattico**
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Testo: Luigi Grippo, Marco Sciandrone, Metodi di Ottimizzazione Non Vincolata, Springer, Unitext 2011



Materiale didattico

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- **Materiale didattico**
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Testo: Luigi Grippo, Marco Sciandrone, Metodi di Ottimizzazione Non Vincolata, Springer, Unitext 2011

Dispense scaricabili dal sito del corso



Materiale didattico

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- **Materiale didattico**
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Testo: Luigi Grippo, Marco Sciandrone, Metodi di Ottimizzazione Non Vincolata, Springer, Unitext 2011

Dispense scaricabili dal sito del corso

Testi in inglese forniti man mano, tra cui :

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- **Materiale didattico**
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Materiale didattico

Testo: Luigi Grippo, Marco Sciandrone, Metodi di Ottimizzazione Non Vincolata, Springer, Unitext 2011

Dispense scaricabili dal sito del corso

Testi in inglese forniti man mano, tra cui :

1. Introduction to Statistical Learning

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- **Materiale didattico**
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Materiale didattico

Testo: Luigi Grippo, Marco Sciandrone, Metodi di Ottimizzazione Non Vincolata, Springer, Unitext 2011

Dispense scaricabili dal sito del corso

Testi in inglese forniti man mano, tra cui :

1. Introduction to Statistical Learning
2. The elements of Statistical Learning

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- **Materiale didattico**
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Materiale didattico

Testo: Luigi Grippo, Marco Sciandrone, Metodi di Ottimizzazione Non Vincolata, Springer, Unitext 2011

Dispense scaricabili dal sito del corso

Testi in inglese forniti man mano, tra cui :

1. Introduction to Statistical Learning
2. The elements of Statistical Learning
3. Deep Learning

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- **Materiale didattico**
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Materiale didattico

Testo: Luigi Grippo, Marco Sciandrone, Metodi di Ottimizzazione Non Vincolata, Springer, Unitext 2011

Dispense scaricabili dal sito del corso

Testi in inglese forniti man mano, tra cui :

1. Introduction to Statistical Learning
2. The elements of Statistical Learning
3. Deep Learning
4. Understanding Machine Learning: from Theory to Algorithms

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- **Materiale didattico**
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Materiale didattico

Testo: Luigi Grippo, Marco Sciandrone, Metodi di Ottimizzazione Non Vincolata, Springer, Unitext 2011

Dispense scaricabili dal sito del corso

Testi in inglese forniti man mano, tra cui :

1. Introduction to Statistical Learning
2. The elements of Statistical Learning
3. Deep Learning
4. Understanding Machine Learning: from Theory to Algorithms

TensorFlow (Keras) su Python materiale distribuito tramite gruppo google del corso (AW18-19)

Materiale didattico

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- **Materiale didattico**
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

Testo: Luigi Grippo, Marco Sciandrone, Metodi di Ottimizzazione Non Vincolata, Springer, Unitext 2011

Dispense scaricabili dal sito del corso

Testi in inglese forniti man mano, tra cui :

1. Introduction to Statistical Learning
2. The elements of Statistical Learning
3. Deep Learning
4. Understanding Machine Learning: from Theory to Algorithms

TensorFlow (Keras) su Python materiale distribuito tramite gruppo google del corso (AW18-19)

Per iscriversi al gruppo, mandare un'email a veronica.piccialli@uniroma2.it

Modalità di esame

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- Materiale didattico
- **Modalità di esame**
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

orale 15 pt

progetto che prevede l'implementazione e addestramento di una rete neurale su dataset fornito da noi 15 pt

Il progetto può essere presentato da gruppi di **massimo 2 studenti** e **deve** essere consegnato entro l'anno accademico (analogamente se si consegna prima il progetto l'esame va fatto entro l'anno accademico).

Programma di Massima

Introduzione al corso

- Vero titolo del corso
- Sito del corso
- Materiale didattico
- Modalità di esame
- Programma di Massima

Esempi di applicazioni

Statistical Learning

Due approcci semplici

1. Statistical Learning (Apprendimento Statistico?)
2. Ottimizzazione Non Vincolata: esistenza, condizioni di ottimo, algoritmi
3. Reti Neurali: Perceptron, Multi Layer Perceptron, Reti RBF
4. Reti deep: Reti convoluzionali per classificazione di immagini

Machine Learning

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Il termine **Machine Learning** (o apprendimento automatico) si riferisce al rilevamento automatico di modelli significativi nei dati. E' diventato uno strumento comune in quasi tutte le attività che richiedono l'estrazione di informazioni da grandi insiemi di dati.

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Il termine **Machine Learning** (o apprendimento automatico) si riferisce al rilevamento automatico di modelli significativi nei dati. E' diventato uno strumento comune in quasi tutte le attività che richiedono l'estrazione di informazioni da grandi insiemi di dati.

1. I motori di ricerca imparano come darci i migliori risultati

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Il termine **Machine Learning** (o apprendimento automatico) si riferisce al rilevamento automatico di modelli significativi nei dati. E' diventato uno strumento comune in quasi tutte le attività che richiedono l'estrazione di informazioni da grandi insiemi di dati.

1. I motori di ricerca imparano come darci i migliori risultati
2. software anti-spam impara a filtrare i nostri messaggi e-mail

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Il termine **Machine Learning** (o apprendimento automatico) si riferisce al rilevamento automatico di modelli significativi nei dati. E' diventato uno strumento comune in quasi tutte le attività che richiedono l'estrazione di informazioni da grandi insiemi di dati.

1. I motori di ricerca imparano come darci i migliori risultati
2. software anti-spam impara a filtrare i nostri messaggi e-mail
3. le transazioni con carta di credito sono protette da un software che impara a rilevare le frodi

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Il termine **Machine Learning** (o apprendimento automatico) si riferisce al rilevamento automatico di modelli significativi nei dati. E' diventato uno strumento comune in quasi tutte le attività che richiedono l'estrazione di informazioni da grandi insiemi di dati.

1. I motori di ricerca imparano come darci i migliori risultati
2. software anti-spam impara a filtrare i nostri messaggi e-mail
3. le transazioni con carta di credito sono protette da un software che impara a rilevare le frodi
4. Le fotocamere digitali imparano a rilevare volti

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Il termine **Machine Learning** (o apprendimento automatico) si riferisce al rilevamento automatico di modelli significativi nei dati. E' diventato uno strumento comune in quasi tutte le attività che richiedono l'estrazione di informazioni da grandi insiemi di dati.

1. I motori di ricerca imparano come darci i migliori risultati
2. software anti-spam impara a filtrare i nostri messaggi e-mail
3. le transazioni con carta di credito sono protette da un software che impara a rilevare le frodi
4. Le fotocamere digitali imparano a rilevare volti
5. applicazioni intelligenti di assistenza personale su smart-phone imparano a riconoscere i comandi vocali

[Introduzione al corso](#)

[Esempi di applicazioni](#)

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

[Statistical Learning](#)

[Due approcci semplici](#)

Il termine **Machine Learning** (o apprendimento automatico) si riferisce al rilevamento automatico di modelli significativi nei dati. E' diventato uno strumento comune in quasi tutte le attività che richiedono l'estrazione di informazioni da grandi insiemi di dati.

1. I motori di ricerca imparano come darci i migliori risultati
2. software anti-spam impara a filtrare i nostri messaggi e-mail
3. le transazioni con carta di credito sono protette da un software che impara a rilevare le frodi
4. Le fotocamere digitali imparano a rilevare volti
5. applicazioni intelligenti di assistenza personale su smart-phone imparano a riconoscere i comandi vocali

In questi casi, a causa della complessità dei modelli che devono essere rilevati, un programmatore umano non può fornire una specificazione esplicita e dettagliata di come tali compiti dovrebbero essere eseguiti.

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Il termine **Machine Learning** (o apprendimento automatico) si riferisce al rilevamento automatico di modelli significativi nei dati. E' diventato uno strumento comune in quasi tutte le attività che richiedono l'estrazione di informazioni da grandi insiemi di dati.

1. I motori di ricerca imparano come darci i migliori risultati
2. software anti-spam impara a filtrare i nostri messaggi e-mail
3. le transazioni con carta di credito sono protette da un software che impara a rilevare le frodi
4. Le fotocamere digitali imparano a rilevare volti
5. applicazioni intelligenti di assistenza personale su smart-phone imparano a riconoscere i comandi vocali

In questi casi, a causa della complessità dei modelli che devono essere rilevati, un programmatore umano non può fornire una specificazione esplicita e dettagliata di come tali compiti dovrebbero essere eseguiti.

Gli strumenti di Machine Learning si occupano di dotare i programmi della capacità di "imparare".

Email Spam

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Si hanno a disposizione un certo numero di email classificate come email o spam, con la frequenza relativa di alcune parole chiave più comuni

Email Spam

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Si hanno a disposizione un certo numero di email classificate come email o spam, con la frequenza relativa di alcune parole chiave più comuni

Si vuole costruire un **predittore** che sia in grado di classificare una nuova email come email o spam utilizzando questi dati per imparare

Email Spam

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Si hanno a disposizione un certo numero di email classificate come email o spam, con la frequenza relativa di alcune parole chiave più comuni

Si vuole costruire un **predittore** che sia in grado di classificare una nuova email come email o spam utilizzando questi dati per imparare

E' un problema di **classificazione** binaria (due classi) **supervisionato** in quanto per il mio training set ho a disposizione le etichette (label)

Cancro alla prostata

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- **Cancro alla prostata**
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Si vuole studiare la correlazione tra 97 misure cliniche (come il volume del tumore al tronco, il peso della prostata, età, etc) e la quantità di antigene prostatico specifico (PSA), che è un enzima prodotto dalla prostata che aumenta solo in casi patologici, come nel caso di una prostata ingrossata (ipertrofia prostatica benigna) o nel caso di un tumore prostatico.

Cancro alla prostata

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- **Cancro alla prostata**
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Si vuole studiare la correlazione tra 97 misure cliniche (come il volume del tumore al tronco, il peso della prostata, età, etc) e la quantità di antigene prostatico specifico (PSA), che è un enzima prodotto dalla prostata che aumenta solo in casi patologici, come nel caso di una prostata ingrossata (ipertrofia prostatica benigna) o nel caso di un tumore prostatico.

In questo caso si vuole costruire un predittore che date le 97 misure predice il logaritmo della PSA.

Cancro alla prostata

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- **Cancro alla prostata**
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Si vuole studiare la correlazione tra 97 misure cliniche (come il volume del tumore al tronco, il peso della prostata, età , etc) e la quantità di antigene prostatico specifico (PSA), che è un enzima prodotto dalla prostata che aumenta solo in casi patologici, come nel caso di una prostata ingrossata (ipertrofia prostatica benigna) o nel caso di un tumore prostatico.

In questo caso si vuole costruire un predittore che date le 97 misure predice il logaritmo della PSA.

Si tratta di un problema di **regressione** perchè il valore da predire assume valori continui

Riconoscimento di numeri scritti a mano

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

I dati provengono dai codici di avviamento postale scritti a mano su buste da lettera degli Stati Uniti. Ogni immagine rappresenta una cifra delle 5 del codice postale.

Riconoscimento di numeri scritti a mano

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

I dati provengono dai codici di avviamento postale scritti a mano su buste da lettera degli Stati Uniti. Ogni immagine rappresenta una cifra delle 5 del codice postale.

Le immagini sono 16×16 in scala di grigi a otto bit, con ogni pixel di intensità compresa tra 0 e 255. Sono state normalizzate per avere approssimativamente le stesse dimensioni e orientamento.



Riconoscimento di numeri scritti a mano

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

I dati provengono dai codici di avviamento postale scritti a mano su buste da lettera degli Stati Uniti. Ogni immagine rappresenta una cifra delle 5 del codice postale.

Le immagini sono 16×16 in scala di grigi a otto bit, con ogni pixel di intensità compresa tra 0 e 255. Sono state normalizzate per avere approssimativamente le stesse dimensioni e orientamento.



Il compito è quello di prevedere, a partire dalla matrice 16×16 di intensità di pixel, ogni immagine quale numero rappresenta in modo rapido e preciso.

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Riconoscimento di numeri scritti a mano

I dati provengono dai codici di avviamento postale scritti a mano su buste da lettera degli Stati Uniti. Ogni immagine rappresenta una cifra delle 5 del codice postale.

Le immagini sono 16×16 in scala di grigi a otto bit, con ogni pixel di intensità compresa tra 0 e 255. Sono state normalizzate per avere approssimativamente le stesse dimensioni e orientamento.



Il compito è quello di prevedere, a partire dalla matrice 16×16 di intensità di pixel, ogni immagine quale numero rappresenta in modo rapido e preciso.

E' un problema di **classificazione multiclasse** per il quale il tasso di errore deve essere mantenuto molto basso per evitare errori nello smistamento della posta.

Microarray DNA

Introduzione al corso

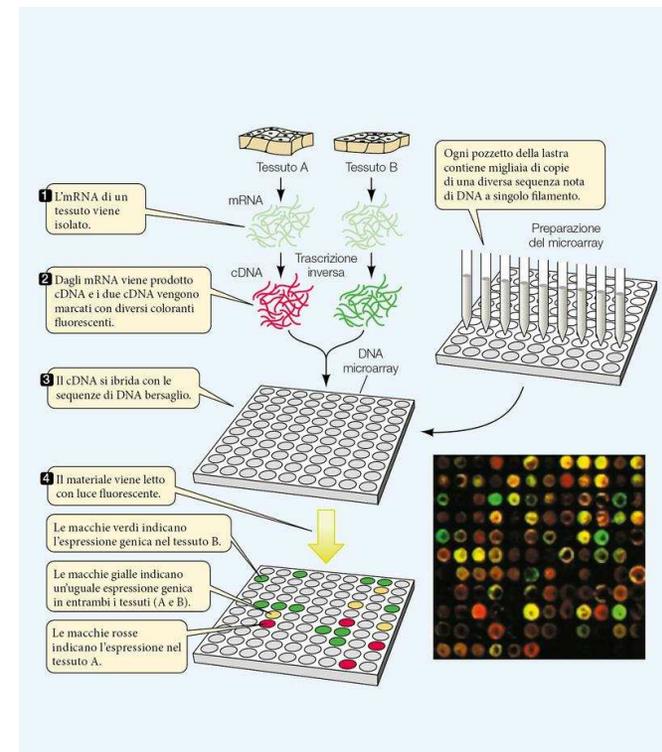
Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- **Microarray DNA**
- Dataset Genico

Statistical Learning

Due approcci semplici

Un microarray di DNA è un insieme di microscopiche sonde di DNA attaccate ad una superficie solida come chip di silicio formanti un array (matrice).



Microarray DNA

Introduzione al corso

Esempi di applicazioni

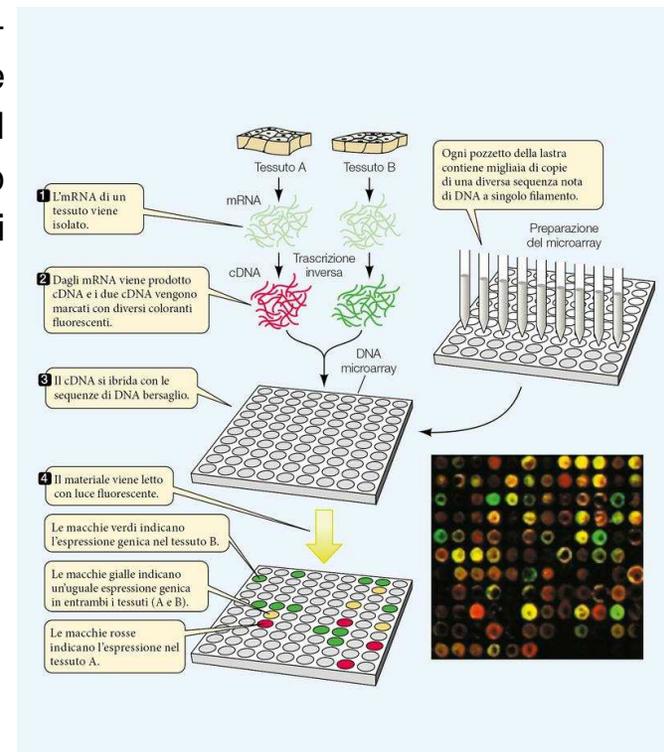
- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- **Microarray DNA**
- Dataset Genico

Statistical Learning

Due approcci semplici

Un microarray di DNA è un insieme di microscopiche sonde di DNA attaccate ad una superficie solida come chip di silicio formanti un array (matrice).

Permettono di esaminare simultaneamente la presenza di moltissimi geni all'interno di un campione di DNA. Un utilizzo tipico è quello di confrontare il profilo di espressione genica di un individuo malato con quello di uno sano per individuare quali geni sono coinvolti nella malattia.



Microarray DNA

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- **Microarray DNA**
- Dataset Genico

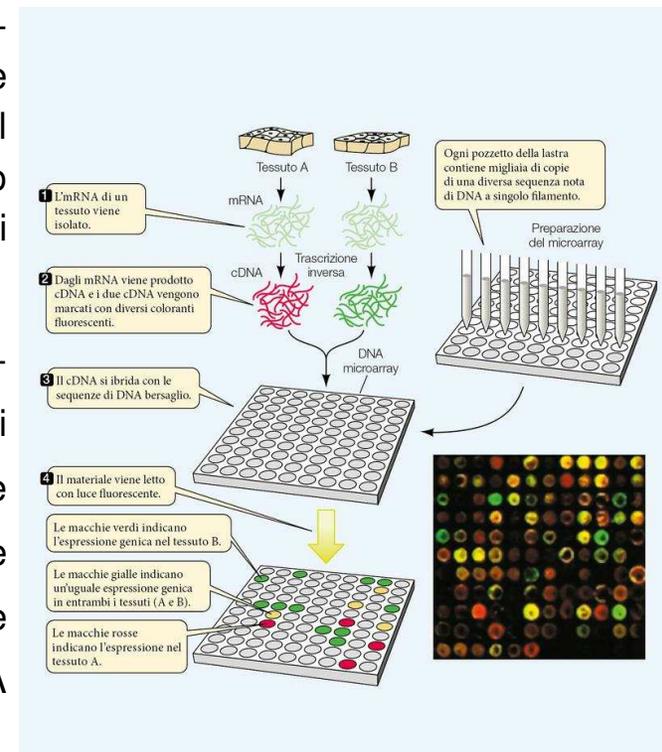
Statistical Learning

Due approcci semplici

Un microarray di DNA è un insieme di microscopiche sonde di DNA attaccate ad una superficie solida come chip di silicio formanti un array (matrice).

Permettono di esaminare simultaneamente la presenza di moltissimi geni all'interno di un campione di DNA. Un utilizzo tipico è quello di confrontare il profilo di espressione genica di un individuo malato con quello di uno sano per individuare quali geni sono coinvolti nella malattia.

I microarray sfruttano una tecnica di ibridazione inversa, che consiste nel fissare tutti i segmenti di DNA (detti probe) su un supporto e nel marcare invece l'acido nucleico che vogliamo identificare (detto target). Permette l'analisi dell'espressione genica monitorando in una sola volta gli RNA prodotti da migliaia di geni.



Dataset Genico

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Un dataset di espressione genica raccoglie insieme i valori di espressione di un serie di esperimenti di DNA microarray, con ogni colonna che rappresenta un esperimento. Ci sono quindi diverse migliaia di righe che rappresentano i geni di singoli individui, e decine di colonne che rappresentano i campioni

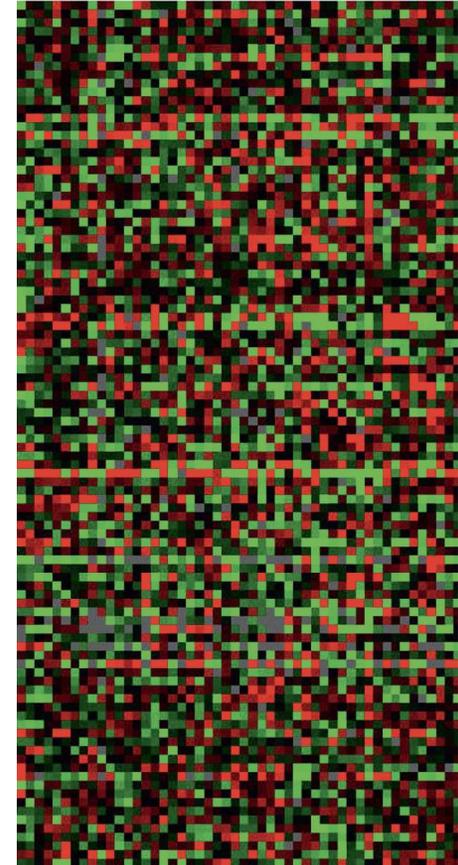


Figure 1: Mappa termica dei dati, che va dal verde (negativo) a rosso (positivo). I campioni sono 64 tessuti tumorali di diversi pazienti.

Dataset Genico

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Un dataset di espressione genica raccoglie insieme i valori di espressione di un serie di esperimenti di DNA microarray, con ogni colonna che rappresenta un esperimento. Ci sono quindi diverse migliaia di righe che rappresentano i geni di singoli individui, e decine di colonne che rappresentano i campioni

La sfida qui è capire come sono organizzati i geni e i campioni.

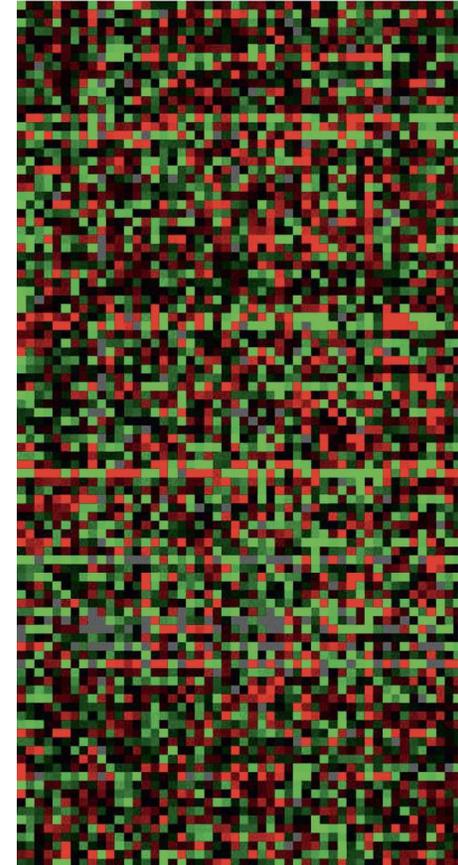


Figure 1: Mappa termica dei dati, che va dal verde (negativo) a rosso (positivo). I campioni sono 64 tessuti tumorali di diversi pazienti.

Dataset Genico

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Un dataset di espressione genica raccoglie insieme i valori di espressione di un serie di esperimenti di DNA microarray, con ogni colonna che rappresenta un esperimento. Ci sono quindi diverse migliaia di righe che rappresentano i geni di singoli individui, e decine di colonne che rappresentano i campioni

La sfida qui è capire come sono organizzati i geni e i campioni.

(a) quali campioni sono più simili l'uno all'altro, in termini di espressione di geni?

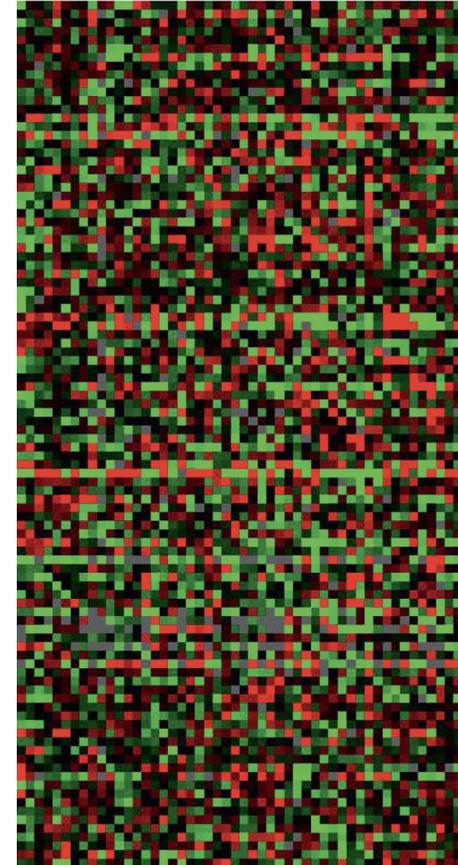


Figure 1: Mappa termica dei dati, che va dal verde (negativo) a rosso (positivo). I campioni sono 64 tessuti tumorali di diversi pazienti.

Dataset Genico

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Un dataset di espressione genica raccoglie insieme i valori di espressione di un serie di esperimenti di DNA microarray, con ogni colonna che rappresenta un esperimento. Ci sono quindi diverse migliaia di righe che rappresentano i geni di singoli individui, e decine di colonne che rappresentano i campioni

La sfida qui è capire come sono organizzati i geni e i campioni.

- quali campioni sono più simili l'uno all'altro, in termini di espressione di geni?
- quali sono i geni più simili tra loro, in termini di espressione tra i campioni?

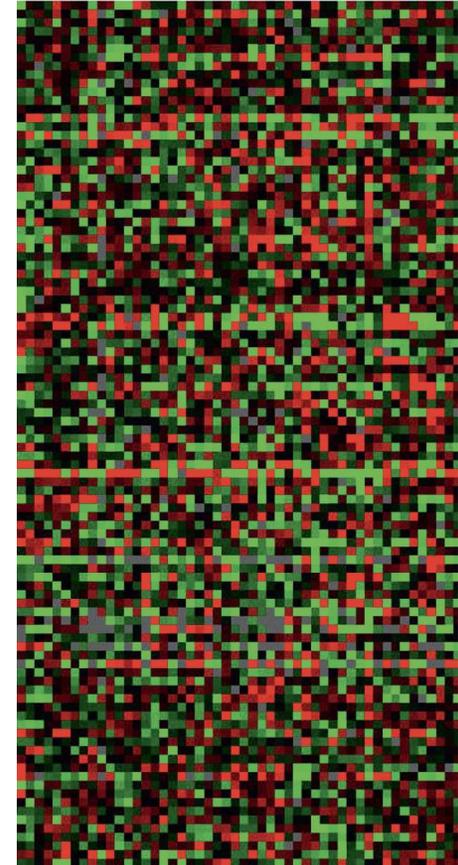


Figure 1: Mappa termica dei dati, che va dal verde (negativo) a rosso (positivo). I campioni sono 64 tessuti tumorali di diversi pazienti.

Dataset Genico

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Un dataset di espressione genica raccoglie insieme i valori di espressione di un serie di esperimenti di DNA microarray, con ogni colonna che rappresenta un esperimento. Ci sono quindi diverse migliaia di righe che rappresentano i geni di singoli individui, e decine di colonne che rappresentano i campioni

La sfida qui è capire come sono organizzati i geni e i campioni.

- quali campioni sono più simili l'uno all'altro, in termini di espressione di geni?
- quali sono i geni più simili tra loro, in termini di espressione tra i campioni?
- certi geni mostrano un'espressione molto alta (o bassa) per certi tipi di cancro?

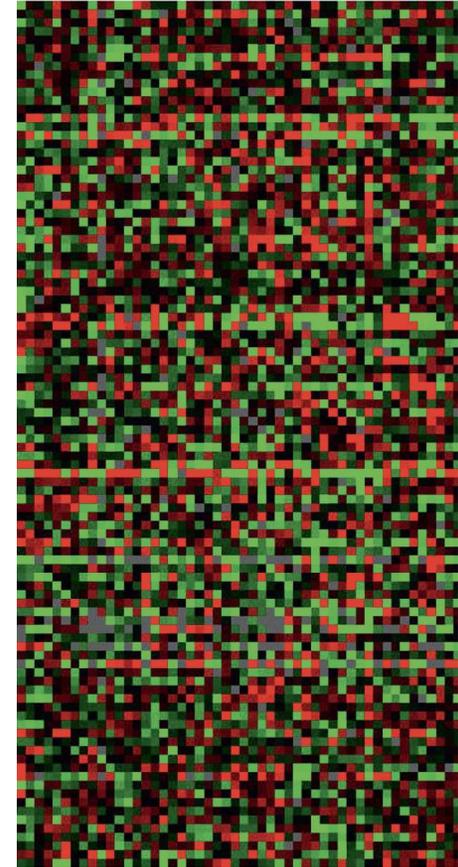


Figure 1: Mappa termica dei dati, che va dal verde (negativo) a rosso (positivo). I campioni sono 64 tessuti tumorali di diversi pazienti.

Dataset Genico

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Un dataset di espressione genica raccoglie insieme i valori di espressione di un serie di esperimenti di DNA microarray, con ogni colonna che rappresenta un esperimento. Ci sono quindi diverse migliaia di righe che rappresentano i geni di singoli individui, e decine di colonne che rappresentano i campioni

La sfida qui è capire come sono organizzati i geni e i campioni.

- quali campioni sono più simili l'uno all'altro, in termini di espressione di geni?
- quali sono i geni più simili tra loro, in termini di espressione tra i campioni?
- certi geni mostrano un'espressione molto alta (o bassa) per certi tipi di cancro?

Si può guardare a questo problema come un problema di regressione, in cui si cerca di predire il livello di espressione per coppie di geni-campioni.

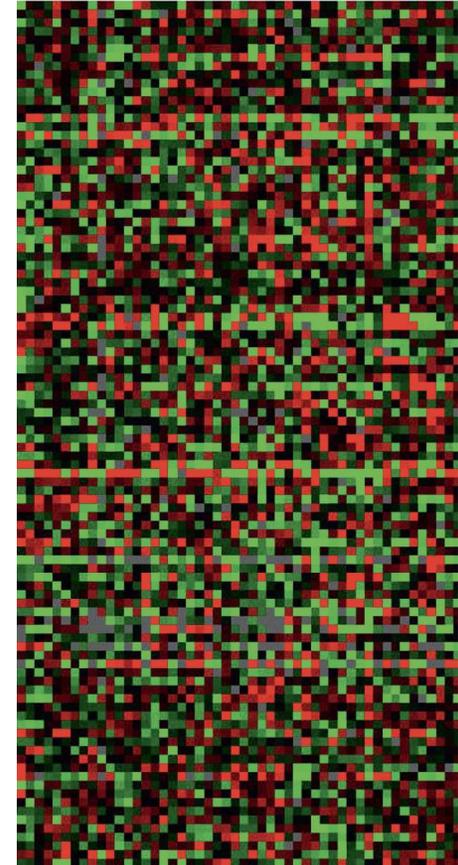


Figure 1: Mappa termica dei dati, che va dal verde (negativo) a rosso (positivo). I campioni sono 64 tessuti tumorali di diversi pazienti.

Dataset Genico

Introduzione al corso

Esempi di applicazioni

- Machine Learning
- Email Spam
- Cancro alla prostata
- Riconoscimento di numeri scritti a mano
- Microarray DNA
- Dataset Genico

Statistical Learning

Due approcci semplici

Un dataset di espressione genica raccoglie insieme i valori di espressione di un serie di esperimenti di DNA microarray, con ogni colonna che rappresenta un esperimento. Ci sono quindi diverse migliaia di righe che rappresentano i geni di singoli individui, e decine di colonne che rappresentano i campioni

La sfida qui è capire come sono organizzati i geni e i campioni.

- quali campioni sono più simili l'uno all'altro, in termini di espressione di geni?
- quali sono i geni più simili tra loro, in termini di espressione tra i campioni?
- certi geni mostrano un'espressione molto alta (o bassa) per certi tipi di cancro?

Si può guardare a questo problema come un problema di regressione, in cui si cerca di predire il livello di espressione per coppie di geni-campioni.

Tuttavia, è meglio vederlo come problema di **apprendimento non supervisionato**, in cui si cerca di costruire cluster di geni e/o campioni.

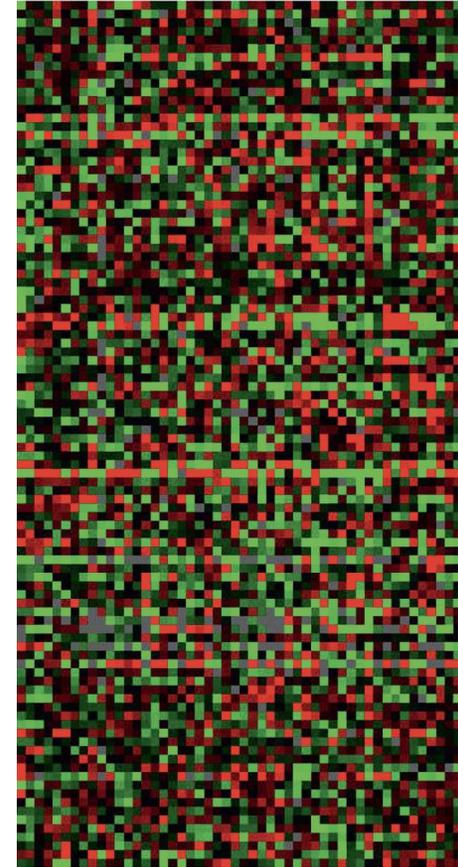


Figure 1: Mappa termica dei dati, che va dal verde (negativo) a rosso (positivo). I campioni sono 64 tessuti tumorali di diversi pazienti.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning

- Spam filtering

- Pericoli del ragionamento induttivo

- Ratti vs Piccioni

- Bias induttivo

- Quando serve il

Machine Learning?

- Tipi di learning

- Supervisionato vs

Non Supervisionato

- Attivo vs Passivo

- Utilità del docente

- Batch vs Online

- Un modello formale

- Generazione dei dati di training

- Misura di Successo

- ERM

- Errore

Due approcci semplici

Che cosa si intende con apprendimento?

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Che cosa si intende con apprendimento?

Bait Shyness (diffidenza verso le esche): Quando i ratti incontrano prodotti alimentari dall'aspetto o dall'odore nuovi, ne mangeranno prima piccole quantità, e se lo rimangiano o meno dipenderà dal sapore del cibo e dal suo effetto su di loro. Se l'alimento produce un effetto nocivo, il nuovo alimento sarà spesso associato con la malattia e, di conseguenza, i topi non lo mangeranno.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

● Statistical Learning

● Spam filtering

● Pericoli del
ragionamento induttivo

● Ratti vs Piccioni

● Bias induttivo

● Quando serve il

Machine Learning?

● Tipi di learning

● Supervisionato vs

Non Supervisionato

● Attivo vs Passivo

● Utilità del docente

● Batch vs Online

● Un modello formale

● Generazione dei dati
di training

● Misura di Successo

● ERM

● Errore

Due approcci semplici

Che cosa si intende con apprendimento?

Bait Shyness (diffidenza verso le esche): Quando i ratti incontrano prodotti alimentari dall'aspetto o dall'odore nuovi, ne mangeranno prima piccole quantità, e se lo rimangiano o meno dipenderà dal sapore del cibo e dal suo effetto su di loro. Se l'alimento produce un effetto nocivo, il nuovo alimento sarà spesso associato con la malattia e, di conseguenza, i topi non lo mangeranno.

Chiaramente, c'è un meccanismo di apprendimento in gioco qui: l'animale ha usato l'esperienza del passato per acquisire esperienza nell'individuazione della sicurezza di questo alimento. Se grazie all'esperienza passata il cibo è stato etichettato negativamente, l'animale prevede che avrà anche un effetto negativo in futuro.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

● Statistical Learning

● Spam filtering

● Pericoli del
ragionamento induttivo

● Ratti vs Piccioni

● Bias induttivo

● Quando serve il

Machine Learning?

● Tipi di learning

● Supervisionato vs

Non Supervisionato

● Attivo vs Passivo

● Utilità del docente

● Batch vs Online

● Un modello formale

● Generazione dei dati
di training

● Misura di Successo

● ERM

● Errore

Due approcci semplici

Che cosa si intende con apprendimento?

Bait Shyness (diffidenza verso le esche): Quando i ratti incontrano prodotti alimentari dall'aspetto o dall'odore nuovi, ne mangeranno prima piccole quantità, e se lo rimangiano o meno dipenderà dal sapore del cibo e dal suo effetto su di loro. Se l'alimento produce un effetto nocivo, il nuovo alimento sarà spesso associato con la malattia e, di conseguenza, i topi non lo mangeranno.

Chiaramente, c'è un meccanismo di apprendimento in gioco qui: l'animale ha usato l'esperienza del passato per acquisire esperienza nell'individuazione della sicurezza di questo alimento. Se grazie all'esperienza passata il cibo è stato etichettato negativamente, l'animale prevede che avrà anche un effetto negativo in futuro.

Proviamo a estendere questo approccio al problema di filtro delle email spam

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo

● Ratti vs Piccioni

● Bias induttivo

● Quando serve il

Machine Learning?

● Tipi di learning

● Supervisionato vs

Non Supervisionato

● Attivo vs Passivo

● Utilità del docente

● Batch vs Online

● Un modello formale

● Generazione dei dati di training

● Misura di Successo

● ERM

● Errore

Due approcci semplici

Una soluzione naive (learning by memorization) sarebbe apparentemente simile al modo in cui i topi imparano ad evitare le esche velenose: la macchina memorizza tutte le e-mail precedenti etichettate come e-mail di spam dall'applicazione utente umano.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
 - Ratti vs Piccioni
 - Bias induttivo
 - Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Una soluzione naive (learning by memorization) sarebbe apparentemente simile al modo in cui i topi imparano ad evitare le esche velenose: la macchina memorizza tutte le e-mail precedenti etichettate come e-mail di spam dall'applicazione utente umano.

Quando arriva una nuova e-mail, la macchina la cercherà nel set di e-mail di spam precedenti. Se corrisponde ad una di esse, sarà distrutta. Altrimenti, sarà spostato nella cartella di posta in arrivo dell'utente.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
 - Ratti vs Piccioni
 - Bias induttivo
 - Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Una soluzione naive (learning by memorization) sarebbe apparentemente simile al modo in cui i topi imparano ad evitare le esche velenose: la macchina memorizza tutte le e-mail precedenti etichettate come e-mail di spam dall'applicazione utente umano.

Quando arriva una nuova e-mail, la macchina la cercherà nel set di e-mail di spam precedenti. Se corrisponde ad una di esse, sarà distrutta. Altrimenti, sarà spostato nella cartella di posta in arrivo dell'utente.

Manca la capacità di etichettare i messaggi email mai visti, ovvero la capacità di chi apprende di generalizzare dagli esempi visti. Questa capacità è indicata con il nome di **ragionamento induttivo**.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Una soluzione naive (learning by memorization) sarebbe apparentemente simile al modo in cui i topi imparano ad evitare le esche velenose: la macchina memorizza tutte le e-mail precedenti etichettate come e-mail di spam dall'applicazione utente umano.

Quando arriva una nuova e-mail, la macchina la cercherà nel set di e-mail di spam precedenti. Se corrisponde ad una di esse, sarà distrutta. Altrimenti, sarà spostato nella cartella di posta in arrivo dell'utente.

Manca la capacità di etichettare i messaggi email mai visti, ovvero la capacità di chi apprende di generalizzare dagli esempi visti. Questa capacità è indicata con il nome di **ragionamento induttivo**.

Nell'esempio dei ratti, dopo che i ratti incontrano un esempio di un certo tipo di cibo, applicano la regola imparata anche a nuovi esempi di cibo con odore e gusto simili.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo

- Ratti vs Piccioni

- Bias induttivo
- Quando serve il Machine Learning?

- Tipi di learning
- Supervisionato vs Non Supervisionato

- Attivo vs Passivo
- Utilità del docente
- Batch vs Online

- Un modello formale
- Generazione dei dati di training

- Misura di Successo

- ERM
- Errore

Due approcci semplici

Una soluzione naive (learning by memorization) sarebbe apparentemente simile al modo in cui i topi imparano ad evitare le esche velenose: la macchina memorizza tutte le e-mail precedenti etichettate come e-mail di spam dall'applicazione utente umano.

Quando arriva una nuova e-mail, la macchina la cercherà nel set di e-mail di spam precedenti. Se corrisponde ad una di esse, sarà distrutta. Altrimenti, sarà spostato nella cartella di posta in arrivo dell'utente.

Manca la capacità di etichettare i messaggi email mai visti, ovvero la capacità di chi apprende di generalizzare dagli esempi visti. Questa capacità è indicata con il nome di **ragionamento induttivo**.

Nell'esempio dei ratti, dopo che i ratti incontrano un esempio di un certo tipo di cibo, applicano la regola imparata anche a nuovi esempi di cibo con odore e gusto simili.

Per il filtraggio dello spam, il learner può estrarre dalle email classificate come spam un insieme di parole ricorrenti. Quando arriva una nuova e-mail, la macchina può controllare se una delle parole sospette compare in essa, e predire l'etichetta di conseguenza.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Superstizione del Piccione

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Superstizione del Piccione

In un esperimento lo psicologo B. F. Skinner ha messo un gruppo di piccioni affamati in una gabbia. Un meccanismo automatico attaccato alla gabbia, consegnava il cibo ai piccioni ad intervalli regolari senza alcun riferimento al comportamento degli uccelli. I piccioni affamati sono andati intorno alla gabbia, e quando il cibo è stato consegnato per la prima volta, ha trovato ogni piccione impegnato in qualche attività (beccare, girare la testa, ecc.).

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Superstizione del Piccione

In un esperimento lo psicologo B. F. Skinner ha messo un gruppo di piccioni affamati in una gabbia. Un meccanismo automatico attaccato alla gabbia, consegnava il cibo ai piccioni ad intervalli regolari senza alcun riferimento al comportamento degli uccelli. I piccioni affamati sono andati intorno alla gabbia, e quando il cibo è stato consegnato per la prima volta, ha trovato ogni piccione impegnato in qualche attività (beccare, girare la testa, ecc.).

L'arrivo del cibo in corrispondenza di una certa azione faceva sì che ogni uccello tendesse a spendere sempre più tempo a fare la stessa cosa. Questo, a sua volta, ha aumentato la possibilità che la prossima consegna casuale di cibo trovasse ogni uccello impegnato in quell'attività.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Superstizione del Piccione

In un esperimento lo psicologo B. F. Skinner ha messo un gruppo di piccioni affamati in una gabbia. Un meccanismo automatico attaccato alla gabbia, consegnava il cibo ai piccioni ad intervalli regolari senza alcun riferimento al comportamento degli uccelli. I piccioni affamati sono andati intorno alla gabbia, e quando il cibo è stato consegnato per la prima volta, ha trovato ogni piccione impegnato in qualche attività (beccare, girare la testa, ecc.).

L'arrivo del cibo in corrispondenza di una certa azione faceva sì che ogni uccello tendesse a spendere sempre più tempo a fare la stessa cosa. Questo, a sua volta, ha aumentato la possibilità che la prossima consegna casuale di cibo trovasse ogni uccello impegnato in quell'attività.

Il risultato è una catena di eventi che rafforza l'associazione dei piccioni della consegna del cibo associata alle azioni che stavano compiendo casualmente quando è stato consegnato per la prima volta. Per questo continuano ad eseguire tali azioni stesse azioni diligentemente.

Ratti vs Piccioni

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo

● **Ratti vs Piccioni**

- Bias induttivo
- Quando serve il Machine Learning?

- Tipi di learning
- Supervisionato vs Non Supervisionato

- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training

- Misura di Successo
- ERM
- Errore

Due approcci semplici

Cosa distingue i meccanismi di apprendimento che portano alla superstizione da quelli utili da Imparare?

Ratti vs Piccioni

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- **Ratti vs Piccioni**
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Cosa distingue i meccanismi di apprendimento che portano alla superstizione da quelli utili da imparare?

Questa domanda è cruciale: una volta che esportiamo il compito di imparare a una macchina, dobbiamo fornire principi precisi e ben definiti che proteggano il programma dal raggiungere conclusioni insensate o inutili.

Ratti vs Piccioni

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- **Ratti vs Piccioni**
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Cosa distingue i meccanismi di apprendimento che portano alla superstizione da quelli utili da imparare?

Questa domanda è cruciale: una volta che esportiamo il compito di imparare a una macchina, dobbiamo fornire principi precisi e ben definiti che proteggano il programma dal raggiungere conclusioni insensate o inutili.

Bait Shyness II - i ratti non riescono ad acquisire condizionamento tra cibo e scossa elettrica. Negli esperimenti effettuati da Garcia (Garcia & Koelling 1996), è stato dimostrato che, se l'operazione di sgradevole stimolo che segue il consumo di cibo è sostituito da, ad esempio, elettro shock, non si verifica alcun condizionamento.

Ratti vs Piccioni

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- **Ratti vs Piccioni**
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Cosa distingue i meccanismi di apprendimento che portano alla superstizione da quelli utili da imparare?

Questa domanda è cruciale: una volta che esportiamo il compito di imparare a una macchina, dobbiamo fornire principi precisi e ben definiti che proteggano il programma dal raggiungere conclusioni insensate o inutili.

Bait Shyness II - i ratti non riescono ad acquisire condizionamento tra cibo e scossa elettrica. Negli esperimenti effettuati da Garcia (Garcia & Koelling 1996), è stato dimostrato che, se l'operazione di sgradevole stimolo che segue il consumo di cibo è sostituito da, ad esempio, elettro shock, non si verifica alcun condizionamento.

Anche dopo ripetute prove in cui il consumo di alcuni alimenti è seguito dalla somministrazione di spiacevole scossa elettrica, i ratti non tendono ad evitare quel cibo.

Ratti vs Piccioni

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- **Ratti vs Piccioni**
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Cosa distingue i meccanismi di apprendimento che portano alla superstizione da quelli utili da Imparare?

Questa domanda è cruciale: una volta che esportiamo il compito di imparare a una macchina, dobbiamo fornire principi precisi e ben definiti che proteggano il programma dal raggiungere conclusioni insensate o inutili.

Bait Shyness II - i ratti non riescono ad acquisire condizionamento tra cibo e scossa elettrica. Negli esperimenti effettuati da Garcia (Garcia & Koelling 1996), è stato dimostrato che, se l'operazione di sgradevole stimolo che segue il consumo di cibo è sostituito da, ad esempio, elettro shock, non si verifica alcun condizionamento.

Anche dopo ripetute prove in cui il consumo di alcuni alimenti è seguito dalla somministrazione di spiacevole scossa elettrica, i ratti non tendono ad evitare quel cibo.

I ratti sembrano avere qualche conoscenza innata preventiva che dice loro che, mentre la correlazione temporale tra cibo e nausea può essere causale, è improbabile che ci sia un rapporto causale tra consumo alimentare e shock elettrici .

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo

- Ratti vs Piccioni

- **Bias induttivo**

- Quando serve il Machine Learning?

- Tipi di learning

- Supervisionato vs Non Supervisionato

- Attivo vs Passivo

- Utilità del docente

- Batch vs Online

- Un modello formale

- Generazione dei dati di training

- Misura di Successo

- ERM

- Errore

Due approcci semplici

La caratteristica distintiva tra l'apprendimento dei ratti e quello dei piccioni è l'aver tenuto conto di qualche innata conoscenza apriori, che ha influenzato il meccanismo di apprendimento. Questo è chiamato anche **bias induttivo**.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- **Bias induttivo**
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

La caratteristica distintiva tra l'apprendimento dei ratti e quello dei piccioni è l'aver tenuto conto di qualche innata conoscenza apriori, che ha influenzato il meccanismo di apprendimento. Questo è chiamato anche **bias induttivo**.

I piccioni sono disposti ad adottare qualsiasi spiegazione per la presenza di cibo, mentre i ratti “sanno” che il cibo non può causare una scossa elettrica. Il processo di apprendimento dei ratti è orientato verso il rilevamento di un certo tipo di modello, ignorando altre correlazioni temporali tra gli eventi

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- **Bias induttivo**
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

La caratteristica distintiva tra l'apprendimento dei ratti e quello dei piccioni è l'aver tenuto conto di qualche innata conoscenza apriori, che ha influenzato il meccanismo di apprendimento. Questo è chiamato anche **bias induttivo**.

I piccioni sono disposti ad adottare qualsiasi spiegazione per la presenza di cibo, mentre i ratti “sanno” che il cibo non può causare una scossa elettrica. Il processo di apprendimento dei ratti è orientato verso il rilevamento di un certo tipo di modello, ignorando altre correlazioni temporali tra gli eventi

L'utilizzo di conoscenze pregresse sul problema, che influenzano l'apprendimento è necessario per il successo degli algoritmi di apprendimento.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- **Bias induttivo**
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

La caratteristica distintiva tra l'apprendimento dei ratti e quello dei piccioni è l'aver tenuto conto di qualche innata conoscenza apriori, che ha influenzato il meccanismo di apprendimento. Questo è chiamato anche **bias induttivo**.

I piccioni sono disposti ad adottare qualsiasi spiegazione per la presenza di cibo, mentre i ratti "sanno" che il cibo non può causare una scossa elettrica. Il processo di apprendimento dei ratti è orientato verso il rilevamento di un certo tipo di modello, ignorando altre correlazioni temporali tra gli eventi

L'utilizzo di conoscenze pregresse sul problema, che influenzano l'apprendimento è necessario per il successo degli algoritmi di apprendimento.

Più forte è la conoscenza pregressa (o ipotesi) immagazzinata nel processo di apprendimento, più è facile imparare da ulteriori esempi.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- **Bias induttivo**
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

La caratteristica distintiva tra l'apprendimento dei ratti e quello dei piccioni è l'aver tenuto conto di qualche innata conoscenza apriori, che ha influenzato il meccanismo di apprendimento. Questo è chiamato anche **bias induttivo**.

I piccioni sono disposti ad adottare qualsiasi spiegazione per la presenza di cibo, mentre i ratti “sanno” che il cibo non può causare una scossa elettrica. Il processo di apprendimento dei ratti è orientato verso il rilevamento di un certo tipo di modello, ignorando altre correlazioni temporali tra gli eventi

L'utilizzo di conoscenze pregresse sul problema, che influenzano l'apprendimento è necessario per il successo degli algoritmi di apprendimento.

Più forte è la conoscenza pregressa (o ipotesi) immagazzinata nel processo di apprendimento, più è facile imparare da ulteriori esempi.

Tuttavia, più forte è il bias induttivo, meno flessibile è l'apprendimento, in quanto vincolato dalle ipotesi induttive

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo

- Ratti vs Piccioni

- Bias induttivo

- Quando serve il Machine Learning?

- Tipi di learning

- Supervisionato vs Non Supervisionato

- Attivo vs Passivo

- Utilità del docente

- Batch vs Online

- Un modello formale

- Generazione dei dati di training

- Misura di Successo

- ERM

- Errore

Quando serve il Machine Learning?

Task abituali per umani/animali Ma troppo difficili per scrivere un programma in grado di riprodurli: riconoscimento immagini, guida automatica, riconoscimento vocale.

Quando serve il Machine Learning?

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Task abituali per umani/animali Ma troppo difficili per scrivere un programma in grado di riprodurli: riconoscimento immagini, guida automatica, riconoscimento vocale.

Task al di là delle capacità umane Analisi di moli di dati troppo grandi e complesse: dati astronomici, medici, transazioni web etc

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Quando serve il Machine Learning?

Task abituali per umani/animali Ma troppo difficili per scrivere un programma in grado di riprodurli: riconoscimento immagini, guida automatica, riconoscimento vocale.

Task al di là delle capacità umane Analisi di moli di dati troppo grandi e complesse: dati astronomici, medici, transazioni web etc

Task che richiedono adattabilità problemi di spam detection, riconoscimento vocale o di scrittura a mano (cambiano al cambiare dell'utente)

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- **Tipi di learning**
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Supervisionato vs Non supervisionato

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- **Tipi di learning**
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Supervisionato vs Non supervisionato

Learning Attivo vs Passivo

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- **Tipi di learning**
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Supervisionato vs Non supervisionato

Learning Attivo vs Passivo

Insegnante di aiuto o no

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- **Tipi di learning**
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Supervisionato vs Non supervisionato

Learning Attivo vs Passivo

Insegnante di aiuto o no

Protocollo Batch vs Online

Supervisionato vs Non Supervisionato

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo

- Ratti vs Piccioni

- Bias induttivo

- Quando serve il

Machine Learning?

- Tipi di learning

● Supervisionato vs
Non Supervisionato

- Attivo vs Passivo

- Utilità del docente

- Batch vs Online

- Un modello formale

- Generazione dei dati di training

- Misura di Successo

- ERM

- Errore

Due approcci semplici

Considerando l'apprendimento come un processo di “utilizzo dell'esperienza per acquisire competenze”, l'apprendimento supervisionato descrive uno scenario in cui l'esperienza contiene informazioni significative (ad esempio, le etichette spam/non spam) che manca negli esempi di test non visti a cui applicare le competenze acquisite.

Supervisionato vs Non Supervisionato

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- **Supervisionato vs Non Supervisionato**
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Considerando l'apprendimento come un processo di “utilizzo dell'esperienza per acquisire competenze”, l'apprendimento supervisionato descrive uno scenario in cui l'esperienza contiene informazioni significative (ad esempio, le etichette spam/non spam) che manca negli esempi di test non visti a cui applicare le competenze acquisite.

Si può pensare all'ambiente come a un insegnante che “supervisiona” il discente fornendo le informazioni aggiuntive (etichette).

Supervisionato vs Non Supervisionato

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- **Supervisionato vs Non Supervisionato**
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Considerando l'apprendimento come un processo di “utilizzo dell'esperienza per acquisire competenze”, l'apprendimento supervisionato descrive uno scenario in cui l'esperienza contiene informazioni significative (ad esempio, le etichette spam/non spam) che manca negli esempi di test non visti a cui applicare le competenze acquisite.

Si può pensare all'ambiente come a un insegnante che “supervisiona” il discente fornendo le informazioni aggiuntive (etichette).

Nell'apprendimento non supervisionato, tuttavia, non vi è distinzione tra i dati di training e i dati dei test. L'allievo elabora i dati di input con l'obiettivo di estrarne informazione o ottenerne una versione compressa.

Supervisionato vs Non Supervisionato

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- **Supervisionato vs Non Supervisionato**
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Considerando l'apprendimento come un processo di “utilizzo dell'esperienza per acquisire competenze”, l'apprendimento supervisionato descrive uno scenario in cui l'esperienza contiene informazioni significative (ad esempio, le etichette spam/non spam) che manca negli esempi di test non visti a cui applicare le competenze acquisite.

Si può pensare all'ambiente come a un insegnante che “supervisiona” il discente fornendo le informazioni aggiuntive (etichette).

Nell'apprendimento non supervisionato, tuttavia, non vi è distinzione tra i dati di training e i dati dei test. L'allievo elabora i dati di input con l'obiettivo di estrarne informazione o ottenerne una versione compressa.

Esempio: spam detection vs anomaly detection

Supervisionato vs Non Supervisionato

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- **Supervisionato vs Non Supervisionato**
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Considerando l'apprendimento come un processo di “utilizzo dell'esperienza per acquisire competenze”, l'apprendimento supervisionato descrive uno scenario in cui l'esperienza contiene informazioni significative (ad esempio, le etichette spam/non spam) che manca negli esempi di test non visti a cui applicare le competenze acquisite.

Si può pensare all'ambiente come a un insegnante che “supervisiona” il discente fornendo le informazioni aggiuntive (etichette).

Nell'apprendimento non supervisionato, tuttavia, non vi è distinzione tra i dati di training e i dati dei test. L'allievo elabora i dati di input con l'obiettivo di estrarne informazione o ottenerne una versione compressa.

Esempio: spam detection vs anomaly detection

Ambiente di apprendimento intermedio in cui, mentre i dati di training contengono più informazioni rispetto agli esempi di test, si vogliono prevedere informazioni ulteriori sugli esempi di test. In questo caso si parla di **apprendimento per rinforzo**

Attivo vs Passivo

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

I paradigmi di apprendimento possono variare a seconda del ruolo giocato

Attivo vs Passivo

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

I paradigmi di apprendimento possono variare a seconda del ruolo giocato

Distinguiamo tra discente “attivo” e “passivo”. Un discente attivo interagisce con l’ambiente durante la formazione, ad esempio, ponendo domande o eseguendo esperimenti, mentre il discente passivo osserva solo le informazioni fornite dall’ambiente (o l’insegnante) senza influenzare o dirigere.

Attivo vs Passivo

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- **Attivo vs Passivo**
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

I paradigmi di apprendimento possono variare a seconda del ruolo giocato

Distinguiamo tra discente “attivo” e “passivo”. Un discente attivo interagisce con l’ambiente durante la formazione, ad esempio, ponendo domande o eseguendo esperimenti, mentre il discente passivo osserva solo le informazioni fornite dall’ambiente (o l’insegnante) senza influenzare o dirigere.

Un filtro antispam è solitamente passivo - in attesa che gli utenti contrassegnino le e-mail che arrivano. In un ambiente attivo, si potrebbe immaginare di chiedere agli utenti di etichettare e-mail specifiche scelte dall’allievo, o addirittura composte dall’allievo, per migliorare la sua comprensione di ciò che è spam.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo

- Ratti vs Piccioni

- Bias induttivo

- Quando serve il

Machine Learning?

- Tipi di learning

- Supervisionato vs

Non Supervisionato

- Attivo vs Passivo

- **Utilità del docente**

- Batch vs Online

- Un modello formale

- Generazione dei dati di training

- Misura di Successo

- ERM

- Errore

Due approcci semplici

Nel caso dell'apprendimento umano (si pensi a un bambino a casa o a scuola) il processo coinvolge spesso un insegnante fondamentale, che sta cercando di fornire all'allievo le informazioni più utili per raggiungere l'obiettivo di apprendimento.

Utilità del docente

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Nel caso dell'apprendimento umano (si pensi a un bambino a casa o a scuola) il processo coinvolge spesso un insegnante fondamentale, che sta cercando di fornire all'allievo le informazioni più utili per raggiungere l'obiettivo di apprendimento.

Al contrario, quando uno scienziato studia la natura, l'ambiente, giocando il ruolo dell'insegnante, può essere pensata come passiva - i fenomeni naturali si manifestano a prescindere dalle esigenze del discente.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- **Utilità del docente**
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Nel caso dell'apprendimento umano (si pensi a un bambino a casa o a scuola) il processo coinvolge spesso un insegnante fondamentale, che sta cercando di fornire all'allievo le informazioni più utili per raggiungere l'obiettivo di apprendimento.

Al contrario, quando uno scienziato studia la natura, l'ambiente, giocando il ruolo dell'insegnante, può essere pensata come passiva - i fenomeni naturali si manifestano a prescindere dalle esigenze del discente.

Questi scenari di apprendimento sono modellati assumendo che i dati del training (o l'esperienza del discente) siano generato da un processo casuale.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- **Utilità del docente**
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Nel caso dell'apprendimento umano (si pensi a un bambino a casa o a scuola) il processo coinvolge spesso un insegnante fondamentale, che sta cercando di fornire all'allievo le informazioni più utili per raggiungere l'obiettivo di apprendimento.

Al contrario, quando uno scienziato studia la natura, l'ambiente, giocando il ruolo dell'insegnante, può essere pensata come passiva - i fenomeni naturali si manifestano a prescindere dalle esigenze del discente.

Questi scenari di apprendimento sono modellati assumendo che i dati del training (o l'esperienza del discente) siano generato da un processo casuale.

In alcuni casi, si assume che l'input dell'allievo sia generato da un "insegnante avversario". Questo può essere il caso dell'esempio di filtraggio dello spam (se lo spammer fa uno sforzo per indurre in errore il progettista del filtro antispam) o nell'imparare a rilevare le frodi.

Utilità del docente

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- **Utilità del docente**
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Nel caso dell'apprendimento umano (si pensi a un bambino a casa o a scuola) il processo coinvolge spesso un insegnante fondamentale, che sta cercando di fornire all'allievo le informazioni più utili per raggiungere l'obiettivo di apprendimento.

Al contrario, quando uno scienziato studia la natura, l'ambiente, giocando il ruolo dell'insegnante, può essere pensata come passiva - i fenomeni naturali si manifestano a prescindere dalle esigenze del discente.

Questi scenari di apprendimento sono modellati assumendo che i dati del training (o l'esperienza del discente) siano generato da un processo casuale.

In alcuni casi, si assume che l'input dell'allievo sia generato da un "insegnante avversario". Questo può essere il caso dell'esempio di filtraggio dello spam (se lo spammer fa uno sforzo per indurre in errore il progettista del filtro antispam) o nell'imparare a rilevare le frodi.

Si usa anche un modello di insegnamento avversario come scenario peggiore, quando non si può supporre con sicurezza una configurazione più blanda.

Batch vs Online

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo

- Ratti vs Piccioni

- Bias induttivo
- Quando serve il Machine Learning?

- Tipi di learning
- Supervisionato vs Non Supervisionato

- Attivo vs Passivo
- Utilità del docente

- **Batch vs Online**
- Un modello formale
- Generazione dei dati di training

- Misura di Successo
- ERM
- Errore

Due approcci semplici

Si deve distinguere tra situazioni in cui il discente deve rispondere online, durante tutto il processo di apprendimento, e le impostazioni in cui l'allievo deve impegnare le competenze acquisite solo dopo aver avuto la possibilità di elaborare grandi quantità di dati.

Batch vs Online

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- **Batch vs Online**
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Si deve distinguere tra situazioni in cui il discente deve rispondere online, durante tutto il processo di apprendimento, e le impostazioni in cui l'allievo deve impegnare le competenze acquisite solo dopo aver avuto la possibilità di elaborare grandi quantità di dati.

Ad esempio, un agente di cambio deve prendere decisioni ogni giorno, sulla base dell'esperienza acquisita finora. Può diventare un esperto nel tempo, ma potrebbe aver commesso errori costosi nel processo. In contrasto, in molte impostazioni di data mining, l'allievo ha grandi quantità di dati di allenamento con cui giocare prima di dover produrre delle uscite definitive.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- **Un modello formale**
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Input dell'algoritmo di apprendimento (supervisionato):

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- **Un modello formale**
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Input dell'algoritmo di apprendimento (supervisionato):

1. Spazio di input (o spazio delle istanze): un insieme arbitrario \mathcal{X} , che rappresenta l'insieme di dati che si vuole etichettare (ad es. le papaya). Di solito i punti di \mathcal{X} vengono rappresentati tramite un vettore di **features** (colore e morbidezza della papaya) che è una variabile random

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- **Un modello formale**
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Input dell'algoritmo di apprendimento (supervisionato):

1. Spazio di input (o spazio delle istanze): un insieme arbitrario \mathcal{X} , che rappresenta l'insieme di dati che si vuole etichettare (ad es. le papaya). Di solito i punti di \mathcal{X} vengono rappresentati tramite un vettore di **features** (colore e morbidezza della papaya) che è una variabile random
2. Insieme di etichette: \mathcal{Y} valori possibili delle etichette, ad es $\mathcal{Y} = \{0, 1\}$ (ad es cattiva, gustosa)

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Input dell'algoritmo di apprendimento (supervisionato):

1. Spazio di input (o spazio delle istanze): un insieme arbitrario \mathcal{X} , che rappresenta l'insieme di dati che si vuole etichettare (ad es. le papaya). Di solito i punti di \mathcal{X} vengono rappresentati tramite un vettore di **features** (colore e morbidezza della papaya) che è una variabile random
2. Insieme di etichette: \mathcal{Y} valori possibili delle etichette, ad es $\mathcal{Y} = \{0, 1\}$ (ad es cattiva, gustosa)
3. Training set: $T = \{(x^1, y^1), \dots, (x^l, y^l)\}$ è un insieme finito di coppie in $\mathcal{X} \times \mathcal{Y}$, ovvero una sequenza di punti dello spazio di input con relativa etichetta (papaya che sono state assaggiate con relativo colore, morbidezza e giudizio sul gusto).

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Un modello formale

Input dell'algoritmo di apprendimento (supervisionato):

1. Spazio di input (o spazio delle istanze): un insieme arbitrario \mathcal{X} , che rappresenta l'insieme di dati che si vuole etichettare (ad es. le papaya). Di solito i punti di \mathcal{X} vengono rappresentati tramite un vettore di **features** (colore e morbidezza della papaya) che è una variabile random
2. Insieme di etichette: \mathcal{Y} valori possibili delle etichette, ad es $\mathcal{Y} = \{0, 1\}$ (ad es cattiva, gustosa)
3. Training set: $T = \{(x^1, y^1), \dots, (x^l, y^l)\}$ è un insieme finito di coppie in $\mathcal{X} \times \mathcal{Y}$, ovvero una sequenza di punti dello spazio di input con relativa etichetta (papaya che sono state assaggiate con relativo colore, morbidezza e giudizio sul gusto).

Output dell'algoritmo di apprendimento: Una regola di predizione $\hat{f} : \mathcal{X} \rightarrow \mathcal{Y}$, che può essere usata per predire l'etichetta dei nuovi punti (nuove papaya)

Generazione dei dati di training

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Si assume che i dati di training siano generati in accordo a una qualche distribuzione congiunta di probabilità $\mathcal{P}(x, y)$.

Generazione dei dati di training

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Si assume che i dati di training siano generati in accordo a una qualche distribuzione congiunta di probabilità $\mathcal{P}(x, y)$.

La distribuzione di probabilità non è nota.

Generazione dei dati di training

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Si assume che i dati di training siano generati in accordo a una qualche distribuzione congiunta di probabilità $\mathcal{P}(x, y)$.

La distribuzione di probabilità non è nota.

In pratica, ogni coppia (x^i, y^i) del training set si assume che sia stata generata in maniera indipendente e identicamente distribuita secondo la distribuzione di probabilità $\mathcal{P}(X, Y)$.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- **Misura di Successo**
- ERM
- Errore

Due approcci semplici

Ci piacerebbe avere il metodo di apprendimento che minimizza l'errore atteso su qualunque esempio estratto dalla distribuzione (**rischio effettivo**). Supponiamo di introdurre una **funzione di Loss** $L(Y, \hat{f}(X))$ per penalizzare errori in predizione. Un esempio naturale di funzione di loss è l'errore quadratico (per problemi di regressione)

$$L(Y, \hat{f}(X)) = (Y - \hat{f}(X))^2,$$

per cui ci piacerebbe minimizzare

$$E(Y - \hat{f}(X))^2 = \int (Y - \hat{f}(X))^2 \mathcal{P}(dx, dy)$$

che però non possiamo calcolare.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- **ERM**
- Errore

Due approcci semplici

Non conoscendo la distribuzione di probabilità quello che si fa è minimizzare il cosiddetto rischio empirico:

$$R_{emp} = \frac{1}{l} \sum_{i=1}^l L(\hat{f}(x_i), y_i)$$

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- **ERM**
- Errore

Due approcci semplici

Non conoscendo la distribuzione di probabilità quello che si fa è minimizzare il cosiddetto rischio empirico:

$$R_{emp} = \frac{1}{l} \sum_{i=1}^l L(\hat{f}(x_i), y_i)$$

In realtà a noi non interessa troppo la performance sul training, quindi si valuta la performance su un insieme di test non utilizzato per l'apprendimento.

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- **ERM**
- Errore

Due approcci semplici

Non conoscendo la distribuzione di probabilità quello che si fa è minimizzare il cosiddetto rischio empirico:

$$R_{emp} = \frac{1}{l} \sum_{i=1}^l L(\hat{f}(x_i), y_i)$$

In realtà a noi non interessa troppo la performance sul training, quindi si valuta la performance su un insieme di test non utilizzato per l'apprendimento.

Tipicamente si fanno delle ipotesi restrittive sul predittore utilizzato

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Non conoscendo la distribuzione di probabilità quello che si fa è minimizzare il cosiddetto rischio empirico:

$$R_{emp} = \frac{1}{l} \sum_{i=1}^l L(\hat{f}(x_i), y_i)$$

In realtà a noi non interessa troppo la performance sul training, quindi si valuta la performance su un insieme di test non utilizzato per l'apprendimento.

Tipicamente si fanno delle ipotesi restrittive sul predittore utilizzato

$$\text{NB: } Bias(\hat{f}(x)) = E(\hat{f}(x) - f(x)),$$

$$Var(\hat{f}(x)) = E \left[\left(\hat{f}(x) - E(\hat{f}(x)) \right)^2 \right]$$

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Un'assunzione ragionevole che si fa spesso in machine learning è quella che si abbia

$$Y = f(X) + \epsilon$$

in cui ϵ è un termine di errore random con media 0 indipendente da X

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Un'assunzione ragionevole che si fa spesso in machine learning è quella che si abbia

$$Y = f(X) + \epsilon$$

in cui ϵ è un termine di errore random con media 0 indipendente da X

Si addestra un predittore $\hat{f}(x)$ in modo da ottenere una stima \hat{Y} di cui si vuole stimare l'accuratezza

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Un'assunzione ragionevole che si fa spesso in machine learning è quella che si abbia

$$Y = f(X) + \epsilon$$

in cui ϵ è un termine di errore random con media 0 indipendente da X

Si addestra un predittore $\hat{f}(x)$ in modo da ottenere una stima \hat{Y} di cui si vuole stimare l'accuratezza

Ci sono due livelli di errore: **errore riducibile** che è quello legato a quanto bene \hat{f} approssima f , e quello **irriducibile** che deriva dalla presenza del termine di errore ϵ

Introduzione al corso

Esempi di applicazioni

Statistical Learning

- Statistical Learning
- Spam filtering
- Pericoli del ragionamento induttivo
- Ratti vs Piccioni
- Bias induttivo
- Quando serve il Machine Learning?
- Tipi di learning
- Supervisionato vs Non Supervisionato
- Attivo vs Passivo
- Utilità del docente
- Batch vs Online
- Un modello formale
- Generazione dei dati di training
- Misura di Successo
- ERM
- Errore

Due approcci semplici

Un'assunzione ragionevole che si fa spesso in machine learning è quella che si abbia

$$Y = f(X) + \epsilon$$

in cui ϵ è un termine di errore random con media 0 indipendente da X

Si addestra un predittore $\hat{f}(x)$ in modo da ottenere una stima \hat{Y} di cui si vuole stimare l'accuratezza

Ci sono due livelli di errore: **errore riducibile** che è quello legato a quanto bene \hat{f} approssima f , e quello **irriducibile** che deriva dalla presenza del termine di errore ϵ

Il termine di errore è per esempio legato al fatto che Y non dipende solo dalle variabili che stiamo misurando X o dal fatto che ci sono variazioni nei dati non misurabili

Minimi Quadrati Lineari

Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

● Minimi Quadrati
Lineari

- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Si fa l'ipotesi che il predittore sia lineare:

$$\hat{f}(x) = \hat{\beta}_0 + \sum_{j=1}^p x_j \hat{\beta}_j = x^T \hat{\beta}$$

Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

● Minimi Quadrati
Lineari

● Esempio

● k -NN

● Tradeoff tra bias e
varianza

● Complessità corretta
del modello

● No free Lunch

Theorem

Si fa l'ipotesi che il predittore sia lineare:

$$\hat{f}(x) = \hat{\beta}_0 + \sum_{j=1}^p x_j \hat{\beta}_j = x^T \hat{\beta}$$

Come si sceglie il parametro $\hat{\beta}$? Il metodo più popolare è quello di minimizzare la somma dei residui rispetto al training set:

$$RSS(\hat{\beta}) = \frac{1}{l} \sum_{i=1}^l (y^i - (x^i)^T \hat{\beta})^2.$$

Minimi Quadrati Lineari

Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

● Minimi Quadrati
Lineari

● Esempio

● k -NN

● Tradeoff tra bias e
varianza

● Complessità corretta
del modello

● No free Lunch

Theorem

Si fa l'ipotesi che il predittore sia lineare:

$$\hat{f}(x) = \hat{\beta}_0 + \sum_{j=1}^p x_j \hat{\beta}_j = x^T \hat{\beta}$$

Come si sceglie il parametro $\hat{\beta}$? Il metodo più popolare è quello di minimizzare la somma dei residui rispetto al training set:

$$RSS(\hat{\beta}) = \frac{1}{l} \sum_{i=1}^l (y^i - (x^i)^T \hat{\beta})^2.$$

Questo problema ammette sempre soluzione e la soluzione può essere calcolata in modo efficiente (ci torneremo)

Introduzione al corso

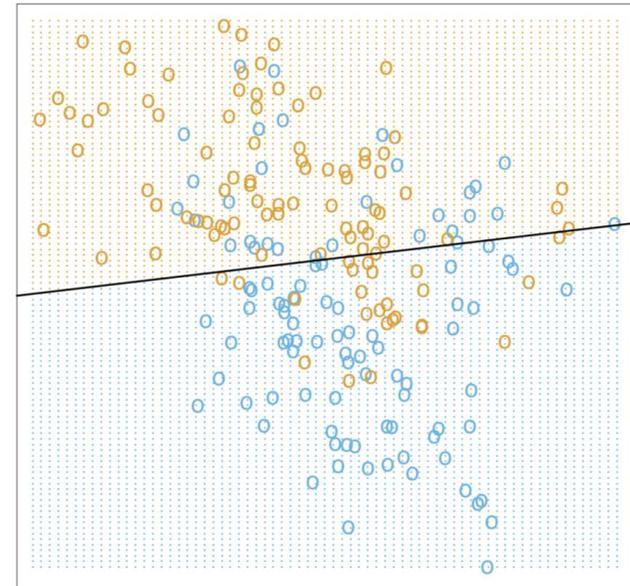
Esempi di applicazioni

Statistical Learning

Due approcci semplici

- Minimi Quadrati Lineari
- **Esempio**
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Consideriamo un problema di classificazione binaria come in figura.



Esempio

Introduzione al corso

Esempi di applicazioni

Statistical Learning

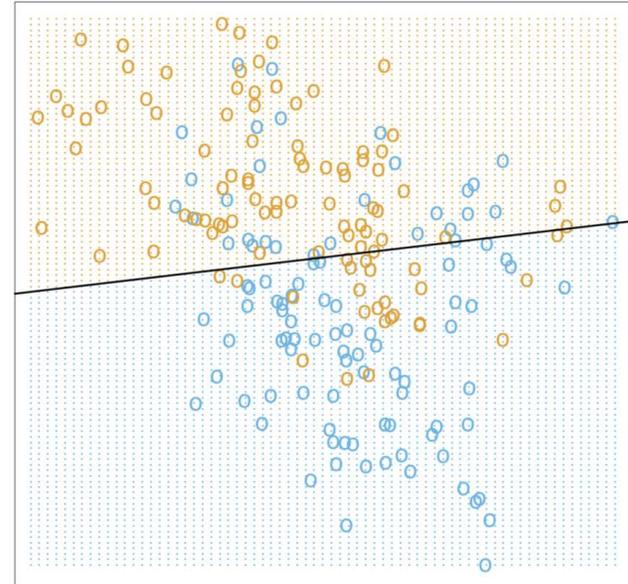
Due approcci semplici

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Consideriamo un problema di classificazione binaria come in figura.

I valori del classificatore lineare $f(x)$ vengono tradotti nella seguente classificazione:

$$g(x) = \begin{cases} 1 & \text{se } \hat{f}(x) \geq 0.5 \\ 0 & \text{se } \hat{f}(x) < 0.5 \end{cases}$$



Esempio

Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

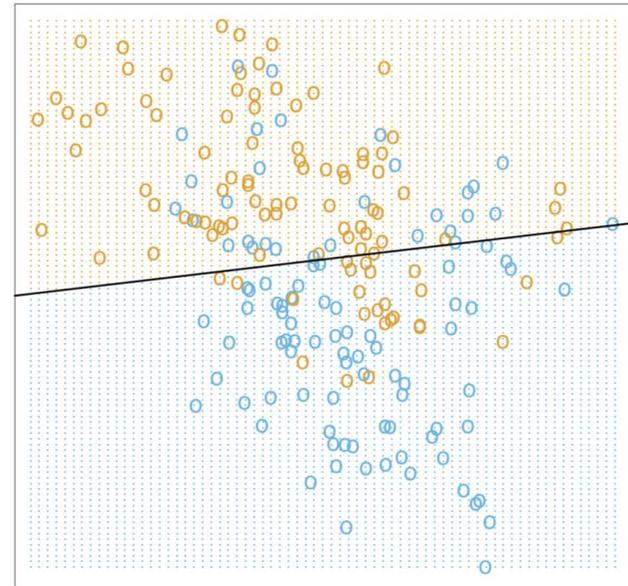
- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Consideriamo un problema di classificazione binaria come in figura.

I valori del classificatore lineare $f(x)$ vengono tradotti nella seguente classificazione:

$$g(x) = \begin{cases} 1 & \text{se } \hat{f}(x) \geq 0.5 \\ 0 & \text{se } \hat{f}(x) < 0.5 \end{cases}$$

Ci sono molti punti mal classificati in entrambe le classi: è perchè il classificatore è troppo rigido o non si potevano evitare?



Esempio

Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

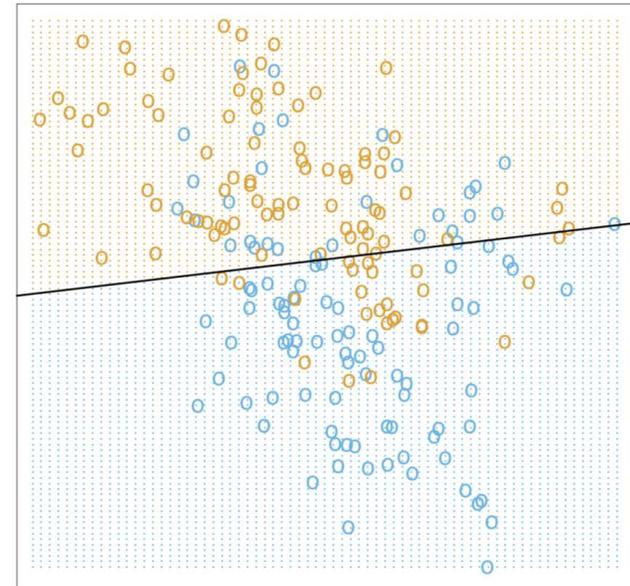
Consideriamo un problema di classificazione binaria come in figura.

I valori del classificatore lineare $f(x)$ vengono tradotti nella seguente classificazione:

$$g(x) = \begin{cases} 1 & \text{se } \hat{f}(x) \geq 0.5 \\ 0 & \text{se } \hat{f}(x) < 0.5 \end{cases}$$

Ci sono molti punti mal classificati in entrambe le classi: è perchè il classificatore è troppo rigido o non si potevano evitare?

Il predittore lineare è regolare (smooth) e stabile da individuare (varianza bassa e alto bias)



Introduzione al corso

Esempi di applicazioni

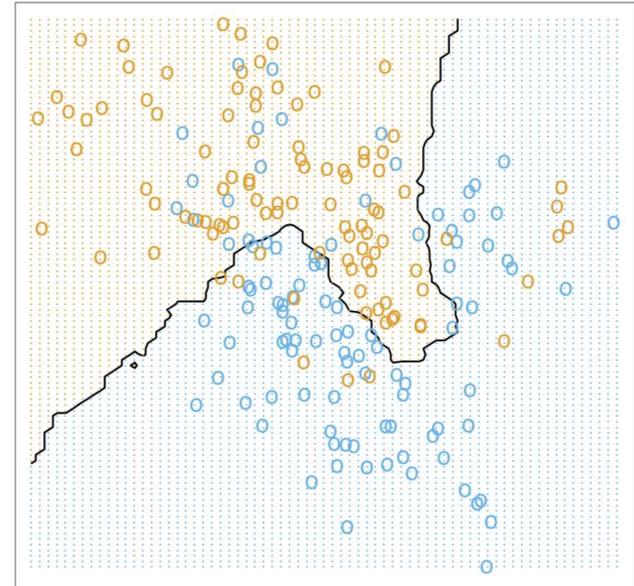
Statistical Learning

Due approcci semplici

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Il predittore prodotto da k -Nearest Neighbor è

$$\hat{f}(x) = \frac{1}{k} \sum_{x^i \in N_k(x)} y^i$$



Introduzione al corso

Esempi di applicazioni

Statistical Learning

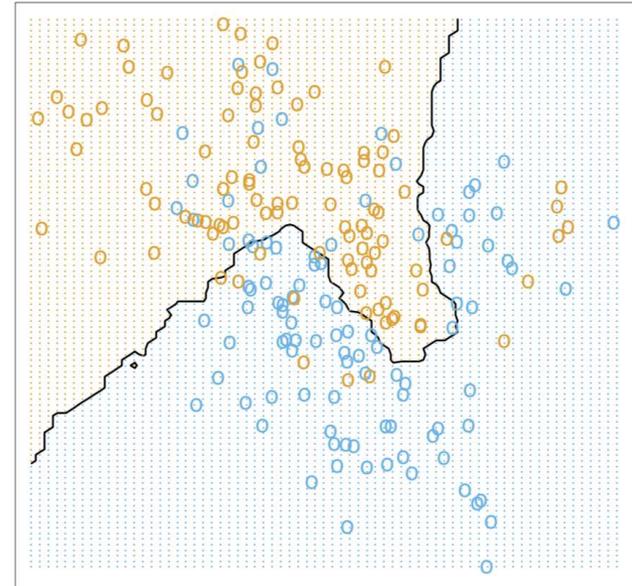
Due approcci semplici

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Il predittore prodotto da k -Nearest Neighbor è

$$\hat{f}(x) = \frac{1}{k} \sum_{x^i \in N_k(x)} y^i$$

Equivale a un voto di maggioranza tra i punti dell'intorno



Introduzione al corsoEsempi di applicazioniStatistical LearningDue approcci semplici

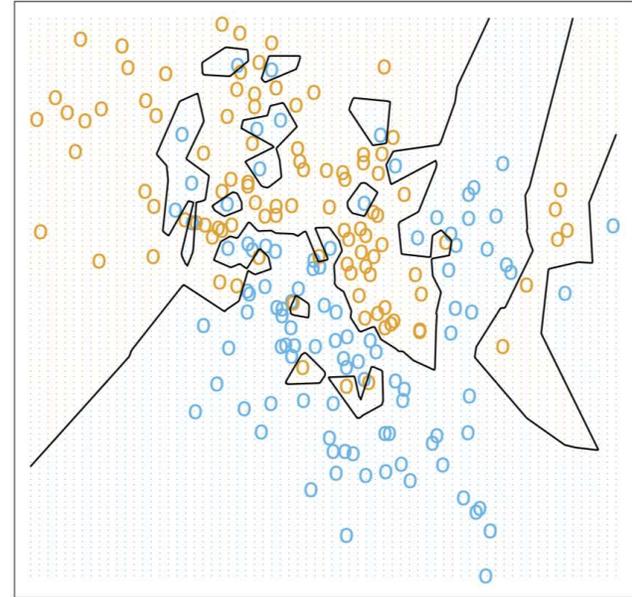
- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Il predittore prodotto da k -Nearest Neighbor è

$$\hat{f}(x) = \frac{1}{k} \sum_{x^i \in N_k(x)} y^i$$

Equivale a un voto di maggioranza tra i punti dell'intorno

Se prendo $k = 1$, ho errore nullo sul TS



Introduzione al corsoEsempi di applicazioniStatistical LearningDue approcci semplici

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

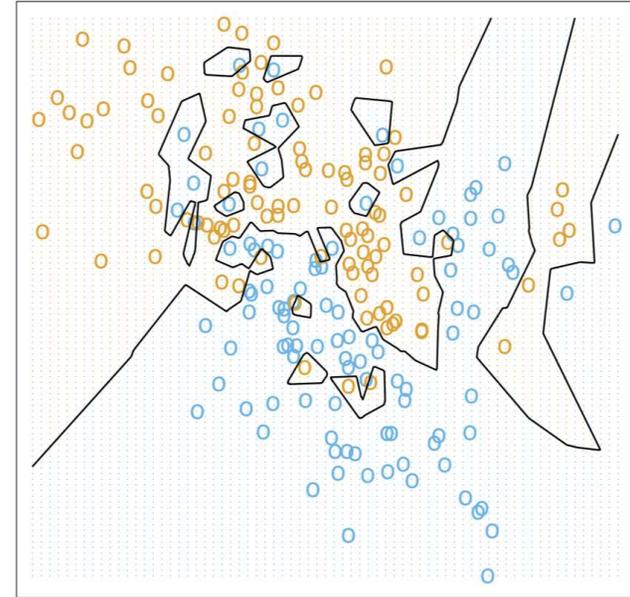
Il predittore prodotto da k -Nearest Neighbor è

$$\hat{f}(x) = \frac{1}{k} \sum_{x^i \in N_k(x)} y^i$$

Equivale a un voto di maggioranza tra i punti dell'intorno

Se prendo $k = 1$, ho errore nullo sul TS

In questo caso ho basso bias e alta varianza: ha alta adattabilità ma è instabile



Tradeoff tra bias e varianza

Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Si può dimostrare che il valore atteso dell'errore quadratico medio su un'istanza di test (x^0, y^0) mai vista nel training si può decomporre nel seguente modo:

$$E(y^0 - \hat{f}(x^0))^2 = \text{Var}(\hat{f}(x^0)) + (\text{Bias}[\hat{f}(x^0)])^2 + \text{Var}(\epsilon)$$

Tradeoff tra bias e varianza

Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Si può dimostrare che il valore atteso dell'errore quadratico medio su un'istanza di test (x^0, y^0) mai vista nel training si può decomporre nel seguente modo:

$$E(y^0 - \hat{f}(x^0))^2 = \text{Var}(\hat{f}(x^0)) + (\text{Bias}[\hat{f}(x^0)])^2 + \text{Var}(\epsilon)$$

Il valore medio dell'MSE si può calcolare facendo la media sugli esempi di test di $E(y^0 - f(x^0))^2$

Tradeoff tra bias e varianza

Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Si può dimostrare che il valore atteso dell'errore quadratico medio su un'istanza di test (x^0, y^0) mai vista nel training si può decomporre nel seguente modo:

$$E(y^0 - \hat{f}(x^0))^2 = \text{Var}(\hat{f}(x^0)) + (\text{Bias}[\hat{f}(x^0)])^2 + \text{Var}(\epsilon)$$

Il valore medio dell'MSE si può calcolare facendo la media sugli esempi di test di $E(y^0 - f(x^0))^2$

Per minimizzare il MSE serve un predittore che abbia bassa sia la varianza che il bias.

Tradeoff tra bias e varianza

Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

Si può dimostrare che il valore atteso dell'errore quadratico medio su un'istanza di test (x^0, y^0) mai vista nel training si può decomporre nel seguente modo:

$$E(y^0 - \hat{f}(x^0))^2 = \text{Var}(\hat{f}(x^0)) + (\text{Bias}[\hat{f}(x^0)])^2 + \text{Var}(\epsilon)$$

Il valore medio dell'MSE si può calcolare facendo la media sugli esempi di test di $E(y^0 - f(x^0))^2$

Per minimizzare il MSE serve un predittore che abbia bassa sia la varianza che il bias.

Per varianza del predittore si intende sostanzialmente quanto cambierebbe \hat{f} variando il training set, mentre per bias si intende l'errore che si commette approssimando un problema complesso tramite un modello più semplice

Complessità corretta del modello

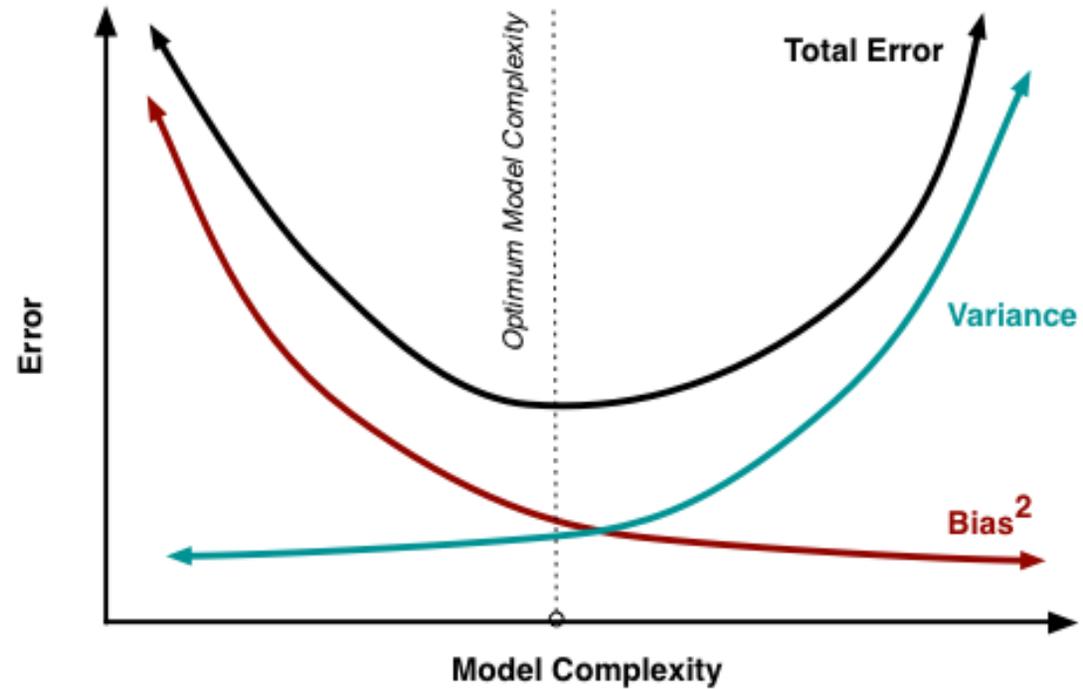
Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- **Complessità corretta del modello**
- No free Lunch Theorem





No free Lunch Theorem

Introduzione al corso

Esempi di applicazioni

Statistical Learning

Due approcci semplici

● Minimi Quadrati

Lineari

● Esempio

● k -NN

● Tradeoff tra bias e
varianza

● Complessità corretta
del modello

● No free Lunch

Theorem

Perchè devo definire classi di predittori diversi? Esiste un predittore ottimo per qualunque distribuzione dei dati?

No free Lunch Theorem

Perchè devo definire classi di predittori diversi? Esiste un predittore ottimo per qualunque distribuzione dei dati?

La risposta è **NO**: si dimostra formalmente che per qualunque predittore esiste una distribuzione dei dati per cui quel predittore performa male.

- Minimi Quadrati
- Lineari
- Esempio
 - k -NN
 - Tradeoff tra bias e varianza
 - Complessità corretta del modello
 - No free Lunch Theorem

No free Lunch Theorem

Perchè devo definire classi di predittori diversi? Esiste un predittore ottimo per qualunque distribuzione dei dati?

La risposta è **NO**: si dimostra formalmente che per qualunque predittore esiste una distribuzione dei dati per cui quel predittore performa male.

Non esiste quindi un modello che funziona meglio per ogni problema. Le ipotesi di un modello ottimo per un dato problema possono non valere per un altro, quindi è comune nell'apprendimento automatico provare più modelli e trovarne uno che funziona meglio per un particolare problema.

- Minimi Quadrati Lineari
- Esempio
- k -NN
- Tradeoff tra bias e varianza
- Complessità corretta del modello
- No free Lunch Theorem

No free Lunch Theorem

Perchè devo definire classi di predittori diversi? Esiste un predittore ottimo per qualunque distribuzione dei dati?

La risposta è **NO**: si dimostra formalmente che per qualunque predittore esiste una distribuzione dei dati per cui quel predittore performa male.

Non esiste quindi un modello che funziona meglio per ogni problema. Le ipotesi di un modello ottimo per un dato problema possono non valere per un altro, quindi è comune nell'apprendimento automatico provare più modelli e trovarne uno che funziona meglio per un particolare problema.

Questo è particolarmente vero nell'apprendimento supervisionato; la **cross validation** viene comunemente usata per valutare le precisioni predittive di modelli multipli di varia complessità per trovare il modello migliore.

- Minimi Quadrati
- Lineari
- Esempio
 - k -NN
 - Tradeoff tra bias e varianza
 - Complessità corretta del modello
 - No free Lunch Theorem

No free Lunch Theorem

Perchè devo definire classi di predittori diversi? Esiste un predittore ottimo per qualunque distribuzione dei dati?

La risposta è **NO**: si dimostra formalmente che per qualunque predittore esiste una distribuzione dei dati per cui quel predittore performa male.

Non esiste quindi un modello che funziona meglio per ogni problema. Le ipotesi di un modello ottimo per un dato problema possono non valere per un altro, quindi è comune nell'apprendimento automatico provare più modelli e trovarne uno che funziona meglio per un particolare problema.

Questo è particolarmente vero nell'apprendimento supervisionato; la **cross validation** viene comunemente usata per valutare le precisioni predittive di modelli multipli di varia complessità per trovare il modello migliore.

A seconda del problema è importante valutare i compromessi tra velocità, accuratezza e complessità dei diversi modelli e algoritmi e trovare il modello che funziona meglio per quel particolare problema.