Il ruolo dei modelli e l'approccio Big Data

Angelo **VULPIANI**

Dipartimento di Fisica, Università Sapienza, Roma,
<angelo.vulpiani@roma1.infn.it>

PERUGIA, Aprile 2023

Perché possiamo sperare di poter fare scienza?

È ben evidente la presenza di una certa regolarità nei fenomeni naturali

Questo appare già nella tradizione biblica:
Il vento soffia a mezzogiorno, poi gira a tramontana;
gira e rigira e sopra i suoi giri il vento ritorna ......
Ciò che è stato sarà e ciò che si è fatto si rifarà;
non c' è niente di nuovo sotto il sole.
(Ecclesiaste)

Possiamo congetturare che nel mondo esistono
"regole" per l'evoluzione dei fenomeni naturali
(i.e. LEGGI della NATURA)

Abbiamo qualche possibilità di capire il mondo che ci circonda.

I tre pilastri tradizionali delle scienze dure e un (ipotetico?) quarto paradigma

- Teoria, in generale uso della matematica.

- Osservazioni e, se possibile, Esperimenti.

- Simulazione numerica.

BIG DATA: verso un quarto paradigma?

Un crescente numero di persone crede che usando la grande quantità di dati a disposizione e una rete di computer, si possa capire (o almeno predire) su base puramente empirica, il comportamento di un qualunque fenomeno:
visto che siamo nell'era dei dati in abbondanza si può fare a meno delle teorie, basta usare i dati (parole del guru informatico C. Anderson).

Teorie, modelli, algoritmi e dati

Siamo veramente davanti ad una nuova rivoluzione scientifica?

Gli Algoritmi, i Big Data e l' Intelligenza Artificiale sono sempre un progresso?

• Il problema delle **Previsioni**:
un esempio concreto sui limiti dei Big Data e dell' Intelligenza Artificiale nella ricerca

• Big Data e Algoritmi nella vita di tutti i giorni

La presunta potenza dell' IA

Negli anni 80 dello scorso secolo alcuni ricercatori dell'Intelligenza Artificiale (IA) costruirono BACON, un programma al computer per realizzare scoperte scientifiche in modo automatico. Apparentemente BACON sembrava essere in grado di "scoprire" la terza legge di Keplero. Però ....

• BACON ha usato come dati di input i valori delle distanze dal Sole D e i periodi di rivoluzione dei pianeti P. Poi il programma ha scoperto la proporzionalità di D^3 con P^2.

• Nel caso di Keplero i dati osservativi non erano D e P, bensì le posizioni dei pianeti a diversi tempi visti dalla Terra, nella sua scoperta Keplero usò le "giuste" variabili D e P e venne guidato da forti convinzioni sull' armonia matematica dell' universo e sulla teoria di Copernico, a quel tempo ancora controversa.

Possiamo dire che l'IA è in grado di rimpiazzare il tradizionale approccio creatività per le scoperte scientifiche?

Tanti dettagli non sono necessariamente un fatto positivo

Borges nel breve racconto *Funes, o della memoria* scrive di un personaggio che, in seguito ad un incidente, ricordava tutto di tutto, sin nei minimi dettagli della più comune delle situazioni. Questo, ben lungi dall'essere un fatto positivo, comportava la quasi incapacità di un pensiero astratto. Funes era infastidito che un cane visto di profilo alle 3:14 fosse lo stesso visto di fronte alle 3:15 e era quasi incapace di idee generali platoniche.

In un problema scientifico la prima cosa (forse la più difficile e importante) da fare è identificare la parte significativa del fenomeno, solo così si ha qualche speranza di capire
Per pensare in modo corretto in primo luogo si deve capire cosa buttar via. Per andare avanti devi sapere cosa lasciar fuori, questa è l' essenza del pensare in modo efficace
(Kurt Gödel).

Galileo GALILEI (1564- 1642)

Per capire il mondo intorno a noi dobbiamo usare la Matematica e

DIFALCARE GLI IMPEDIMENTI

cioè concentrarsi su problemi ("semplici") dopo averne identificato gli aspetti essenziali.

Anche le "teorie paludate" (ad es. elettromagnetismo o meccanica quantistica) in fondo non sono altro che modelli di lusso.

Ci occuperemo dei "veri" modelli, che non pretendono di fornire una descrizione generale, in particolare nei casi in cui manca una "teoria paludata".

Prima cosa da capire:

Quali sono le variabili rilevanti?

Questo è un problema MOLTO difficile; ricordiamo che è stato necessario aspettare Galileo e Newton per capire che in meccanica le variabili "giuste" sono la posizione e la velocità.

Seconda cosa da capire:

Che tipo di legge regola il sistema?

Il sistema è deterministico?

Cioè dato lo stato iniziale $\mathbf{x}(0)$ il futuro è determinato; questo accade in meccanica.

Oppure c'è una regola probabilistica?

Diverse situazioni:
A: Le leggi di evoluzione esistono e si conoscono (elettromagnetismo, meccanica).
B: Le leggi di evoluzione esistono e non si conoscono (terremoti).
C: Non sappiamo se le leggi di evoluzione esistono (finanza, fenomeni sociali).

IN OGNI CASO È INEVITABILE FAR RICORSO A MODELLI

Le principali tipologie in cui possono essere classificati i modelli:
I- Modelli molto semplificati che danno risposte giuste ad alcuni aspetti del problema considerato;
II- Modelli per analogia;
III- Modelli a grande scala;
IV- Modelli dai dati.

Modelli per tutti i gusti, o quasi...

Nella classe **I** c'è modello di Lorenz che nasce da una semplificazione (brutale) delle equazioni della fluidodinamica, e ha permesso di capire che il moto irregolare può essere dovuto al caos deterministico che è presente anche in sistemi a bassa dimensionalità.

Il modello di Lotka-Volterra è nella classe **II** e ha fatto scuola in biomatematica.

Nei modelli della classe **III** troviamo le equazioni efficaci usate in meteorologia e ingegneria, in cui vengono prese in considerazioni solo le "variabili rilevanti" del problema.

Nell classe **IV** i problemi forse più interessanti e certamente più difficili, in cui si cerca di costruire equazioni a partire dai dati senza avere una struttura teorica di riferimento.

Il modello di Lotka-Volterra

Dal suo futuro genero Umberto D'Ancona, Vito Volterra venne a conoscenza di alcuni dati della pesca nell'Adriatico tra il 1903 e il 1923, che mostravano strano comportamenti, in particolare il fatto, apparentemente paradossale, che la riduzione dell' attività di pesca durante la guerra aveva portato ad una diminuzione di certi tipi di pesci.

Volterra semplifica drasticamente il problema descrivendo il sistema ecologico con due sole variabili: $x =$ quantità di pesci erbivori che si cibano di alghe e $y =$ quantità di pesci predatori che mangiano i pesci erbivori.

Per costruire il modello Volterra ragiona per analogia con la teoria cinetica: *il pesce grosso "collide" con il pesce piccolo e con una certa probabilità lo mangia*. Giunge così a identificare due equazioni differenziali (già trovate in un contesto diverso da Alfred Lotka) come modello capace di rendere conto delle osservazioni di D'Ancona.

Le equazioni che governano la dinamica delle due popolazioni sono le seguenti:

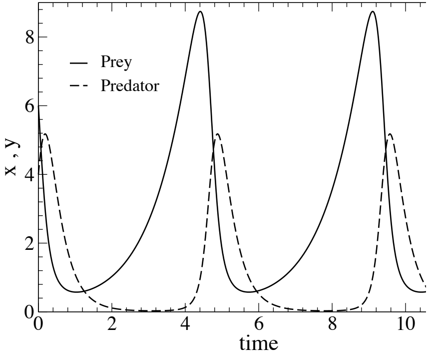
$$\frac{dx}{dt} = ax - bxy \ , \ \frac{dy}{dt} = -cx + dxy \ ,$$

le costanti a, b, c e d sono positive.

I termini lineari non hanno bisogno di molte spiegazioni: assumendo risorse alimentari illimitate, l'assenza di predatori porta alla proliferazione esponenziale delle prede; analogamente, in assenza di prede, i predatori si estinguono.

Notare la struttura (bilineare) simile a quella dell' equazione di Boltzmann nella teoria cinetica dei gas, non è una coincidenza: i termini non lineari sono stati introdotti da Volterra notando l'analogia tra l'interazione preda/predatore e la collisione tra due atomi.

L' andamento periodico nel tempo di $x(t)$ e $y(t)$ nel modello di Lotka-Volterra, $a = 1$, $b = 1$, $c = 3$ e $d = 1$.



L'importanza del modello di Lotka-Volterra

• In ecologia non c'è niente di simile alle equazioni di Newton per la meccanica, o le equazioni di Maxwell per l' elettromagnetismo, e quindi la costruzione di modelli non può basarsi sui principi primi di teorie consolidate.

• Il modello di Lotka-Volterra è in grado di spiegare l'andamento qualitativo dei dati, la periodicità, e l'effetto controintoitivo del ruolo della pesca.

• Rilevanza degli aspetti qualitativi, che non devono essere intesi come di livello inferiore rispetto a quello quantutativi.

E se non bastasse spesso abbiamo anche il Caos

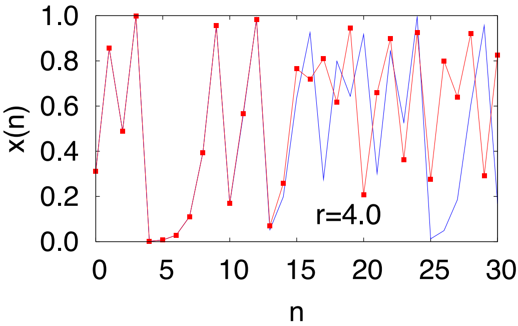
Se pure accadesse che le leggi della natura non avessero più alcun segreto per noi, anche in questo caso potremmo conoscere la situazione iniziale solo approssimativamente .... può accadere che piccole differenze nelle condizioni iniziali ne producano di grandissime nei fenomeni finali. Un piccolo errore nelle prime produce un errore enorme nei secondi. La previsione diventa impossibile e si ha un fenomeno fortuito.
(H. Poincaré)

Ci sono sistemi deterministici che, nonostante la natura che sembrerebbe suggerire un comportamento regolare, mostrano un' evoluzione temporale piuttosto irregolare (come ci si aspetta nei processi stocastici).
Può esserci una forte dipendenza dalla condizione iniziale: piccole differenze al tempo iniziali vengono amplificate in modo esponenziale.
Questo è l'effetto farfalla: una farfalla che batte le ali in Brasile può provocare un tornado in Texas dopo un paio di settimane.

Il famigerato effetto farfalla

Un modello molto semplice: la mappa logistica

$$x(n+1) = 4x(n)(1 - x(n)),$$

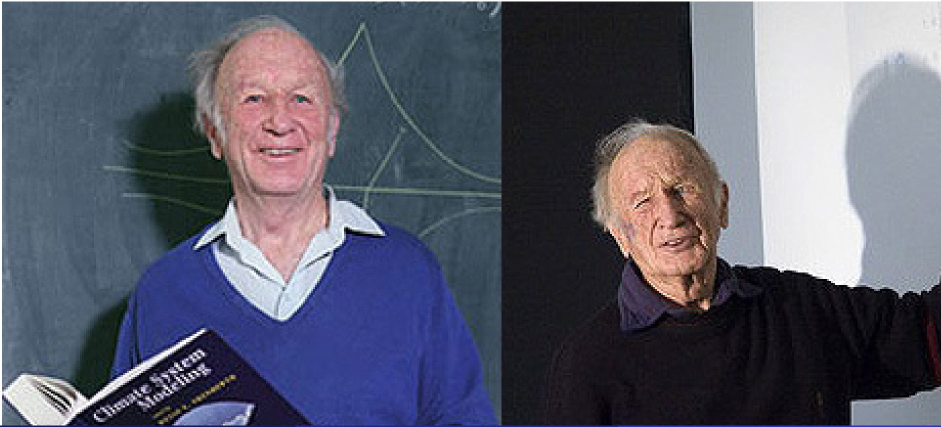


Due traiettorie con condizioni iniziali molto vicine
$|x(0) - x'(0)| = 4 \times 10^{-6}$.
Dopo un tempo breve le due traiettorie sono completamente diverse.

Edward LORENZ (1917-2008)

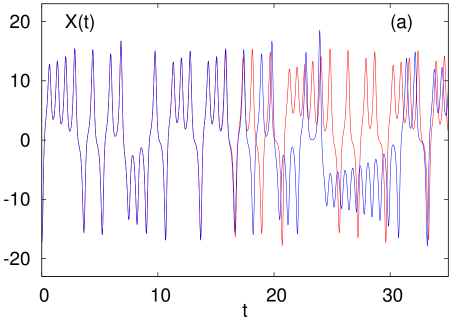
Nel 1963 Lorenz introdusse il suo famoso modello **con solo 3 variabili**, per la circolazione atmosferica, e (ri)scroprì il caos.

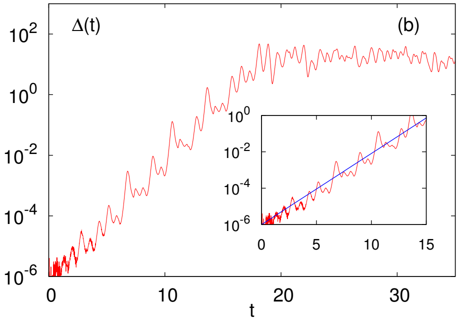
$$\frac{dx}{dt} = \sigma(y - x)\,, \quad \frac{dy}{dt} = -xz + rx - y\,, \quad \frac{dz}{dt} = -bz + xy$$

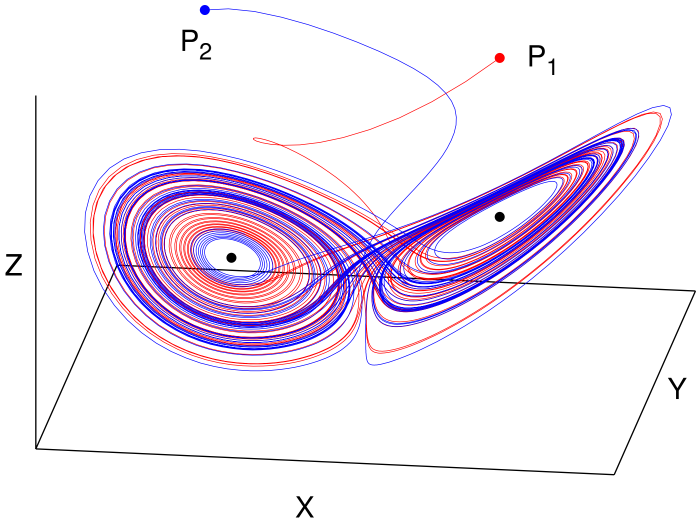


Effetto farfalla nel modello di Lorenz

Due traiettorie con condizioni iniziali molto vicine si allontanano molto velocemente: non c'è speranza di fare previsioni meteo a tempi lunghi... (questo anche se il modello di Lorenz descrivesse esattamente l'atmosfera).






P2
P1
Z
Y
X

Cosa abbiamo capito dal modello di Lorenz

*** Il determinismo non implica la possibiltà di fare previsioni accurate.**
*** Sistemi "semplici" e con poche variabili possono essere caotici.**
Due traiettorie con condizioni iniziali molto vicine si allontanano molto velocemente: non c'è speranza di fare previsioni meteo a tempi lunghi. Questo è quello che viene chamato **effetto farfalla**.
*** Ci può essere "complessità" anche in un sistema con solo 3 variabili.**

Le equazioni del modello di Lorenz non sono state inventate, bensì ricavate, anche se con un troncamento molto brutale, dalle equazioni della fluidodinamica.

Un problema molto ambizioso: Come costruire un modello?

Purtroppo non tutti abbiamo le capacità di astrazione e l'intuizione di Volterra. E non sempre abbiamo una teoria generale di riferimento dalla quale, come nel caso di Lorenz, ricavare un modello in grado di descrivere almeno qualche aspetto importante.

Per questo sarebbe auspicabile identificare un metodo generale che ci permetta di determinare le equazioni che regolano il fenomeno di interesse. In questo senso la ricerca di metodi algoritmici per la generazione di "teorie" a partire dai dati è perfettamente comprensibile.

È possibile realizzare questo sogno ultrainduttivista?

Cosa fare se non abbiamo una teoria di riferimento?

In prima approssimazione possiamo distinguere due situazioni con difficoltà crescente:

I - si conoscono le "variabili giuste" $\mathbf{X}$ ma non la struttura delle equazioni di evoluzione e abbiamo a disposizione solo una (lunga) serie storica $\{\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_M\}$ (per semplictà ci limitiamo al caso con tempo discreto);

II - non si conoscono le "variabili giuste" e si hanno solo delle serie storiche di una qualche variabile $\{u_1, u_2, ..., u_M\}$.

Verso una formalizzazione della saggezza biblica

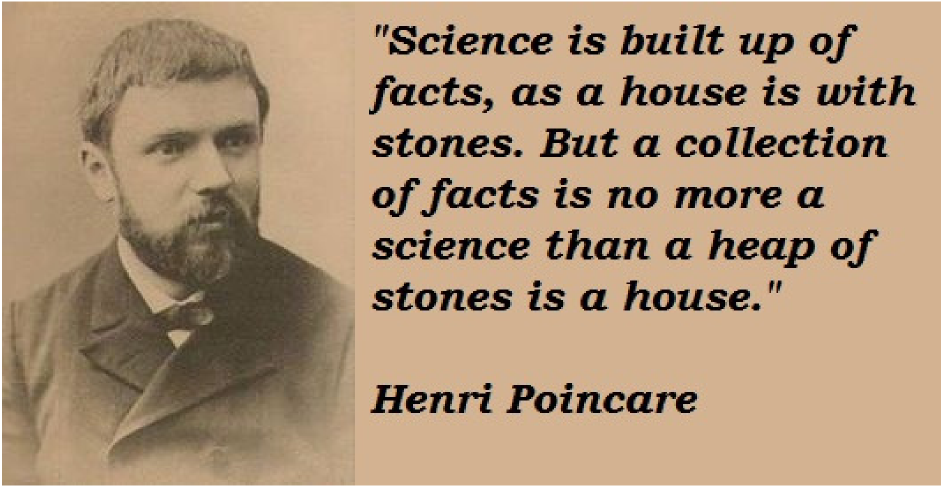
Il metodo degli analoghi

* sappiamo che lo stato del sistema è dato da un vettore $\mathbf{x}$
* conosciamo il passato del sistema, cioè una serie storica $\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_M\}$ ove $\mathbf{x}_j = \mathbf{x}(j\Delta t)$
* vogliamo predire il futuro, cioè $\mathbf{x}_{M+t}$ for $t > 0$.

Se il sistema è deterministico, per capire il futuro basta guardare al passato e cercare un "analogo", cioè un vettore $\mathbf{x}_k$ with $k < M$ tale che $|\mathbf{x}_k - \mathbf{x}_M| < \epsilon$, quindi, poiché "dallo stesso antecedente segue lo stesso conseguente", possiamo predire il futuro al tempo $M + t > M$:

$$\mathbf{x}_{M+t} \simeq \mathbf{x}_{k+t}$$

Già Poincaré aveva idee chiare a riguardo il reale impatto dei dati nella scienza



Il sogno (duro a morire) di una scienza solo induttiva

Vale la pena discutere il caso delle previsioni meteo per mostrare chiaramente come, per un problema non banale, sia decisamente troppo ottimistico puntare solo sull' uso dei dati osservativi, ma sia necessaria una combinazione di tecniche matematiche, intuizione fisica e sviluppo tecnologico.
Assumiamo (cosa non sempre vera) di sapere che il fenomeno che vogliamo studiare sia descritto da un set di variabili $\mathbf{X}(t)$ con evoluzione deterministica. Per fare una previsione del futuro si potrebbe pensare di cercare nel passato una situazione "vicina" a quella di oggi, se la si trova al giorno 25 gennaio 1923 allora è sensato assumere che domani il sistema sarà "vicino" a quello del 26 gennaio 1923, e quindi sembrerebbe che abbiamo un metodo empirico di previsione.

Un po' di teoria aiuta a capire cosa (non) si può fare

Sembrerebbe tutto facile, in particolare ora che siamo nell' era dei Big Data e, quindi, potremmo non perdere tempo con la teoria. Per prima cosa chiediamoci se sia sempre possibile individuare un analogo, cioè un giorno nel passato in cui il sistema è "vicino" ad oggi. Da un punto di vista matematico il problema è strettamente collegato ad un risultato classico della fine del diciannovesimo secolo (**il teorema di ricorrenza di Poincaré**):
In un sistema deterministico in uno spazio delle fasi limitato, il sistema torna (quasi) sempre vicino alla condizione iniziale.
Da questo si conclude che un analogo sicuramente deve esistere.

C' è però un problema pratico:

Quanto indietro si deve andare per trovarlo?

La maledizione esponenziale

La risposta è nel **lemma di Kac** un risultato della teoria dell' ergodicità. La difficoltà di trovare un analogo dipende dalla dimensione D (in parole povere D è il numero minimo di variabili necessarie per descrivere il problema):
Per trovare un analogo con precisione percentuale ϵ si deve andare indietro di un tempo T_m

$$T_m \sim \left(\frac{1}{\epsilon}\right)^D .$$

Se D è grande (diciamo oltre 6), già per precisioni non enormi serve una serie temporale di lunghezza gigantesca, ad esempio se vogliamo una precisione 5% in 10 dimensioni la lunghezza deve essere almeno 10^{13}.
Quindi la limitata lunghezza delle serie dei Big Data, per quanto grandi, in situazioni non banali, non permette di usare per le previsioni un approccio puramente induttivo e senza teoria.

Lewis Fry Richardson (1881- 1953), pacifista coerente, grande scienziato visionario, ingiustamente poco noto

Previsioni al giorno d'oggi

Nella realtà la situazione è quasi sempre molto più complicata di quella sopra descritta, infatti tipicamente non si conoscono nemmeno le "variabili giuste", e molto spesso non sappiamo neanche se il sistema evolve con regole deterministiche o stocastiche.

L' idea di base dell' approccio attualmente usato, proposto negli anni 20 da **L.F. Richardson**, che aveva capito come l'approccio in termini di analoghi era destinato a fallire, è il seguente:
l'atmosfera evolve in accordo con le equazioni dell' idrodinamica e la termodinamica. Quindi dalla conoscenza dello stato presente dell'atmosfera, risolvendo un sistema di equazioni alle derivate parziali, si può (almeno in linea di principio) effettuare una previsione del tempo. Ovviamente le equazioni in questione possono essere risolte solo numericamente.

Per la realizzazione del progetto visionario di Richardson si è aspettato fino agli anni 50 con lo sviluppo di tre "ingredienti" assolutamente non banali:
a) la messa a punto di equazioni efficaci;
b) algoritmi numerici veloci;
c) computer per i calcoli numerici.

I punti b) e c) non hanno bisogno di particolari commenti, il punto a) è l' aspetto concettualmente importante per capire la necessità di un uso non banale della matematica per descrivere un fenomeno complesso.
Negli anni 50 Von Neuman ed i suoi collaboratori notarono che le equazioni originariamente proposte da Richardson, benchè corrette, non sono adatte per le previsioni meteo. Il motivo, apparentemente paradossale, è che sono **troppo accurate**, infatti descrivono anche moti ondosi ad alta frequenza che sono irrilevanti in ambito meteorologico e difficili da trattare numericamente.
È quindi necessario costruire **equazioni efficaci** in cui non compaiano le variabili veloci, quelle relative alle alte frequenze.

Verso una previsione dei comportamenti sociali?

L' esempio dell' Economia

Per molto tempo la formulazione matematica dell'economia è stata basata sull' assunzione di comportamenti regolari, e la convinzione di poterli controllare modificando i parametri di controllo, cioè agendo sui tassi di interesse, svalutazione, tasse etc. Sostanzialmente è la stessa idea di von Neumann che, negli anni 50, credeva di poter controllare l'atmosfera. Ma l'economia e la finanza non mostrano affatto andamenti regolari, i politici (e i loro consulenti strapagati) non sono in grado di controllare molto.

Perché, in assenza di una teoria accettabile, non usare i Big Data?

Problemi veramente difficili

Se provassimo a generalizzare e sistematizzare l'intuizione di Richardson ci confronteremmo con problemi prevedibili e probabilmente insormontabili. È difficile pensare a un metodo generale per la scelta delle variabili "giuste". Questo è un aspetto molto delicato (purtroppo spesso trascurato), ben chiaro nell'ambito della meccanica statistica,

S.K. Ma è molto esplicito :
the hidden worry of thermodynamics is: we do not know how many coordinates or forces are necessary to completely specify an equilibrium state.

Analogamente Onsager e Machlup per i problemi di non equilibrio:
how do you know you have taken enough variables, for it to be Markovian?
(Traduzione: *come sai che le variabili usate sono quelle "opportune"?*)

Una digressione sulla scienza nel periodo vittoriano: le previsioni delle maree

Già nella prima metà del 19-mo secolo esistevano metodi empirici per compilare tavole numeriche per le previsioni delle maree in un dato porto.

Lord Kelvin e George Darwin (figlio del famoso naturalista Charles) mostrarono che il livello dell'acqua poteva essere predetto con pochi termini della serie di Fourier (diciamo 10 o 20). I coefficienti della serie dipendono dal porto e sono determinati dalle serie storiche del porto di interesse.

Kelvin e il fratello (un ingegnere meccanico) costruirono una sorta di computer dedicato, uno dei primi esempi di business scientifico.

Un esempio della macchina per predire le maree, un computer analogico di oltre una tonnellata

Lord Kelvin and George Darwin erano molto bravi, ma pure fortunati...

Anche se le maree sono caotiche, la loro previsione dal passato è relativamente facile.
Il motivo principale, capito negli anni 90 del 20-mo secolo usando una tecnica piuttosto elaborata (dovuta a Takens), è che la dimensione è piccola $O(3-4)$.
Questo spiega a posteriori il successo del metodo empirico.

Al contrario per l'atmosfera non c' è speranza perché D è molto grande $O(10^3-10^4)$.

Un primo tentativo di conclusioni

• Bisogna conoscere bene sia la fenomenologia che la teoria di riferimento, nonché la matematica e le tecniche computazionali; e in più bisogna avere occhio, un'intuizione che viene solo dall' esperienza.

• Analizzando specifici problemi (come quello delle previsioni) con un opportuno spirito critico si capisce che, pur costituendo un'interessante sfida scientificamente e di elevato impatto potenziale, le recenti tecniche per la raccolta e l'analisi (algoritmica) di grandi quantità di dati non sono una panacea.

• Meglio evitare di cedere a facili entusiasmi, in particolare alla tentazione di pensare che la soluzione a tutti i problemi, dai grandi progetti scientifici alla diagnosi medica, dipenda prevalentemente dallo sviluppo di una qualche tecnologia.

Algoritmici oggettivi o pregiudizi?

L'esempio della polizia Chicago:
Dividere la città in quadrati (ciascuno di circa 2 ettari) e fare controlli in base alle informazioni della banca dati, la frequenza dei controlli è proporzionale al numero dei crimini verificatisi negli ultimi anni.

Risultato
Nelle zone verificate più frequentemente venne riscontrato un maggior numero di crimini. Sembrerebbe che l'algoritmo abbia fatto un ottimo lavoro.

Ma l'algoritmo non è affatto neutro.
Che tipo di crimini erano stati presi in considerazione?
Quelli tipici dei quartieri abitati da afroamericani e ispanici (spaccio, scippo, atti di vandalismo) ma non quelli dei colletti bianchi (illeciti di natura fiscale etc).

Con gli algoritmi l' imparzialità non è cosa facile

Nella contea di Broward, in Florida, un software aiuta a decidere se una persona accusata di un reato debba essere rilasciata su cauzione prima del processo. Per gli individui di colore l' algoritmo prediceva un numero sproporzionato di falsi positivi, individui classificati ad alto rischio che però successivamente non commettevano un altro reato.
La società sviluppatrice del software sosteneva che lo strumento non aveva pregiudizi, e classificava gli individui ad alto rischio, sia bianchi che di colore, con la stessa accuratezza.
Purtroppo è possibile dimostrare che esistono almeno 20 plausibili definizioni di imparzialità, che in molti casi sono mutualmente esclusive.

È certamente vero che gli algoritmi istruiti dai dati raggiungono una tale complessità che sfugge ai creatori. In alcuni casi se venissero aperte le black box non capiremmo nulla.
(Fosca Giannotti, Sole 24 Ore, 21 luglio 2019, pagina 12)

Le banalità, o peggio, che si trovano sui quotidiani...

Il 14 agosto 2019 Repubblica dedica un'intera pagina (la 25) alla vincita di 209 milioni di euro al superenalotto: **Il CNR ha calcolato che c' è una possibilità su oltre 622 milioni di indovinare la sestina giusta.**
In verità questo risultato non è altro che un'applicazione elementare del calcolo combinatorio, cosa che uno studente di liceo scientifico dovrebbe sapere.
In una colonna dal titolo **IL TRIONFO DEL CASO CHE UMILIA LA CABALA**, **Marino Niola** (un antropologo), discute, con grande sorpresa, come un algoritmo avrebbe fatto vincere il fortunato scommettitore, che aveva giocato una schedina precompilata.
Questo algoritmo semplicemente genera numeri a caso, la versione digitale della tombola di Natale: si prende un sacchetto con 90 numeri, se ne tirano fuori 6. La strategia è buona, molto di più: è l'unica che ha senso in quanto le estrazioni sono indipendenti, quindi NON ci può essere alcuna strategia basata sulle estrazioni precedenti.

Un tentativo di sommario

• Uno degli aspetti più importanti della Scienza è capire cosa si può fare e cosa no.

• I metodo basati su Big Data e Intelligenza Artificiale, pur utili, non hanno ancora mostrato capacità predittive paragonabili a quelle dei metodi tradizionali.

• Nessuna persona sana di mente mangerebbe un panino trovato su una panchina, allora perché fidarsi di un algoritmo senza saperne qualcosa?

• Un po' di buona matematica classica (quella nel 19-mo secolo per intendersi) è sempre utile.

• **Una Perla di Saggezza** *Sapere che sai quando sai, e che non sai quando non sai, questa è la vera conoscenza* (**Confucio**).

Qualche Referenza

* C. Anderson *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete* http://www.wired.com/2008/06/pb-theory/
* G. Boffetta e A. Vulpiani *Probabilità in Fisica* (Springer Italia, 2012)
* S. Chibbaro, L. Rondoni and A. Vulpiani *Reductionism, Emergence and Levels of Reality* (Springer–Verlag, 2014)
* I. Ekeland *Come funziona il caos* (Bollati Boringhieri, 2010)
* L. Gammaitoni e A. Vulpiani *Perché è difficile prevedere il futuro* (Dedalo, 2019)
* A. Guerraggio 15 *grandi idee matematiche* (Bruno Mondadori, 2013)
* T. Hey, S. Tansley and K. Tolle *The Fourth Paradigm: Data-Intensive Scientific Discovery* (Microsoft Research: Redmond, WA, USA, 2009)
* H. Hosni and A. Vulpiani *Forecasting in Light of Big Data.* Philos. Technol. 2017, **31**, 557-569 (2018)

* M. Kac *Gli enigmi del caso* (Bollati Boringhieri, 1996)
* P. Lynch *The Emergence of Numerical Weather Prediction: Richardson's Dream* (Cambridge University Press, 2006)
* E. Noël (Editore) *Aggiornamenti sull'idea di "caso"* (Bollati Boringhieri, 1992)
* H. Poincaré *Geometria e caso* (Bollati Boringhieri, 1995)
* K. Pomian (Editore) *Sul determinismo* (Il Saggiatore, 1991)
* C. O'Neil *Armi di distruzione matematica* (Bompiani, 2016)
* D. Ruelle *Caso e caos* (Bollati Boringhieri, 1992)
* A. Vespignani *L'algoritmo e l'oracolo* (Il Saggiatore, 2019)
* A. Vulpiani *Determinismo e caos* (Nuova Italia Scientifica 1994, ristampa Carocci 2004)
* A. Vulpiani *Caso, probabilità e complessità* (Ediesse, 2014)