

Sommario

<u>Premessa.</u>	5
<u>Introduzione.</u>	13
<u>Trading intelligente.</u>	22
<u>Alle prese con R..</u>	34
<u>Machine learning.</u>	43
<u>Tipi di machine learning.</u>	49
<u>Base dati</u>	54
<u>Decision trees.</u>	65
<u>Utilizzo del modello..</u>	91
<u>Random forests.</u>	100
<u>Neural nets.</u>	108
<u>S.V.M. - Support vector machines.</u>	119
<u>Knn – k-nearest neighbors.</u>	128
<u>Valutazione grafica.</u>	133
<u>Ensemble.</u>	143
<u>Integrazione.</u>	148
<u>Un pizzico di follia.</u>	155
<u>Caret.</u>	165
<u>Labels multiple.</u>	168
<u>Conclusioni</u>	177
<u>Bibliografia.</u>	181

Premessa

La genesi di questo libro risale ad oltre venti anni fa. Una prima stesura, evidentemente al netto delle evoluzioni intervenute nel frattempo, è stata realizzata nel 1997. All'epoca la considerai piuttosto ingenua e poco credibile. Mi astenni perciò dal cercare un editore o, comunque, dal diffonderla sotto qualsiasi forma. In effetti, non essendo io un matematico o un professionista dell'intelligenza artificiale, mi è sembrato poco opportuno presentare delle conclusioni non giustificate da specifiche competenze sull'argomento.

Ho sbagliato. Già allora, avendo manifestato attraverso la rete internet il mio interesse a questo settore di ricerca, fui invitato a partecipare a delle conferenze internazionali qualificate^[1], nel corso delle quali ho ricevuto degli apprezzamenti veramente inaspettati. Evidentemente sono stato premiato da un atteggiamento corretto. E' stato sufficiente chiarire la mia posizione nel contesto. Non mi sono presentato da esperto, per insegnare qualcosa a dei veri esperti, dai quali avevo tutto da imparare. Mi sono presentato come un apprendista stregone, desideroso di applicare al proprio campo di interesse, cioè al trading di azioni e derivati, le competenze di accademici molto preparati, ma inclini soprattutto allo studio di realizzazioni astratte in un mondo che, allora, si interrogava sulle prospettive di un impiego pratico dell'intelligenza artificiale.

Insomma, ho formulato delle ipotesi di utilizzo di strumenti dei quali conoscevo solo superficialmente il funzionamento. La classica metafora che si potrebbe evocare è quella di chi guida una macchina senza aver mai aperto il cofano del motore. Metafora che non è tale, almeno per quanto mi riguarda. Infatti guido, e credo anche bene, senza sapere assolutamente nulla di congegni meccanici ed ingranaggi ai quali altri, invece, si interessano con passione. E così, tornando alla questione delle conferenze, ho incuriosito chi mi ascoltava, perché parlare di soldi incuriosisce sempre.

Erano ovviamente già in circolazione alcuni studi sulle supposte potenzialità dell'intelligenza artificiale in ambito finanziario. Tuttavia, senza eccezione alcuna, si trattava di elaborazioni matematiche teoriche, nelle quali spiccava l'assenza di una effettiva conoscenza della pratica del trading. Un mondo, quello accademico, strizzava l'occhio all'altro, l'allora mondo dei "borsini" affiancato dal nascente "trading on line", senza che nella maggioranza dei casi, sono pronto a scommettere, fosse stato investito un solo centesimo in operazioni reali su indicazione di modelli "artificialmente intelligenti". E' vero che io, da trader, pretendevo di parlare di intelligenza artificiale senza conoscerla compiutamente; però è altrettanto vero che, specularmente, tanti teorici di intelligenza artificiale tentavano di immergersi nel mondo del trading senza conoscerne il funzionamento; nel migliore dei casi, armati della sola esperienza dell'acquisto e della vendita di qualche azione passando dal funzionario bancario dell'ufficio titoli.

Tuttavia, benché fossimo alla pari, mi sentivo un po' in soggezione di fronte ad argomentazioni che facevano uso di paroloni altisonanti e di formule matematiche fuori della

mia portata. Fu questa situazione di diffusa ambiguità, animata da abili ricercatori dalle grandi aspettative, che mi trattenne dal sembrare presuntuoso con un testo che coniugasse competenze del mondo del trading, nel quale mi muovevo con sufficiente disinvoltura, con altre che non mi appartenevano.

Una seconda ragione, forse più sentita della precedente, mi ha trattenuto dal dare alle stampe il frutto del mio lavoro. Perché mai, mi si sarebbe potuto chiedere allora e mi si potrebbe chiedere tuttora, ti cimenti nella stesura di un libro, sprecando tempo e fatica, quando potresti andare bellamente a goderti nei mari tropicali i miliardi guadagnati con i tuoi sistemi di intelligenza artificiale? La risposta è davvero molto semplice ma, in un'epoca non ancora adeguatamente matura in questo campo, quando si guardava all'intelligenza artificiale come ad un qualcosa di magico, sarebbe apparsa certamente poco convincente. Oggi, spero, verrà capita meglio. E del resto, avendo adesso preso la decisione di attuare il mio progetto editoriale, non posso lasciare l'interrogativo in sospeso.

Un programma di intelligenza artificiale non è in grado intrinsecamente di realizzare dei profitti maggiori di quanti ne possa realizzare ciascuno di noi, con le proprie tecniche ed intuizioni, infarcite di errori. Bisogna diradare le nebbie, esaminandone la logica di funzionamento e guardando le cose nella giusta prospettiva. Un tale programma non viene codificato con rigide istruzioni del tipo “se si verifica questo evento, compra; se si verifica quest'altro evento, vendi”. Infatti, invece di essere strutturato su una serie di classici e stupidi “if... then...”, viene alimentato con degli esempi concreti. Le istruzioni che riceve non sono direttamente esecutive; sono invece algoritmi che istruiscono la macchina affinché possa “imparare” autonomamente dagli esempi che ad essa vengono sottoposti. In poche parole, tra un programma tradizionale ed uno intelligente passa la stessa identica differenza che separa un lavoratore da uno studente. Il primo partecipa già all'attività produttiva mentre il secondo si sta preparando.

Questa insolita particolarità dell'autoapprendimento non deve indurre a confidare necessariamente in un esito finale mirabolante, così come non tutti gli studenti concludono il loro percorso con un esito brillante. Quando avrà finito di imparare, un programma intelligente potrà essere più o meno efficace di un sistema classico di trading, più o meno profittevole. Presenterà dei punti di forza e dei punti di debolezza. E' il percorso che lo ha condotto ad essere operativo che è particolare, non necessariamente la sua capacità di generare profitti.

Capiremo meglio quando ci addentreremo nelle dinamiche del suo funzionamento. Per il momento ci basti sapere che, esattamente come per gli algoritmi tradizionali fondati sull'analisi tecnica o su altro, i risultati prodotti da macchine che apprendono sono funzione di una serie di fattori estranei. Essenzialmente, l'adeguatezza degli esempi utilizzati, l'efficacia del programma di addestramento, la possibilità di rilevare schemi di mercato invisibili all'occhio umano, ammesso che schemi di questo tipo siano presenti.

Sui primi, cioè sugli elementi di base utilizzati per l'apprendimento automatico, le perplessità abbondano. E' sufficiente leggere solo alcune delle innumerevoli “papers” in circolazione per rendersi conto che, quasi sempre, al programma di intelligenza artificiale applicato al trading vengono sottoposti dei tradizionali indicatori di analisi tecnica, nella speranza che il computer sappia interpretarli meglio di noi. Purtroppo, come si dice, “garbage

in, garbage out”[\[2\]](#). Il nostro amico “artificialmente intelligente” non può trarre conclusioni definitive e soddisfacenti da ingredienti che, per loro stessa natura, soggiacciono ad una sconfinata elasticità interpretativa.

Qualcuno ha pensato di alimentare il programma con una lunga serie storica di prezzi, nella speranza che vengano individuati dei pattern ripetitivi. Purtroppo, se tale tecnica può risultare estremamente efficace in altri settori, nell’ambito delle quotazioni finanziarie i pattern ripetitivi non possono essere evocati su comando. Peraltro, non esiste alcun valido motivo per credere che, se le dieci quotazioni più recenti hanno avuto lo stesso andamento di dieci quotazioni di sei mesi fa, l’undicesima mantenga la regolarità della successione. E che dire delle dimensioni dei pattern? Fissiamo costantemente il confronto al decimo step? Su quali basi discriminanti? Perché non al settimo o al ventesimo?

Mi astengo dal menzionare altre pseudo-soluzioni più o meno fantasiose. Allo stato attuale delle conoscenze, non c’è ragione di confidare nella presenza di precisi schemi comportamentali dei mercati finanziari; per ovvia conseguenza, non c’è programma automatico, intelligente o meno, che possa farli emergere. E’ naturalmente possibile che vengano rilevate delle tendenze di prezzo o alcune situazioni di particolare volatilità. Finché persistono... e sempre guardando nello specchietto retrovisore. E’ piuttosto semplice far prevedere ad un sistema intelligente la futura direzione di una media mobile. Ma la previsione è basata sul ben noto e scontato concetto della persistenza: un trend rimane tale finché non cambia. Insomma, è difficile estrarre da un programma intelligente qualcosa di più di quanto viene prodotto da un buon algoritmo tradizionale, che richiede minor fatica e minor dispendio di energie.

E sì! Anche questo va detto. Un programma di intelligenza artificiale è sempre assai costoso, in termini di tempo e fatica, per la sua realizzazione. Esistono delle ottime soluzioni a portata di mano, e ne parleremo, ma queste servono a poco se non vengono nutrite con una base dati impegnativa e laboriosa. E richiedono sempre domestichezza con dei linguaggi informatici. Non spaventatevi, però; arriveranno anche le buone notizie.

Prescindendo dai dati informativi, lo stesso programma di intelligenza artificiale presta il fianco a molte disillusioni. Non si può fare affidamento, infatti, su un modello ben definito. Piuttosto, bisogna scegliere tra una pluralità di proposte dai nomi simpatici rigorosamente inglesi (neural networks, random forests, decision trees, support vector machines, bagging, boosting, deep learning, etc.), ciascuna delle quali con i propri punti di forza ed i propri punti di debolezza.

Non è agevole orientarsi tra le possibili alternative, individuando esattamente la soluzione più appropriata alle proprie esigenze operative ed alla base dati di cui si dispone. Si può ricorrere, come si può facilmente supporre, a delle possibilità di screening automatico, anch’esso frutto dell’intelligenza artificiale, ma questo presuppone la presenza di un incontestabile risultato ottimale sul quale parametrizzare gli obiettivi. A differenza di quanto si possa credere, il potenziale profitto non è una misura della consistenza di un sistema di trading. Il percorso per conseguirlo è sempre costellato di una miriade di fattori che col sistema in sé hanno poco a che vedere; se non altro, l’esito è condizionato dalle varie tecniche di money management.

E, allora, siamo costretti a confrontarci nuovamente con l'assenza di schemi utili dei mercati, in grado di conferire un'affidabilità validamente misurabile alle previsioni fornite dai vari algoritmi intelligenti. Il sistema che funziona meglio oggi fallisce domani. Quello che sembra andar bene per il cross euro-dollaro non va bene per il cross sterlina-yen. E non si può risolvere la questione, come si potrebbe istintivamente pensare, realizzando modelli specialistici. Il problema non risiede nella presenza di schemi comportamentali diversi, ma nell'assenza di schemi.

Se dunque l'idea di un profitto stabile e costante è illusoria, come si fa a sottoporre alla macchina un obiettivo ben definito da raggiungere? Il programma finisce col "non convergere", come si dice. Si accorge presto di non essere in grado di costruire un modello soddisfacente, cioè di non poter sviluppare delle equazioni che soddisfino tutti, o quasi, gli esempi che riceve per lo studio, a meno che non vengano allargati i parametri di apprendimento; che ci si accontenti, cioè, di un modello che memorizzi i dati sotto analisi, restando incapace di generalizzare le sue conclusioni su dati mai visti prima.

Sforzo inutile quindi? Stiamo perdendo del tempo? Non necessariamente. Se, ridimensionando le aspettative, rinunciamo all'idea che un programma di intelligenza artificiale sia una macchina che produce quattrini a palate ed in maniera scontata, mentre noi ce ne stiamo in spiaggia ad arrostarci al sole, possiamo benissimo esplorare altre diverse possibilità interessanti: ad esempio, l'uso di algoritmi capaci di fornire indicazioni statisticamente accettabili che affianchino, con funzioni di semplice confronto o di eventuale convalida, un qualsiasi vecchio modello meno "artificialmente intelligente", anch'esso non privo di inevitabili incongruenze ed errori; oppure possiamo far scegliere al computer il mercato (azionario, delle commodities, delle valute) che, in un determinato momento, risponde meglio al nostro tipo di trading. Cose di questo genere, insomma, nell'attesa che qualcuno scopra il Sacro Graal del trading, che naturalmente terrà per se stesso.

Più in particolare, in questo libro seguiremo una strada piuttosto affollata allo stadio attuale dell'intelligenza artificiale: mettere insieme un gruppo di algoritmi diversi, considerandoli come degli esperti che, a conclusione delle loro valutazioni, esprimeranno un voto in un sistema a maggioranza. E' quanto di più promettente credo si possa fare in un campo così impreciso come il nostro.

Questa tecnica, che prende il nome di "ensemble", ha il pregio di aumentare significativamente, secondo dei principi statistici ben definiti, le fragili probabilità di successo dei singoli componenti.

Quella che troverete unica e originale, mi auguro, sarà l'attenzione che verrà prestata ai risvolti pratici ai quali il trader è interessato.

L'autore

[1] Third international symposia on intelligent industrial automation and soft computing – International Computer Science Conventions Canada/Switzerland - Genova 1-4 giugno 1999.

Financial application of neural nets and fuzzy systems – Department of applied mathematics – University “Ca’ Foscari”, Venezia – 6 maggio 1999.

[\[2\]](#) Se immetti spazzatura, viene fuori spazzatura.