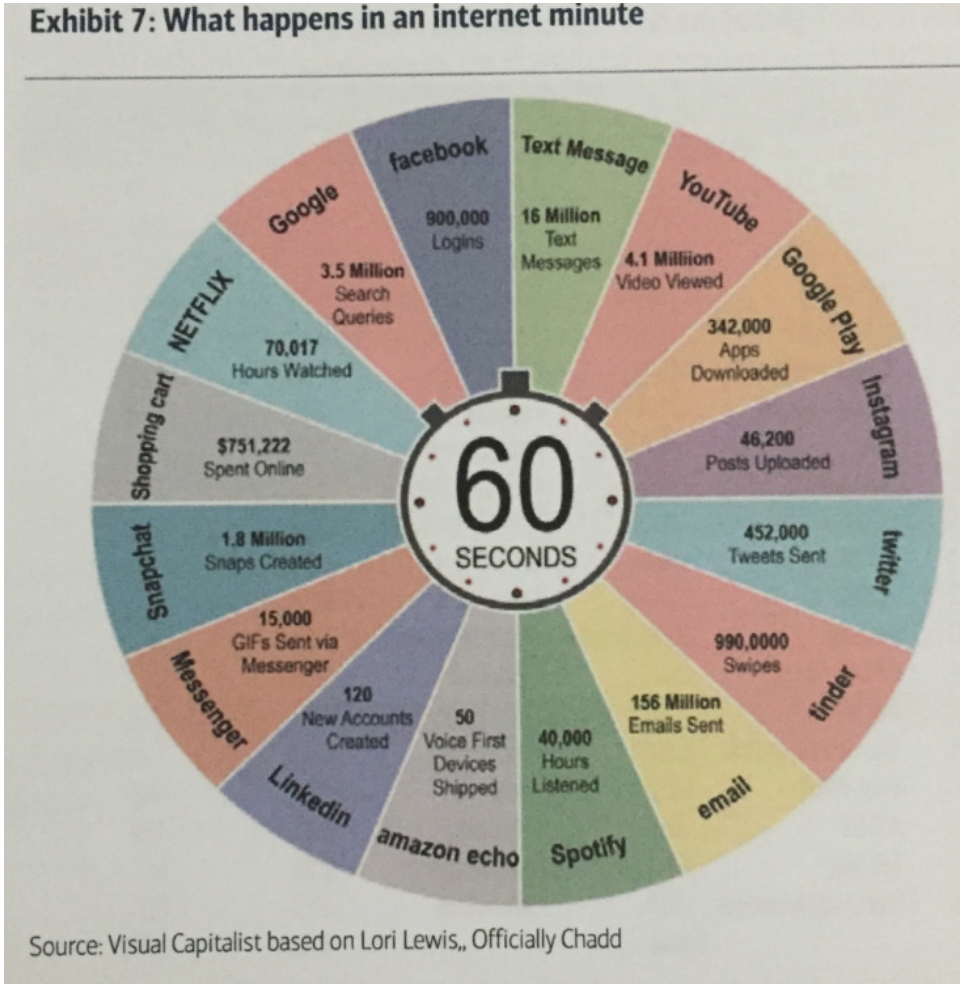
«BIG DATA E ARTIFICIAL INTELLIGENCE APPLICATI AGLI INVESTIMENTI»

UNA (PARZIALE) REVIEW DELLA LETTERATURA

Elisabetta Basilico
Novembre 2017

Big Data: Premessa



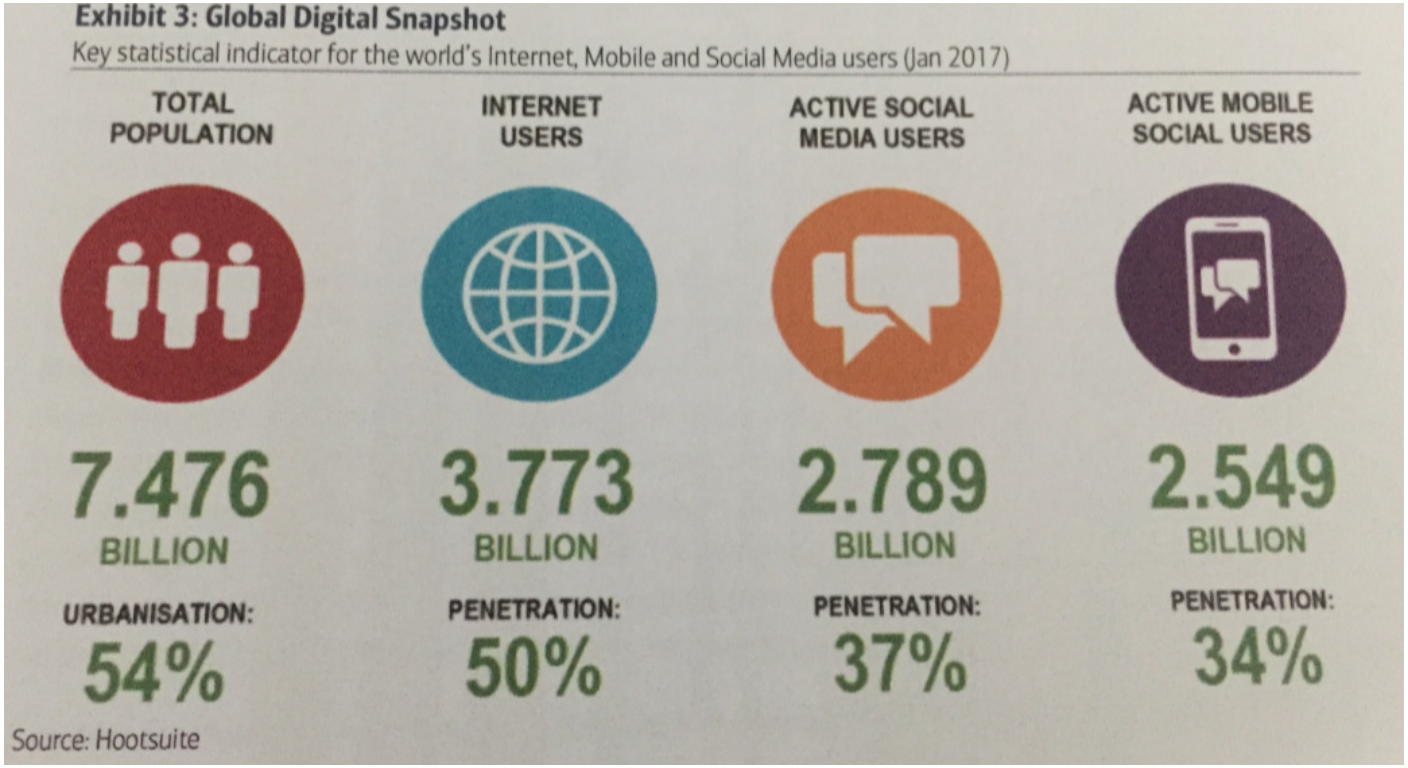
- Ad oggi esistono più di 1 miliardo di siti internet e 100 milioni di nuovi vengono aggiunti in media ogni anno per un totale di 10,000 miliardi di singole pagine web(1)

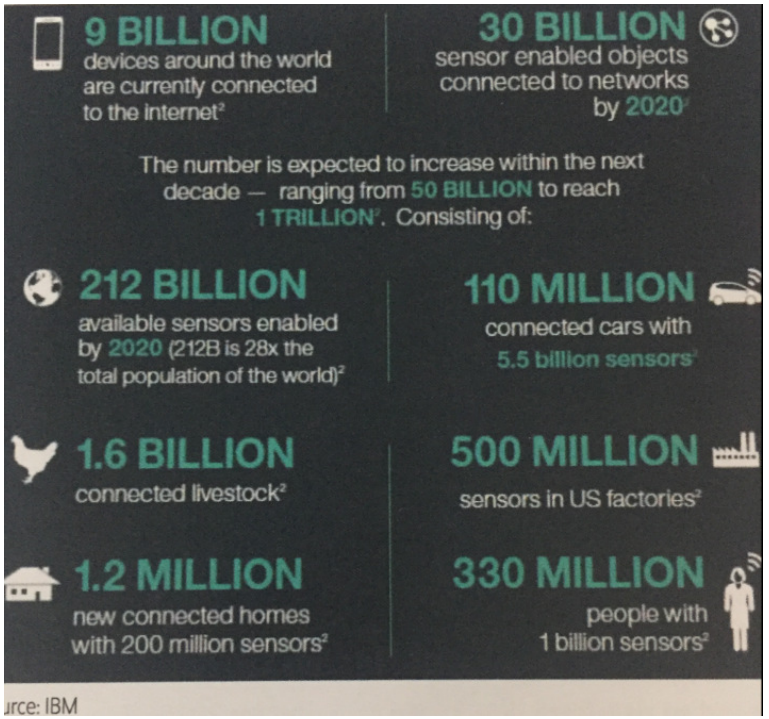
- Google ha indicizzato una quantità di dati pari solo allo 0,01% di quelli disponibili (2). Il 90% dei dati disponibili oggi a livello globale sono stati generati negli ultimi due anni e si prevede che il volume di dati generato raddoppierà ogni 2-3 anni (3)

- Studi di Accenture e McKinsey stimano che Big Data &AI possano aggiungere 0.8-1.4% alla crescita produttiva globale nel lungo periodo (entro il 2030). Sono paragonati all'effetto di internet a partire dagli anni '90 e dei telefoni cellulari degli inizi anni '2000.

(1),(2) Fonte: Big Data in Investment Management, Deutsche Bank Market Research, 2016
(3) Fonte: IBM, IDC, 2016

Big Data: Cos'è cambiato dagli anni '90?

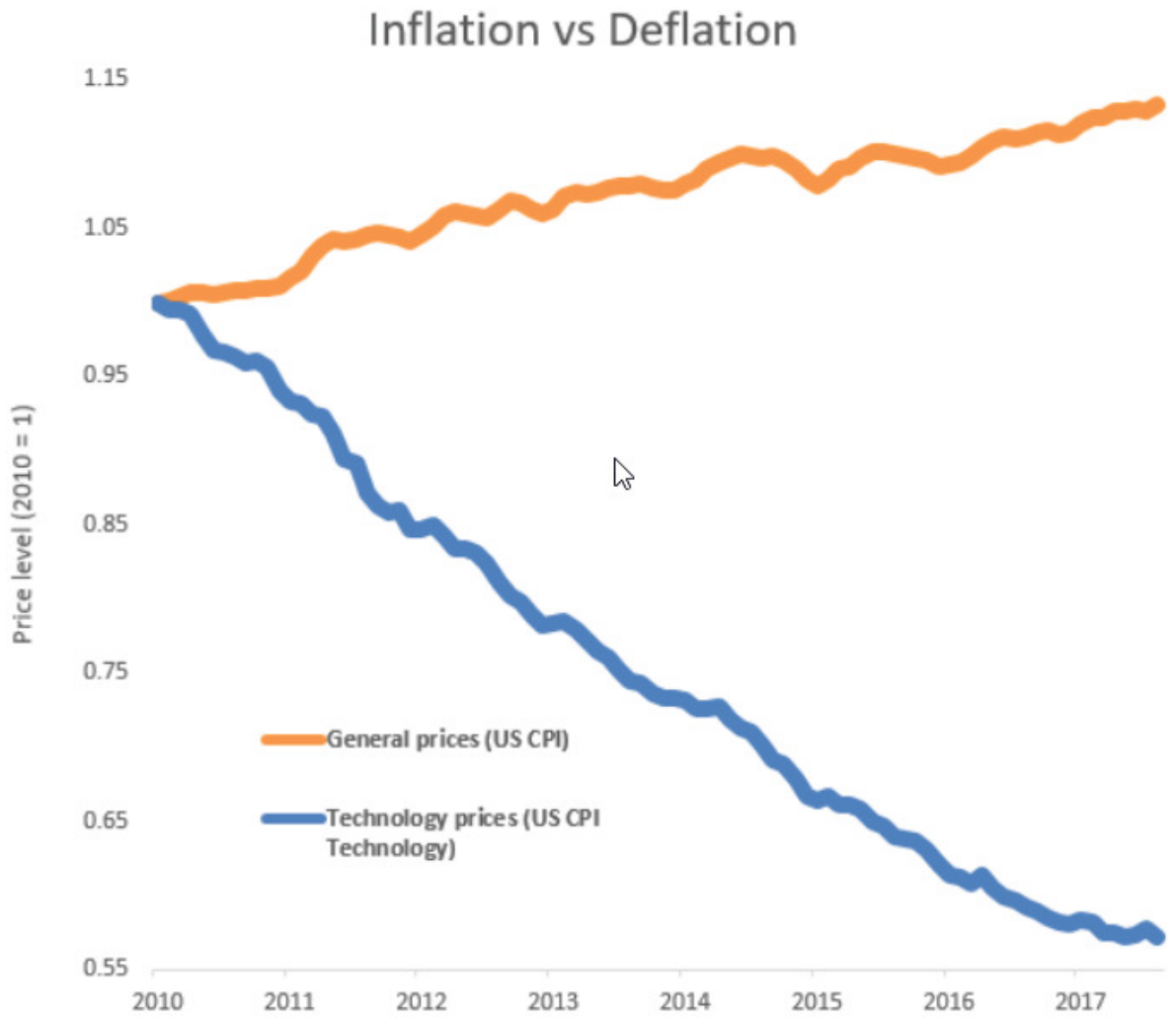




- 3.7 miliardi o il 50% della popolazione mondiale utilizza internet (rispetto a 14 milioni nel 1993)
- 2+ miliardi di « utilizzatori di smartphone» (rispetto a 0 nel 1993 e 80 milioni di telefoni nel 1995)
- L'87% degli utilizzatori di smartphone dice di esserne «dipendente»
- Il costo dello stoccaggio dati (storage) è diventato 24,450 volte più economico
- Il costo del potere computazionale è diventato 10,540 volte più economico

Fintech. Big Data. Machine Learning. AI. Data Science. And so on.

Well, basically it's all here, in this simple chart:



Fonte: Zenti, R., Making Sense of Fintech and Data Science in One Simple Chart, Medium, 2017

Big Data: Cosa sono?

- I dati di bilancio, i prezzi delle varie asset class , i volumi scambiati in Borsa sono definiti dati «strutturati», spesso già ben organizzati in database commerciali e di facile analisi attraverso software quali SQL e Excel. Questi dati rappresentano il 10% dei dati disponibili (4).

- A differenza dei dati strutturati, i Big Data sono definiti «dati non strutturati» e caratterizzati da elevati «volumi, varietà e velocità (3Vs).

- Alcuni esempi di Big Data sono: social media, text, immagini satellitari e non.

- La maggior parte delle aziende in diversi settori parla di Big Data e AI ma pochi sono veramente «up to speed». Si stima che solo il 10% utilizzi già i Big Data per decisioni aziendali e che solo il 10% faccia uso di tecniche di intelligenza artificiale.

- Nel settore dell'asset management è fondamentale avere accesso a informazioni e saperle analizzare. Le informazioni non mancano. Naturale che questo settore sia terreno fertile per la ricerca e per verificare se le nuove tecniche di AI possano veramente creare valore con nuove strategie di investimento.

(4) Fonte: Thematic Investing, BAML, 2017

Big Data: Esempi

- INRIX (GPS and Mobile network data)
- PREMISE (real time inflation signals)
- ORBITAL INISGHTS (Satallite Imaginery)
- TWITTER (Stockwits; DataMinr)
- **THE BIG DATA FOREST**: according to EagleAlpha there are 500 alternative data providers

The Wisdom of Twitter Crowds: *Predicting Stock Market Reactions to FOMC Meetings via Twitter Feeds*

PABLO D. AZAR AND ANDREW W. LO

FORMAT

Journal of Portfolio Management, 2016

The Information Content of Forward-looking Statements in Corporate Filings—a Naïve Bayesian Machine Learning Approach *

Feng Li

Journal of Accounting Research, 2010

Papers -3

Deep Learning for Finance: Deep Portfolios

15 Pages • Posted: 14 Sep 2016

J.B. Heaton
Conjecture LLC; J.B. Heaton, P.C.

Nick Polson
University of Chicago - Booth School of Business

Jan Hendrik Witte
University of Oxford - Mathematical Institute

Applied Stochastic Models in Business and Industry, 2016

Papers -4

News Versus Sentiment: Predicting Stock Returns from News Stories

FEDS Working Paper No. 2016-048

36 Pages • Posted: 9 Jun 2016

Steven L. Heston
University of Maryland - Department of Finance

Nitish Ranjan Sinha
Board of Governors of the Federal Reserve System

Financial Analyst Journal, 2017

Papers -5

Academic Research Insight: The Social Media Factor

May 27, 2017 Academic Insights Blog

- **Title**: THE SIXTH FACTOR- A SOCIAL MEDIA FACTOR DERIVED DIRECTLY FROM MEDIA SENTIMENTS
- **Authors**: JIM LIEW AND TAMAS BUDAVARI
- **Publication**: THE JOURNAL OF PORTFOLIO MANAGEMENT, SPRING 2017 (version here)

Genetic programming optimization for a sentiment feedback strength based trading strategy

- Steve Y.Yang, Sheung Yin Kevin Mo, AnqiLiu, Andrei A.Kirilenko

Neurocomputing, 2017

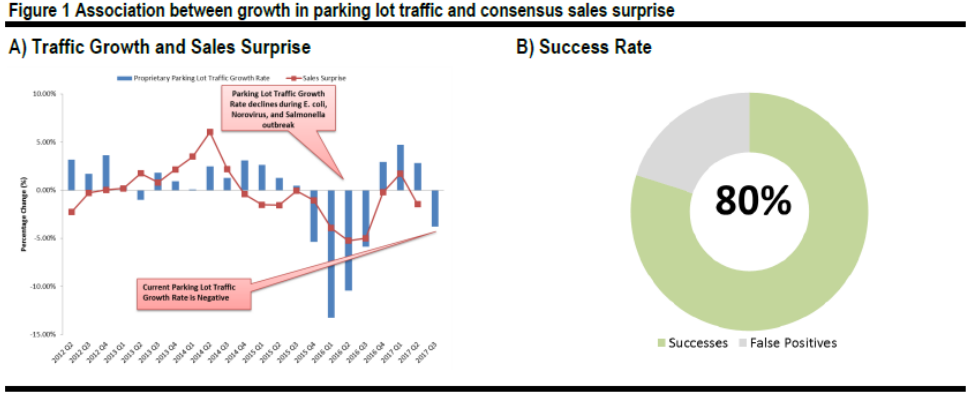
Big Data: Esempi

t_0 — Parking Lot Traffic Data (Immagini Satellitari)

t_1 — Quarterly Earnings Announcements

capacità di predirre le *«sales surprises»* (differenza tra il revenue consensus e i reported revenues)

Figure 1 Association between growth in parking lot traffic and consensus sales surprise



Sources: IBES, S&P Capital IQ, FTSE Russell, RS Metrics, Wolfe Research Luo's QES

Sub-Industry Group	T-statistics	Number of Observations	Hit Rate
Restaurants	4.50	104	71%
Home Improvement Retail	3.64	67	65%
Specialty Stores	2.89	120	64%
Department Stores	0.64	83	54%
Homefurnishing Retail	0.62	33	67%
General Merchandise Stores	0.37	71	56%
Computer & Electronics Retail	0.37	31	45%
Food Retail	0.21	25	50%
Apparel Retail	0.12	20	25%
Automotive Retail	(0.21)	24	56%
Specialty Chemicals	(0.24)	21	33%
Retail REITs	(0.70)	4	0%
Hypermarkets & Super Centers	(1.50)	26	53%

Stronger Statistical Significance

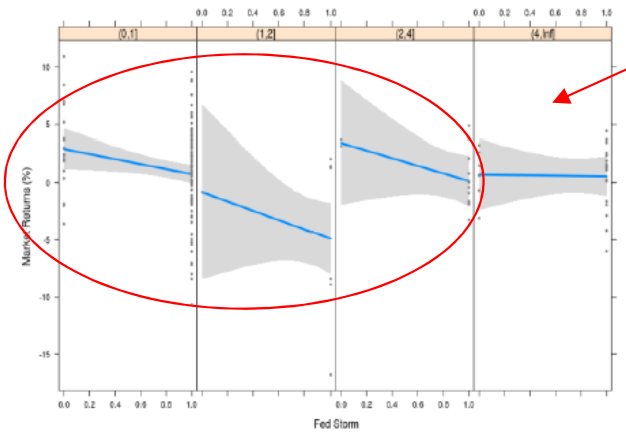
Stock Selection

Big Data: Esempi

- Prattle è una società che analizza dati in formato «testo» attraverso la tecnica del *Natural Language Processing* (NLP). Copre informazioni su 3,000 società e 15 banche centrali.

- E' possibile calcolare un *sentiment score* per ciascuna comunicazione. Nel caso delle banche centrali, si definiscono comunicazioni «hawkish» (in caso di sentiment positivo) oppure «dovish» (in caso di sentiment negativo).

- Studi empirici (L. Qes, 2017) mostrano che sia possibile costruire degli indicatori che sono predittivi dell'andamento futuro di mercato



La pendenza negativa della retta di regressione indica che all'aumentare di comunicazioni estreme da parte delle banche centrali, il mercato subirà dei movimenti al ribasso

Big Data: Esempi

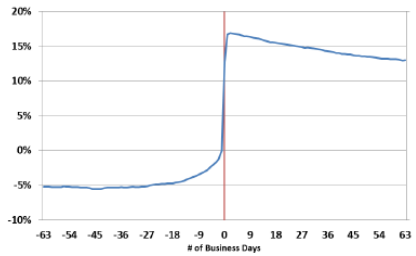
U.S. SECURITIES AND EXCHANGE COMMISSION

- Edgard è il database della SEC dove le società quotate negli USA devono compilare una serie di documenti con svariate informazioni sulla società stessa, i propri azionisti etc.

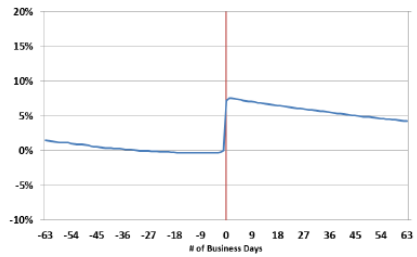
- Queste informazioni possono essere utilizzate per diverse strategie di investimento.

Figure 5 Event Study

A) Average Cumulative Excess Returns



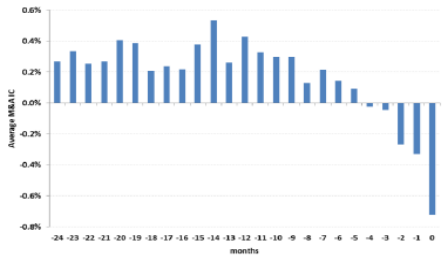
B) Median Cumulative Excess Returns



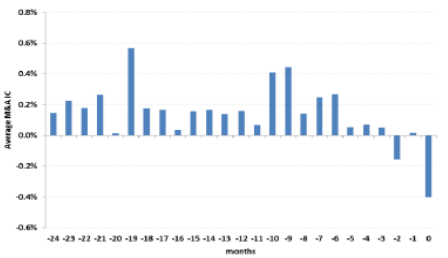
Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Figure 16 Average M&A IC, # of Filings, by Month Prior to M&A Announcement

A) Form 4



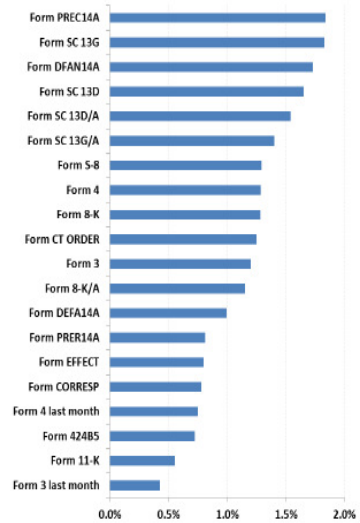
B) Form 3



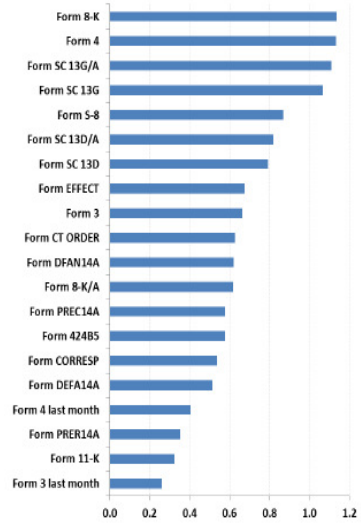
Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Figure 32 Performance, EDGAR Filing Factors

A) Average M&A IC



B) Risk Adjusted M&A IC



La tabella indica quali report che ogni società americana è obbligata a depositare presso la SEC sia maggiormente predittivo di futuri eventi di M&A.

FORM 4

☐ Check this box if no longer subject to Section 16. Form 4 or Form 5 obligations may continue. *See* Instruction 1(b).

(Print or Type Responses)

UNITED STATES SECURITIES AND EXCHANGE COMMISSION
Washington, D.C. 20549

STATEMENT OF CHANGES IN BENEFICIAL OWNERSHIP

Big data and AI: Caveat da tenere presente

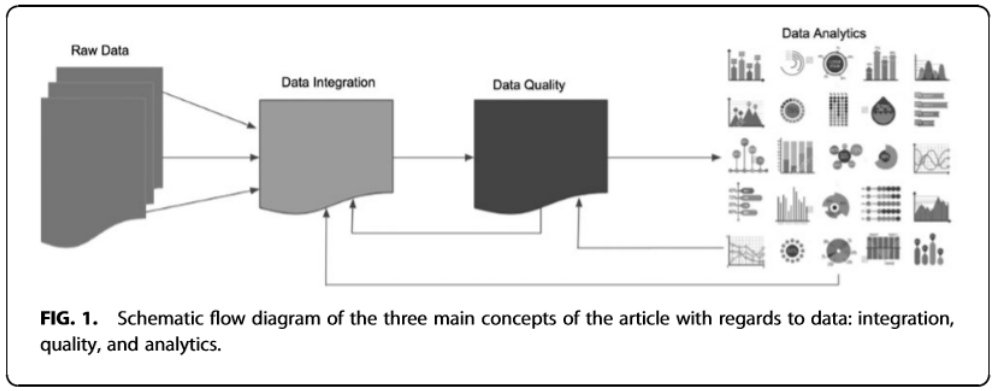
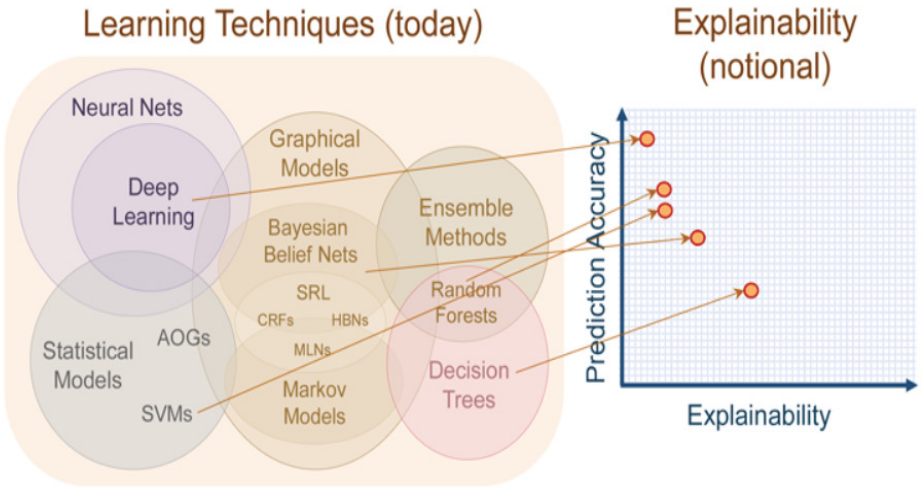


FIG. 1. Schematic flow diagram of the three main concepts of the article with regards to data: integration, quality, and analytics.

Fonte: Alexander et al., 2017, Research Challenges in Financial Data Modeling and Analysis, *Big Data*



WHAT VS. WHY: Modern learning algorithms show a tradeoff between human interpretability, or explainability, and their accuracy. Deep learning is both the most accurate and the least interpretable.

Darpa

- Complessità dei dati: sono necessarie tecniche di standardizzazione e pulizia
- Overfitting/Data Mining (servono serie di dati sufficientemente lunghe)
- Trasparenza non sempre ottimale

QUESTIONS?

Academic Insights on Investing

Scouting the latest academic research that backs successful investing ideas

THANKS!!!

Academicinisghtsoninvesting.com