



BANCA D'ITALIA
EUROSISTEMA

Big data e tecnologie di machine learning per l'attività delle banche centrali

Intervento di apertura del Vice Direttore Generale della Banca d'Italia
Fabio Panetta

Roma, 26 marzo 2018

Sono lieto di aprire i lavori di questo seminario su “Big data e tecnologie di machine learning per l’attività delle banche centrali” e di dare il benvenuto ai relatori e a tutti i partecipanti. Permettetemi di cogliere questa occasione per ringraziare gli organizzatori per aver definito un programma particolarmente ricco e attuale.

I big data e il machine learning derivano dalla diffusione delle tecnologie digitali, che stanno trasformando le modalità di comunicazione, di apprendimento, di trasmissione della conoscenza. La digitalizzazione sta determinando importanti cambiamenti nella nostra società. Circa l’80 per cento dei cittadini europei possiede uno smartphone, un computer o un tablet. Internet è ampiamente usato per raccogliere informazioni, comunicare e svolgere attività che influenzano direttamente o indirettamente il comportamento degli individui.

In Italia l’80 per cento dei consumatori utilizza Internet attraverso uno smartphone; la metà di essi accede direttamente al proprio conto corrente e acquista beni e servizi online. Circa un terzo delle aziende italiane ha automatizzato le attività amministrative e di *back-office*.

Gli effetti di questi cambiamenti sono particolarmente rilevanti nel settore finanziario. A livello mondiale, solo negli ultimi due anni il numero di coloro che utilizzano soluzioni di pagamento offerte da grandi aziende tecnologiche, quali Apple, Google, Samsung o Android, è passato da 18 a 144 milioni. Amazon Lending¹ nell’ultimo anno (da giugno 2016 a giugno 2017) ha concesso prestiti per oltre un miliardo di dollari, un volume tre volte superiore alla media dei cinque anni precedenti.

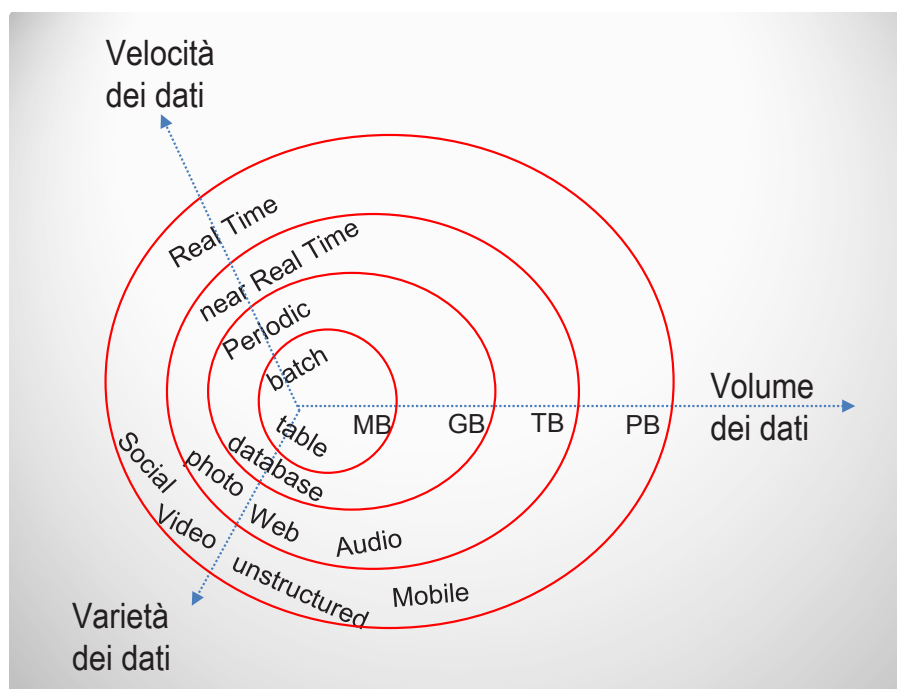
Per Big data si intende l’applicazione di tecniche innovative che consentono di elaborare volumi di informazioni digitali su una scala di gran lunga superiore rispetto a quella degli approcci tradizionali². Il termine “big” fa riferimento alle dimensioni e alla diversità dei dati, nonché alla rapidità con cui questi vengono generati. Ciò si sintetizza nelle tre V: volume, velocità e varietà, come illustrato nella figura.

Le banche centrali possono e sicuramente devono svolgere un ruolo attivo nello sfruttamento delle tecnologie digitali e dell’enorme mole di dati che esse generano.

¹ Amazon Lending è la società del gruppo Amazon che offre servizi di finanziamento alle imprese che vendono attraverso Amazon.

² Cfr. Lohr, S., “How Big Data Became So Big”, *The New York Times*, 11 agosto 2012.

Big data: una rapida espansione su tre fronti



Nell'attuale contesto le autorità pubbliche possono migliorare la propria capacità di azione facendo leva su una grande quantità di informazioni assai varie e tempestive sulle preferenze e sulle scelte degli agenti economici. Insieme ai dati contenuti in archivi tradizionali, oggi i social media e i dispositivi mobili generano infatti una enorme quantità di informazioni sia in forma semi-strutturata (come XML o JSON) sia non strutturate (audio, video, testi). Questi dati, insieme a quelli tradizionali, possono essere usati per costruire indicatori dell'attività economica più accurati e tempestivi.

Vi sono importanti esempi di utilizzo dei big data da parte delle autorità pubbliche, a fini di policy. I big data sono impiegati per stimare il tasso di disoccupazione o il tasso di inflazione, per migliorare le previsioni di variabili rilevanti a fini di policy, per misurare il clima di fiducia di consumatori e imprese³.

³ D'Amuri e Marcucci (2017) mostrano che i modelli che utilizzano i dati di Google Trends forniscono previsioni più accurate sulla disoccupazione negli Stati Uniti rispetto a quelli che impiegano i tradizionali indicatori anticipatori. Avvalendosi dei dati online desunti dal Billion Prices Project (BPP), Cavallo (2013) trova che nel periodo 2007-11 – a fronte di un tasso medio di inflazione dichiarato dal governo argentino pari all'8 per cento sui dodici mesi – i dati online suggerivano una dinamica dei prezzi superiore al 20 per cento, in linea con le stime di alcune amministrazioni provinciali e con i risultati tratti dalle indagini sulle aspettative di inflazione delle famiglie. Goolsbee e Klenow (2018) utilizzano le transazioni online relative a milioni di prodotti nel periodo 2014-17 per dimostrare che negli Stati Uniti l'inflazione ricavata da dati online è inferiore (di circa 1,3 punti percentuali all'anno) a quella calcolata in base all'indice dei prezzi al consumo per le medesime categorie di prodotti. Daas e Puts (2014) elaborano una misura del clima di fiducia dei cittadini olandesi utilizzando i messaggi dei social media, che risultano essere strettamente correlati e cointegrati con l'indice di fiducia dei consumatori ricavato dalle indagini tradizionali.

Dal punto di vista delle banche centrali i possibili utilizzi sono numerosi, vi sono tuttavia diverse sfide da affrontare. Gli aspetti tecnici verranno esaminati a fondo durante il seminario, ma vorrei sin d'ora sottolineare un problema metodologico di carattere generale: prima di utilizzare a fini di policy le relazioni stimate utilizzando Internet o i social media, occorre essere certi che tali stime siano robuste, rappresentative e affidabili. Per conseguire questo obiettivo occorrono cospicui investimenti in ricerca.

Un importante settore influenzato da vicino dalle nuove tecniche di estrazione di informazioni dai big data è quello della stabilità finanziaria. Le imprese fintech utilizzano i big data e le tecnologie digitali per creare nuove linee di attività e garantire un accesso continuo ai propri servizi finanziari. Si tratta di un'evoluzione positiva, perché capace di stimolare la concorrenza e la produttività. Tuttavia, la diffusione di nuovi prodotti e i legami diretti fra banche, nuovi intermediari e investitori potrebbero influenzare la stabilità del sistema. In un tale contesto, le autorità pubbliche sono chiamate a esaminare a fondo i cambiamenti indotti dalle tecnologie digitali e a valutare attentamente i progetti innovativi al fine di preservare la stabilità e l'efficienza del settore finanziario.

Il grado di complessità che deriva da questi sviluppi è accentuato dal fatto che le banche centrali sono a un tempo fruitrici e produttrici di informazioni e di big data. Si pensi, ad esempio, alle informazioni granulari sui singoli prestiti bancari raccolte dal Meccanismo di vigilanza unico (la cosiddetta piattaforma AnaCredit), alle segnalazioni statistiche sui mercati monetari raccolte giornalmente dal Sistema europeo di banche centrali e ai dati sui contratti derivati la cui raccolta è disciplinata dal regolamento sulle infrastrutture di mercato europee (EMIR).

La enorme dimensione di queste informazioni granulari rende evidente che le necessarie procedure di controllo e di validazione dei dati non potranno essere effettuate manualmente, ma richiederanno l'utilizzo di tecniche e algoritmi complessi, quali il machine learning o l'intelligenza artificiale.

Le banche centrali devono conoscere a fondo tali nuove tecnologie. La Banca d'Italia ha costituito al suo interno un gruppo di lavoro sui big data composto da economisti, statistici e informatici provenienti da diversi servizi, operanti in stretta collaborazione con il Dipartimento di Informatica. Il gruppo ha predisposto un'infrastruttura hardware e software per il trattamento di diverse tipologie di big data da utilizzare per temi sia macroeconomici sia microeconomici. Questo seminario offre l'occasione per esaminare alcuni dei risultati sin qui raggiunti. Le analisi si concentrano sull'utilizzo di dati testuali non strutturati desunti dai social media, in particolare da Twitter, per calcolare

le aspettative di inflazione⁴ o per misurare la fiducia dei depositanti al dettaglio⁵. I social media sono inoltre impiegati per valutare l'attitudine dei clienti nei confronti di particolari aziende e l'impatto che essa ha sui rendimenti azionari, sulla volatilità e sui volumi scambiati⁶. Twitter e le notizie che si rendono via via disponibili sono usati per misurare l'incertezza relativa alle politiche economiche e per esaminare le truffe relative alle carte di pagamento al fine di misurarne gli effetti sulle scelte di pagamento dei consumatori. Un ulteriore filone di ricerca si avvale dei dati ricavati da singoli annunci immobiliari online per esaminare la microstruttura del mercato immobiliare italiano⁷.

In conclusione, vorrei ringraziare ancora una volta i relatori, gli esperti, i moderatori e tutti i partecipanti per essere intervenuti a questo seminario. Seguirò con interesse le vostre analisi su big data e sul machine learning, i suggerimenti circa le possibili applicazioni per le banche centrali, in particolare ai fini di policy.

È assai positivo che questo seminario riunisca ricercatori provenienti sia dalle banche centrali sia dal mondo accademico, al fine di ottenere contributi preziosi da diversi punti di vista. Sono certo che il dibattito di questi due giorni risulterà di grande interesse e altamente produttivo.

⁴ Cfr. C. Angelico, J. Marcucci, M. Miccoli, e F. Quarta, "Can We Measure Inflation Expectations Using Twitter?"

⁵ Cfr. M. Accornero e M. Moscatelli in "Listening to the buzz: social media sentiment and retail depositors' trust", una misura della fiducia dei depositanti al dettaglio ricavata dai commenti Twitter, Banca d'Italia, *Temi di Discussione (Working Papers)*, n. 1165, 2018.

⁶ Cfr. G. Bruno, P. Cerchiello, J. Marcucci, e G. Nicola: *Twitter Sentiment and Banks' Equities: Is there any causal link?*

⁷ Cfr. M. Loberto, A. Luciani, and M. Pangallo, *The Potential of Big Housing Data: an application to the Italian Real Estate Market*, basato su dati forniti da www.immobiliare.it.

Riferimenti bibliografici

- Accornero M. e M. Moscatelli, 2018, *Listening to the buzz: social media sentiment and retail depositors' trust*, Banca d'Italia, Temi di Discussione (Working Papers), n. 1165, gennaio.
- Angelico C., J. Marcucci, M. Miccoli e F. Quarta, 2018, *Can We Measure Inflation Expectations Using Twitter?*, mimeo.
- Bruno G., P. Cerchiello, J. Marcucci e G. Nicola, 2018, *Twitter Sentiment and Banks' Equities: Is there any causal link?*, mimeo.
- Cavallo A., 2013, *Online and official price indexes: Measuring Argentina's inflation*, Journal of Monetary Economics, vol. 60(2), pp. 152-165, marzo.
- Daas P. J. H. e M. J. H. Puts, 2014, *Social Media Sentiment and Consumer Confidence*, ECB Statistics Paper Series, n. 5, settembre.
- D'Amuri F. e J. Marcucci, 2017, *The predictive power of Google searches in forecasting US unemployment*, International Journal of Forecasting, vol. 33(4), pp. 801-816, ottobre.
- Goolsbee A. D. e P. J. Klenow, 2018, *Internet Rising, Prices Falling: Measuring Inflation in a World of E-Commerce*, Stanford University working paper, January.
- Loberto M., A. Luciani e M. Pangallo, 2018, *The Potential of Big Housing Data: an application to the Italian Real Estate Market*, mimeo.
- Lohr S., 2012, *How Big Data Became So Big*, The New York Times, 11 Agosto 2012.

