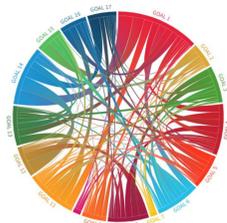


# AGENDA 2030 E INTERLINKAGES: UN'ESPLORAZIONE DELLE RELAZIONI TRA INDICATORI SDGs ATTRAVERSO IL CONFRONTO TRA TECNICHE DI MACHINE LEARNING E TECNICHE TRADIZIONALI

MARCO RAO - ENEA marco.rao@enea.it | PAOLA UNGARO - Istat ungaro@istat.it

L'Agenda 2030 per lo sviluppo sostenibile, varata nel 2015 dall'Assemblea Generale delle Nazioni Unite e sottoscritta dai 193 Paesi dell'ONU, viene monitorata attraverso un sistema di indicatori riferiti ai diversi ambiti sociale, economico, ambientale ed istituzionale dello sviluppo sostenibile, definiti e aggiornati dal UN-Inter Agency Expert Group on SDGs (UN-IAEG). Nella versione attuale, gli indicatori UN-IAEG-SDGs sono 248, 231 dei quali unici (non ripetuti in più Goal). Nel processo di costruzione dell'informazione statistica richiesta a livello UN per il monitoraggio dello stato di avanzamento degli SDGs, l'Istat, così come gli altri istituti nazionali di statistica, è investito del ruolo di coordinamento delle attività di produzione del Sistema statistico nazionale. Dal 2016, l'Istat pubblica le misure statistiche per l'Italia due volte all'anno, unitamente ad un Report che ha cadenza annuale.



E' importante dedicare particolare attenzione alle analisi degli *interlinkages* tra i 17 Goals, i 169 target dell'Agenda e gli indicatori che li rappresentano, allo scopo di monitorare le condizioni di compatibilità tra progresso sociale ed economico e salvaguardia ambientale, nell'ambito di quella complessità fenomenologica, concettuale e statistica che caratterizza la tematica dello sviluppo sostenibile. «Le interconnessioni degli SDGs sono di importanza cruciale nell'assicurare che lo scopo della nuova Agenda venga realizzato. Se noi realizzeremo le nostre ambizioni abbracciando l'intera Agenda, le vite di tutti verranno profondamente migliorate e il nostro mondo sarà trasformato al meglio» (UN. 2015. Agenda 2030). L'individuazione delle relazioni concettuali, metodologiche e sostantive tra target e indicatori è dunque funzionale alla lettura integrata dello sviluppo sostenibile, nell'ambito di un monitoraggio congiunto, supporto essenziale a politiche che tengano conto di criticità e sinergie per migliorare la coerenza.



Obiettivo di questo poster è una prima esplorazione dei legami tra alcune misure statistiche adottate dal Sistema informativo Istat-SDGs per l'Italia, per individuare la connessione tra il Goal 7 (Energia pulita), 12 (Consumo e produzione responsabile) e 13 (Lotta al cambiamento climatico), utilizzando approcci legati alla statistica tradizionale, confrontati con approcci del *machine learning* (ML). Si pongono a confronto, in particolare, diverse metodologie per la stima delle relazioni tra le misure Istat-SDGs che fanno capo a modelli lineari tradizionali e modelli basati sull'uso di macchine a vettori di supporto (SVM), un approccio di ML scelto per la sua particolare adattabilità ad operare regressioni.



Le regressioni lineari possono gestire relazioni estremamente complesse tra variabili, tuttavia un possibile svantaggio del modello lineare è che queste relazioni devono essere specificate affinché la regressione possa rappresentarle con un buon livello di adattamento. I modelli complessi di apprendimento automatico, come le macchine a vettori di supporto, le foreste casuali, le reti neurali, ecc., sanno invece auto-adattarsi a comprendere tali relazioni.



L'adozione di modelli di produzione sostenibili riduce il rischio di erosione quali-quantitativa delle risorse ambientali e comporta vantaggi in termini di salute e benessere sociale, costituendo, al contempo, uno strumento di competitività economica. Incrementare la produttività delle risorse, promuovere il ricorso a processi e tecnologie puliti, limitare la cessione di carichi inquinanti agli ecosistemi sono obiettivi che concorrono a ridurre l'impronta ecologica dei sistemi socio-economici, che possono essere monitorati da indicatori afferenti a più Goal dell'Agenda 2030.

In questa sede si è deciso di concentrarsi, in particolare, su due target del Goal 7 (7.2 e 7.3), finalizzati alla promozione, rispettivamente, delle energie rinnovabili e dell'efficienza energetica, sul target 12.2, finalizzato alla gestione sostenibile e all'uso efficiente delle risorse naturali e sul target 13.2 finalizzato al rafforzamento delle misure di contrasto ai cambiamenti climatici.

Figura 1 - Intensità energetica e quota di energia da fonti rinnovabili. Anni 2012-2022 (Fonte: Istat)

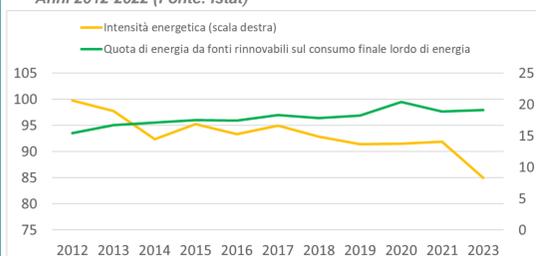


Figura 2 - Intensità energetica e Consumo materiale interno/Pil. Anni 2012-2022 (Fonte: Istat)

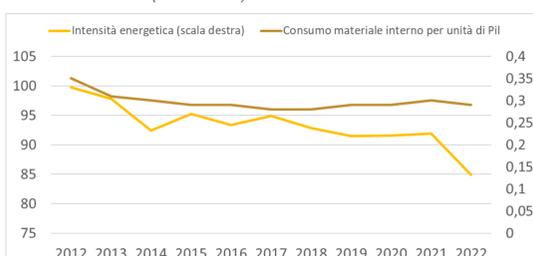
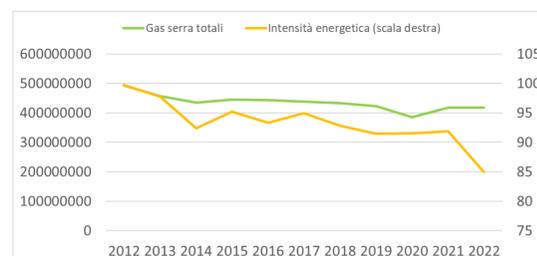
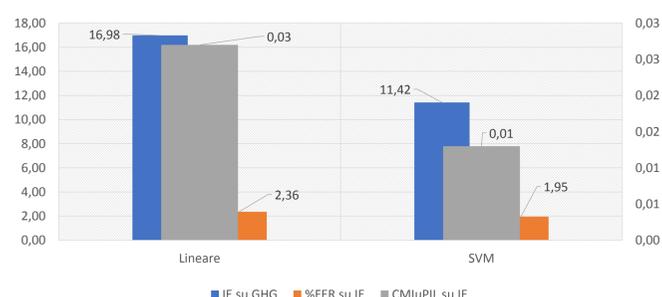


Figura 3 - Intensità energetica e gas serra totali. Anni 2012-2022 (Fonte: Istat)



Le misure statistiche selezionate per l'analisi sono:  
7.2.1.- Quota di energia da fonti rinnovabili sul consumo finale lordo di energia (%FER)  
7.3.1 - Intensità energetica (IE)  
12.2.2 - Consumo di materiale interno per unità di Pil (CMI)  
13.2.2 - Emissioni totali di gas serra per anno (GHG).

Figura 4 - RMSE nella previsione con modello lineare e SVM per le regressioni dell'intensità energetica rispetto alle emissioni di gas serra. Anni 2015 - 2022 (Fonte: elaborazioni su dati Istat)



Si tratta di indicatori tra loro correlati. Nel corso degli ultimi dieci anni, le molteplici misure adottate a livello internazionale e nazionale per raggiungimento della neutralità climatica al 2050 hanno fissato target emissivi ed energetici sempre più ambiziosi, stimolando la produzione di energia da fonti rinnovabili e il risparmio energetico. La contrazione dell'intensità energetica si è accompagnata all'incremento della quota di consumi energetici da fonti rinnovabili, con ricadute positive sulle emissioni di sostanze climalteranti. Il forte impulso alla promozione di modelli di produzione e consumo sostenibile finalizzati al passaggio a un'economia efficiente nell'impiego delle risorse ha determinato un progressivo *decoupling* tra consumo di materia e attività economiche.

Le analisi effettuate mostrano come, in media, la SVM fornisca migliori performance rispetto ai modelli tradizionali, in quanto ad efficacia esplicativa, in particolare nel trattare dati complessi. L'efficacia comparata dei modelli è stata misurata in termini di misure di errore statistico (RMSE - *Root Mean Square Error*), una misura che presenta una più immediata interpretabilità, in quanto espressa nella stessa unità di misura della variabile di risposta.

Per il dato nazionale, la regressione a vettori di supporto riduce l'errore di circa un terzo rispetto ai modelli tradizionali: 17% per il rapporto tra intensità energetica e quota di FER; 33% per il rapporto tra IE e gas serra; 52% per il rapporto tra IE e consumo di materia.

Per i dati regionali, disponibili solo per la quota di FER e il CMI, la riduzione media è invece più contenuta: circa il 13% per il rapporto tra IE e FER (con una variabilità marcata, passando da valori medi intorno al 5% per regioni come il Friuli-Venezia Giulia e il Molise a valori intorno al 40-50% per la Valle d'Aosta e la Campania), e 7% per il rapporto tra IE e CMI (con riduzioni marcate e intorno al 20% per il Meridione e intorno al 5% al Nord). Le performance delle SVM ottengono in una significativa percentuale di casi risultati notevolmente superiori ai modelli tradizionali: per il rapporto tra IE e CMI, oltre il 40% dei risultati riduce l'errore del 10% o più e circa il 20% riduce l'errore di oltre il 20%; per il rapporto tra IE e FER, oltre il 60% dei risultati riduce l'RMSE del 10% o più e il 30% circa riduce di oltre il 20%.

Questa prima esplorazione, che potrebbe fornire risultati più ampi se contestualizzata alla disponibilità di lunghe serie storiche, dimostra come le macchine a vettori di supporto offrano un ottimo compromesso tra semplicità d'uso ed efficacia predittiva. Sviluppi futuri del lavoro andranno nella direzione di un'analisi più approfondita delle relazioni casuali tra variabili e nel confronto con approcci di *machine learning* alternativi a quello prescelto, quali le reti neurali e il *random forest*.

Figura 5 - RMSE della regressione della quota di energia da fonti rinnovabili rispetto all'intensità energetica, per regione. Anni 2012 - 2022. (Fonte: elaborazioni su dati Istat)

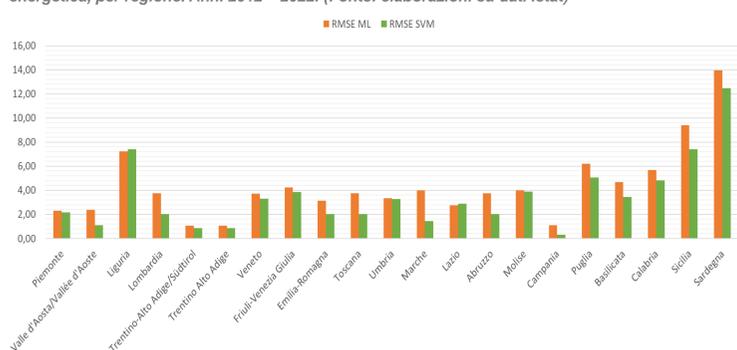
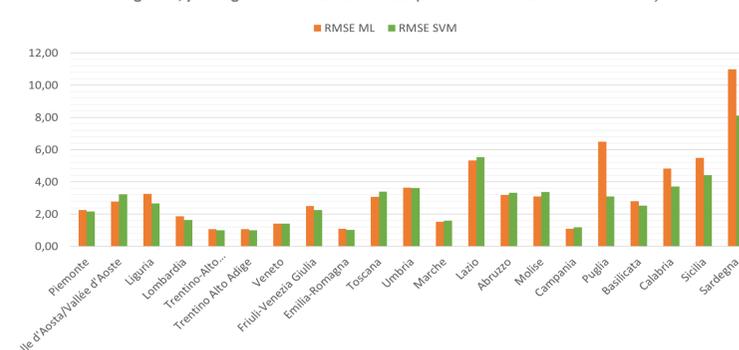


Figura 6 - RMSE della regressione del consumo di materiale interno per unità di Pil rispetto all'intensità energetica, per regione. Anni 2012 - 2022. (Fonte: elaborazioni su dati Istat)



## Bibliografia

- Istat. 2023. *Rapporto SDGs 2023. Informazioni statistiche per l'Agenda 2030 in Italia*. Roma, Istat. <https://www.istat.it/it/archivio/285778>.
- Kavitha S., Varuna S., Ramya R.. 2016. «A comparative analysis on linear regression and support vector regression», *Online International Conference on Green Engineering and Technologies (IC-GET)*. Coimbatore. pp. 1-5.
- Ghosh S., Dasgupta A., Swetapadma A.. 2019. «A Study on Support Vector Machine based Linear and Non-Linear Pattern Classification», *International Conference on Intelligent Sustainable Systems (ICISS)*. Palladam. pp. 24-28
- Petruseva S. et al.. 2017. «Construction costs forecasting: comparison of the accuracy of linear regression and support vector machine models»
- Sretenović A. et al.. 2017. «Multiple linear regression, support vector machines and neural networks for prediction of commercial building energy consumption». *Proceedings of the 46th International HVAC&R Congress*. Belgrade. pp. 383-393.