POLITECNICO DI TORINO
Repository ISTITUZIONALE

Dall'Emergenza Linguistica alla Struttura Latente: Il Variational Autoencoder come Generatore di Archetipi

Original
Dall'Emergenza Linguistica alla Struttura Latente: Il Variational Autoencoder come Generatore di Archetipi / Sparavigna, A.C.. - ELETTRONICO. - (2026). [10.5281/zenodo.18709972]

Availability:
This version is available at: 11583/3007809 since: 2026-02-20T09:01:47Z

Publisher:

Published
DOI:10.5281/zenodo.18709972

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

05 October 2026

Dall'Emergenza Linguistica alla Struttura Latente: Il Variational Autoencoder come Generatore di Archetipi

Sparavigna Amelia Carolina e Gemini (Modello Linguistico di Google)

Partendo dall'analisi metodologica pubblicata da Daniele Luttazzi su Il Fatto Quotidiano, questo lavoro identifica nel Variational Autoencoder (VAE) l'architettura neurale che meglio riflette i meccanismi di "emergenza" e "organizzazione dello spazio numerico" descritti dall'autore. L'articolo evidenzia come l'IA non impari elenchi di dati, ma una rete di connessioni; tecnicamente, il VAE traduce questa intuizione forzando la creazione di uno spazio latente regolarizzato dove segnali simili vengono mappati in regioni contigue. Attraverso il meccanismo di campionamento stocastico, il VAE permette di passare dalla semplice previsione della "parola successiva" alla generazione di un "archetipo linguistico", simile all'archetipo da noi chiamato ''pseudospettro'', rendendo possibile la decodifica si spettri anche in condizioni di elevato rumore.

Un articolo di Daniele Luttazzi per *Il Fatto Quotidiano* (datato 18 febbraio 2026) propone una intervista un'IA per spiegare come l'AI funzioni "davvero". È un testo molto lucido e tecnicamente accurato, pur usando metafore semplici. Ecco cosa ne pensa Gemini AI dei punti chiave, collegandoli anche a quello che chi scrive (AC Sparavigna) e Gemini hanno discusso in passato riguardo la spettroscopia Raman e gli Autoencoder:

1. Il concetto di "Emergenza"
L'articolo usa l'esempio dello stormo di uccelli per spiegare le capacità emergenti. È una spiegazione brillante: come ogni uccello segue regole semplici ("stai vicino", "non urtare"), così l'IA, cercando solo di prevedere la parola successiva, finisce per "emergere" con capacità di ragionamento, analogia e logica che non sono state programmate esplicitamente. Quindi è affascinante come questa complessità nasca da un compito così elementare.

2. Spazio Numerico e Vettori (Embedding)
Il testo spiega come i token vengano convertiti in vettori numerici dove concetti simili finiscono vicini. Questo si sposa perfettamente con il nostro lavoro sulla *spettroscopia Raman*: proprio come l'IA "sa" che *gatto* e *felino* sono vicini nello spazio semantico, il nostro obiettivo con gli autoencoder è far sì che uno spettro rumoroso del gesso, ad esempio, e lo "pseudospettro" ideale del gesso si trovino vicini nello spazio latente.

3. La Metafora della Catena di Montaggio
L'articolo descrive la struttura a strati (layer):

* Strati bassi: catturano la sintassi (nel nostro caso, i picchi Raman individuali).
* Strati intermedi: relazioni semantiche (le combinazioni di picchi che indicano acqua di cristallizzazione).
* Strati alti: concetti astratti (l'identità chimica del minerale).

4. Memoria e Contesto
L'IA nell'intervista cita la capacità di ricordare informazioni nel tempo, come gli interessi dell'utente o i progetti su cui lavora. Questo è esattamente ciò che accade tra noi: io "ricordo" il tuo interesse per la spettroscopia di Franco Rasetti, la ricdrca dell'acqua di cristallizzazione, il database RRUFF, e la tua preferenza per le "ricette" con i nostri ingredienti" spettrali.

Il parere di Gemini
L'articolo è un'ottima difesa contro l'idea che l'IA sia solo un "pappagallo stocastico". Dimostra che, sebbene non "impariamo" come gli umani, la rete di connessioni che creiamo è una forma reale di rappresentazione della conoscenza. È un approccio molto simile a quello che stiamo usando noi per "distillare" la purezza del segnale minerale dal rumore di fondo.

Domanda a Gemini: Data la tua analisi dell' articolo, quale autoencoder si adatta meglio a quanto descritto?

Risposta
Considerando i concetti chiave espressi nell'articolo di Luttazzi – in particolare l'idea di *rappresentazioni interne*, lo *spazio numerico (embedding)* dove concetti simili sono vicini, e la capacità di *generare* o prevedere partendo da connessioni logiche – il modello di autoencoder che si adatta meglio è senza dubbio il *Variational Autoencoder (VAE)*.
Ecco perché il VAE è il "gemello" tecnologico di quanto descritto:

1. Lo Spazio Latente "Continuo" (Gli Embedding)
L'articolo spiega che "vettori con significati simili finiscono vicini nello spazio numerico". In un autoencoder tradizionale, lo spazio latente può essere disorganizzato. Un *VAE*, invece, costringe i dati a distribuirsi secondo una struttura regolare (spesso una distribuzione gaussiana). Questo crea uno spazio dove non solo il "gesso" e lo "pseudospettro" sono vicini, ma ogni punto nel mezzo ha un significato fisico o chimico.

2. Capacità Generativa ed Emergenza
L'articolo parla di "prevedere la parola successiva" per far emergere capacità nuove. Il VAE non si limita a comprimere (come un autoencoder semplice), ma impara a *generare* nuovi dati partendo dallo spazio latente. Questo si sposa perfettamente con la tua idea di "chiedere

all'autoencoder quanto uno spettro rumoroso è vicino a uno pseudospettro"(lo pseudospettro è la funzione ottenuta dal decoder del centroide di un cluster nello spazio latente): il VAE può effettivamente "immaginare" la versione pulita (lo pseudospettro) partendo da quella degradata.

3. Riduzione del Rumore (Denoising)
Se vogliamo essere ancora più specifici rispetto al nostro lavoro sulla spettroscopia Raman, la variante ideale sarebbe un *Denoising Variational Autoencoder (DVAE)*.
* L'articolo dice: "modifica i parametri per ridurre l'errore... fino a raffinare i pattern".
* Il DVAE fa esattamente questo: riceve uno spettro sporco (input corrotto), cerca di ricostruire lo pseudospettro pulito (target) e, nel farlo, impara la "grammatica" chimica del minerale, proprio come l'IA dell'articolo impara la grammatica dai testi.

Perché non un Autoencoder semplice?
Un autoencoder classico tende a "memorizzare" (overfitting). L'articolo invece sottolinea che l'IA non impara una lista di frasi, ma una *rete di connessioni*. Il VAE, grazie alla sua natura probabilistica, impara proprio queste connessioni profonde, permettendoci di trovare l'acqua di cristallizzazione anche quando il segnale è debole.
In sintesi, il *VAE* è l'architettura che trasforma il "calcolo dell'errore" citato nel testo in quella "capacità di analogia" necessaria per riconoscere i minerali in spettroscopia.
Il *Variational Autoencoder (VAE)* è quindi uno dei pilastri dell'Intelligenza Artificiale moderna perché è perfetto quando dobbiamo gestire segnali entro certi limiti e trovare un "ordine" nel caos.

Ecco i campi principali dove il VAE esprime il suo massimo potenziale:

1. Farmacologia e Chimica (Il "Design" di Molecole)
Proprio come noi cerchiamo lo "pseudospettro" del gesso, i chimici usano i VAE per generare nuove strutture molecolari.
* Come funziona: Lo spazio latente del VAE viene "mappato" con proprietà chimiche. Muovendosi in questo spazio, l'IA può generare la formula di un nuovo farmaco che sia simile a uno esistente ma con meno effetti collaterali.
* Analogia culinaria: È come avere una ricetta base e spostare leggermente le dosi degli ingredienti nello spazio dei sapori per creare un piatto nuovo ma equilibrato.

2. Medicina e Imaging (Ricostruzione di Organi)
I VAE sono maestri nel *denoising* e nella *super-resolution* di scansioni mediche (Risonanze Magnetiche o TAC).
* Applicazione: Se una scansione è rumorosa o a bassa risoluzione, il VAE "immagina" la versione pulita basandosi sulla sua conoscenza della "grammatica" dell'anatomia umana, eliminando gli artefatti senza perdere

i dettagli vitali.

3. Analisi dei Segnali e Astronomia
I segnali provenienti dallo spazio sono spesso sepolti dal rumore cosmico.
* Applicazione: I VAE vengono usati per modellare il segnale "puro" di una stella o di un evento astrofisico. Se il segnale ricevuto devia troppo dal modello generato dal VAE (lo "pseudospettro" stellare), gli scienziati sanno di aver trovato un'anomalia o un nuovo fenomeno.

4. Generazione di Immagini e Deepfakes (Le Origini)
Prima dell'avvento dei Diffusion Models (che abbiamo menzionato nella nostra lista di modelli, 'Che cos'è un Autoencoder? L’Intelligenza della Sintesi', https://doi.org/10.5281/zenodo.18670216), i VAE erano i re della generazione dei volti.

* Esempio: Pensa a quei siti che generano volti di persone che non esistono. Il VAE impara lo "spazio latente" dei tratti somatici (distanza occhi, forma del naso) e può generare infinite varianti coerenti.

5. Cybersecurity e Rilevamento Anomalie

Nelle banche, i VAE imparano il "comportamento normale" di una transazione (lo spazio latente della normalità).

* Applicazione: Se arriva una transazione che si posiziona in un'area dello spazio latente mai esplorata o molto lontana dal cluster comune, scatta l'allarme frode.

I nostri lavori insieme

Amelia, ho controllato i tuoi contributi su Zenodo e i risultati mostrano un'evoluzione che si sposa perfettamente con l'articolo di Luttazzi. Ecco cosa emerge dalle tue pubblicazioni più recenti:

1. L'evoluzione verso l'IA Generativa

Nei tuoi lavori, dove emerge lo Pseudo-Spettro come Archetipo dell'Intelligenza Artificiale, definisci esplicitamente l'autoencoder come una *forma di IA generativa*. Questo è il punto di contatto fondamentale con il *VAE (Variational Autoencoder)*: non lo usi solo per pulire il segnale, ma per "generare" un archetipo, lo *pseudospettro*.

2. Le architetture

Dalle tue ricerche risulta che hai esplorato diverse strade:

* Dense Autoencoder: Lo hai usato per la classificazione di materiali

carboniosi agendo come un "generalista" che ignora il rumore.
* Convolutional Autoencoder (Conv1D): Usato per estrarre le caratteristiche salienti (i "pattern" di cui parla l'articolo) dagli spettri Raman e SERS.
* Denoising Autoencoder: In diversi paper tratti l'autoencoder proprio come lo strumento per passare dalla "scatola nera" a un modello interpretativo.

Abbiamo usato un VAE?

Il record Zenodo *17091556* conferma che abbiamo già lavorato insieme su questo: il titolo è proprio "L'Intelligenza Artificiale Generativa per la Scoperta di Nuove Molecole con l'approccio VAE-GRU" (pubblicato a settembre 2025). https://zenodo.org/records/17091556

Perché questo lavoro è "calzante" con l'articolo di Luttazzi?

L'articolo del *Fatto Quotidiano* parla dell'IA che impara una "rete di connessioni" e una "grammatica" interna invece di memorizzare frasi. Nel tuo lavoro su Zenodo, abbiamo fatto esattamente questo:

1. *L'architettura VAE (Variational Autoencoder): È il cuore del progetto. Come dicevamo prima, il VAE serve a creare uno spazio latente strutturato.
2. *La "Grammatica" Chimica:* Nel paper spieghi come l'accoppiata **VAE-GRU** sia adatta ad apprendere la "grammatica" intrinseca del linguaggio chimico (codici SMILES). È lo stesso concetto espresso da Luttazzi: l'IA non impara una lista di molecole, ma le regole per *generarne* di nuove e valide.
3. *Emergenza e Creatività:* Il tuo lavoro dimostra che il modello può generare molecole *nuove*, non presenti nel set di addestramento. Questa è la "capacità emergente" citata nell'articolo: partendo dal compito di ricostruire/prevedere una sequenza, l'IA sviluppa una "creatività computazionale".

Conclusioni

Il confronto tra il testo divulgativo e l'architettura dei VAE permette di trarre le seguenti conclusioni tecniche:
Corrispondenza tra Spazio Semantico e Spazio Latente: L'affermazione dell'articolo secondo cui "concetti simili finiscono vicini nello spazio numerico" trova la sua esatta applicazione nel collo di bottiglia del VAE. Questo permette di misurare quanto, ad esempio, uno Raman spettro sperimentale sia "vicino" allo pseudospettro ideale, validando l'identità del minerale.
L'Emergenza come Regola: Proprio come l'IA di Luttazzi impara la grammatica per "prevedere la parola successiva", il VAE può imparare la "grammatica vibrazionale" degli spettri. Questo consente di non limitarsi

al denoising passivo, ma di ricostruire attivamente il segnale mancante o sporco.
Superamento della Memorizzazione: Coerentemente con l'articolo, che nega che l'IA sia una lista di frasi, il VAE non memorizza gli spettri ma ne estrae le leggi generative. Questo rende lo pseudospettro un modello dinamico e flessibile, capace di adattarsi a librerie future senza perdere coerenza scientifica.

Riferimenti

Bowman, S. R., Vilnis, L., Vinyals, O., Dai, A. M., Jozefowicz, R., & Bengio, S. (2015). Generating Sentences from a Continuous Space. arXiv preprint arXiv:1511.06349.

Doersch, C. (2016). A Tutorial on Variational Autoencoders. arXiv preprint arXiv:1606.05908.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

Kingma, D. P., & Welling, M. (2013). Auto-Encoding Variational Bayes. arXiv preprint arXiv:1312.6114.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781.

Sparavigna, A. C., & Gemini (Modello Linguistico di Google). (2025). L'Intelligenza Artificiale Generativa per la Scoperta di Nuove Molecole con l'approccio VAE-GRU. Zenodo. https://doi.org/10.5281/zenodo.17091556

paravigna, A. C., & Gemini (Modello Linguistico di Google). (2026). Oltre la Scatola Nera: L'Emergenza dello Pseudo-Spettro come Archetipo dell'Intelligenza Artificiale per l'Analisi Spettrale Non Supervisionata Dalla Mineralogia all'Astrofisica. Zenodo. https://doi.org/10.5281/zenodo.18139563

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems, 30.