

KOMPUTIKA

July 2025
Issue

NEWSLETTER

The power of variational disentangled autoencoders

INSIDE

—

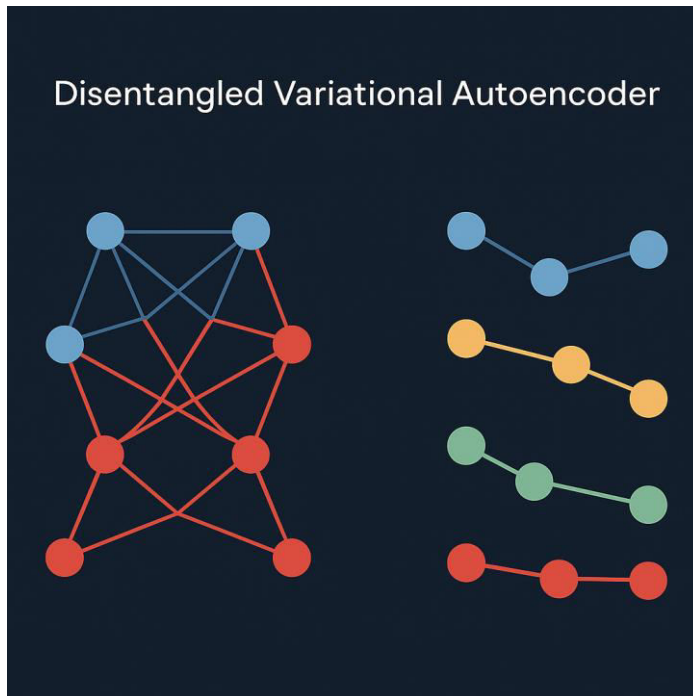
TAG

[Artificial Intelligence]
[Machine Learning]

—

AFFILIATION

Software Engineering
Department,
Faculty of Computer
Science and Information
Technology,
Universiti Malaya



Variational disentangled autoencoder

EDITED BY

Raja Jamilah Raja Yusof

—

The power of variational disentangled autoencoders

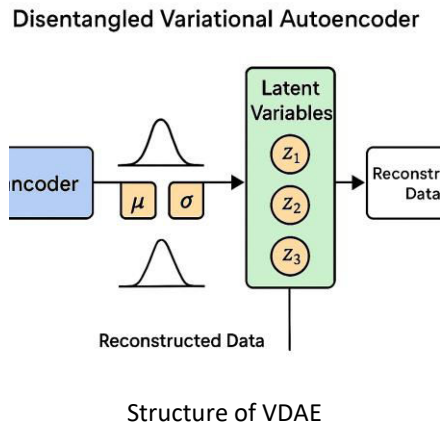
— By Pradhiksha Jogaraj, Siti Hafizah Ab Hamid

What Are Variational Disentangled Autoencoders?

Have you ever wished artificial intelligence could understand not just data, but the reasons behind it? Variational Disentangled Autoencoders offer exactly that. An approach where machines learn meaningful, independent concepts rather than tangled representations. Variational Disentangled Autoencoders (VDAEs) are an extension of Variational Autoencoders (VAEs), a popular type of generative model in deep learning. While VAEs compress input data into a latent space for reconstruction, VDAEs go one step further. They aim to disentangle this latent space, ensuring that each dimension captures a single, distinct factor such as shape, size, color, or user preference.

How Do They Work?

In traditional models, latent variables often combine multiple factors, making interpretation and control difficult. Disentangled representations allow artificial intelligence models to isolate and manipulate specific attributes—leading to more transparent, controllable, and transferable results. Encoder–Decoder Structure: The encoder transforms input data into latent codes, and the decoder reconstructs the data from them. The β -VAE introduces a weighted regularization term to encourage separation of factors. Mutual information penalties minimize overlap between latent variables. Contrastive learning helps separate related but distinct influences, such as user interest versus social behavior.



Real-World Applications

Medical Imaging: Removing unwanted elements (like hair) from skin images. Recommendation Systems : Separating user preferences from peer influence. Creative Generation: Modifying a single visual trait without affecting others.

Conclusion

Variational Disentangled Autoencoders empower artificial intelligence systems to learn in a structured, interpretable, and modular fashion. By isolating distinct factors of variation within data, they promote greater transparency, controllability, and generalization. This not only enhances the reliability of machine learning models but also brings them one step closer to understanding the world in a more human-aligned and explainable way.

For additional information you can contact the author at sitihafizah@um.edu.my from the Department of Software Engineering at Universiti Malaya.