Derivation and Discussion of Variational Autoencoder Model

Jiarui Li
Shanghai Jiao Tong University
Student ID: 522030910119

[1]School of Artificial Intelligence, Shanghai Jiao Tong University

Abstract

Variational Autoencoder (VAE) is a powerful generative model based on auto-generated-encoder. This paper presents a detailed derivation about how the VAE is derived, and discusses the advantages and disadvantages of the model. In the experiment part, I implement the VAE model on MNIST dataset, which hidden dimension are 1 *or* 2*. The minimal loss of validation set is* 000 *and* 000 *for hidden dimension* 1 *and* 2 *respectively.*

Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Praesent porttitor arcu luctus, imperdiet urna iaculis, mattis eros. Pellentesque iaculis odio vel nisl ullamcorper, nec faucibus ipsum molestie. Sed dictum nisl non aliquet porttitor. Etiam vulputate arcu dignissim, finibus sem et, viverra nisl. Aenean luctus congue massa, ut laoreet metus ornare in. Nunc fermentum nisi imperdiet lectus tincidunt vestibulum at ac elit. Nulla mattis nisl eu malesuada suscipit.

Aliquam arcu turpis, ultrices sed luctus ac, vehicula id metus. Morbi eu feugiat velit, et tempus augue. Proin ac mattis tortor. Donec tincidunt, ante rhoncus luctus semper, arcu lorem lobortis justo, nec convallis ante quam quis lectus. Aenean tincidunt sodales massa, et hendrerit tellus mattis ac. Sed non pretium nibh. Donec cursus maximus luctus. Vivamus lobortis eros et massa porta porttitor. Donec laoreet nisl vel risus lacinia elementum non nec lacus. Nullam luctus, nulla volutpat ultricies ultrices, quam massa placerat augue, ut fringilla urna lectus nec nibh. Vestibulum efficitur condimentum orci a semper. Pellentesque ut metus pretium lacus maximus semper.

Fusce varius orci ac magna dapibus porttitor. In tempor leo a neque bibendum sollicitudin. Nulla pretium fermentum nisi, eget sodales magna facilisis eu. Praesent aliquet nulla ut bibendum lacinia. Donec vel mauris vulputate, commodo ligula ut, egestas orci. Suspendisse commodo odio sed hendrerit lobortis. Donec finibus eros erat, vel ornare enim mattis et. Donec finibus dolor quis dolor tempus consequat. Mauris fringilla dui id libero egestas, ut mattis neque ornare. Ut condimentum urna pharetra ipsum consequat, eu interdum elit cursus. Vivamus scelerisque tortor et nunc ultricies, id tincidunt libero pharetra. Aliquam eu imperdiet leo. Morbi a massa volutpat velit condimentum convallis et facilisis dolor.

$$\cos^3\theta = \frac{1}{4}\cos\theta + \frac{3}{4}\cos 3\theta \tag{1}$$

Automatically referencing an equation number using its label: Equation 1.

In hac habitasse platea dictumst. Curabitur mattis elit sit amet justo luctus vestibulum. In hac habitasse platea dictumst. Pellentesque lobortis justo enim, a condimentum massa tempor eu. Ut quis nulla a quam pretium eleifend nec eu nisl. Nam cursus porttitor eros, sed luctus ligula convallis quis. Nam convallis, ligula in auctor euismod, ligula mauris fringilla tellus, et egestas mauris odio eget diam. Praesent sodales in ipsum eu dictum. Aenean vel enim ipsum. Fusce ut felis at eros sagittis bibendum mollis lobortis libero.

Derivation

PCA and VAE

At the beggining, we should find an encoder and decoder function to encode image into a low dimension, and decode it back to the original image as close as possible. We know that PCA seems to be an example of this kind of function, whose encoder is a matrix W and decoder is W^T. Can we just change the matrix W and W^T to two neural networks?

That's a great idea, but there are some questions. If we just do like this, the model will not generate new images from unknown low dimension vectors. For instance, we have low dimension vectorα_1 for full moon, and α_2 for half moon. But when I input vector $(\alpha_1 + \alpha_2)/2$ in decoder, the model will not output 3/4 moon image, which is what we want. That

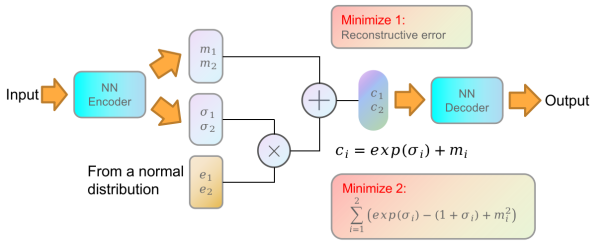

Figure 1: *VAE model*

Table 1: *Example single column table.*

Location		
East Distance	West Distance	Count
100km	200km	422
350km	1000km	1833
600km	1200km	890

is because the model only learn reconstructing image close to the original image, not generating new images from low dimension vectors.

So what VAE do to solve this problem is to add global noise distruction to the low dimension vector, so that the model will generate new images from unknown low dimension vectors. The core idea of VAE is how to add this noise distruction.

Noise Distruction

Figure 1 shows the structure of VAE which hidden dim is 2. To add proper noise distribution, the encoder should output two vectors, m_i is the origin encoded vector, and e_i aims to control the influence of each noise, and e_i is a random vector from normal distribution. So the noise is added into vector as follows 2:

$$\begin{aligned} e_i &\sim \mathcal{N}(0,1) \\ noise &= \exp(\alpha_i) \cdot e_i \\ c_i &= noise + m_i \end{aligned} \tag{2}$$

For the loss function, in addition to the existing image reconstruction loss, VAE also add a new loss funtion. If we do not add this loss function, the model will finally get a -inf in σ_i to make the noise to zero, which is not what we want. So we should make a loss to σ_i to be to small.

$$loss_2 = \sum_{i=1}^{2}(\exp(\sigma_i) - (1+\sigma_i) + m_i^2) \tag{3}$$

Results

Referencing a table using its label: Table 1.

Figure 2: *Anther of thale cress (Arabidopsis thaliana), fluorescence micrograph. Source: Heiti Paves, https://commons.wikimedia.org/wiki/File:Tolmukapea.jpg.*

Aenean feugiat pellentesque venenatis. Sed faucibus tristique tortor vel ultrices. Donec consequat tellus sapien. Nam bibendum urna mauris, eget sagittis justo gravida vel. Mauris nisi lacus, malesuada sit amet neque ut, venenatis tempor orci. Curabitur feugiat sagittis molestie. Duis euismod arcu vitae quam scelerisque facilisis. Praesent volutpat eleifend tortor, in malesuada dui egestas id. Donec finibus ac risus sed pellentesque. Donec malesuada non magna nec feugiat. Mauris eget nibh nec orci congue porttitor vitae eu erat. Sed commodo ipsum ipsum, in elementum neque gravida euismod. Cras mi lacus, pulvinar ut sapien ut, rutrum sagittis dui. Donec non est a metus varius finibus. Pellentesque rutrum pellentesque ligula, vitae accumsan nulla hendrerit ut.

Referencing a figure using its label: Figure 2.

Aenean porttitor eros non pharetra congue. Proin in odio in dolor luctus auctor ac et mi. Etiam euismod mi sed lectus fringilla pretium. Phasellus tristique maximus lectus et sodales. Mauris feugiat ligula quis semper luctus. Nam sit amet felis sed leo fermentum aliquet. Mauris arcu dui, posuere id sem eget, cursus pulvinar mi. Donec nec lacus non lectus fermentum scelerisque et at nibh. Sed tristique, metus ac vestibulum porta, tortor lectus placerat lorem, et convallis tellus dolor eget ante. Pellentesque dui ligula, hendrerit a purus et, volutpat tempor lectus. Mauris nec purus nec mauris rhoncus pellentesque. Quisque quis diam sed est lacinia congue. Donec magna est, hendrerit sed metus vel, accumsan rutrum nibh.

Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Etiam cursus lectus purus, tempus iaculis quam dictum tristique. Nam interdum sapien nec tempor mattis. Quisque id sapien nisi. Mauris vehicula ornare eros

Table 2: *Example two column table with fixed-width columns.*

Location		
East Distance	West Distance	Count
100km	200km	422
350km	1000km	1833
600km	1200km	890

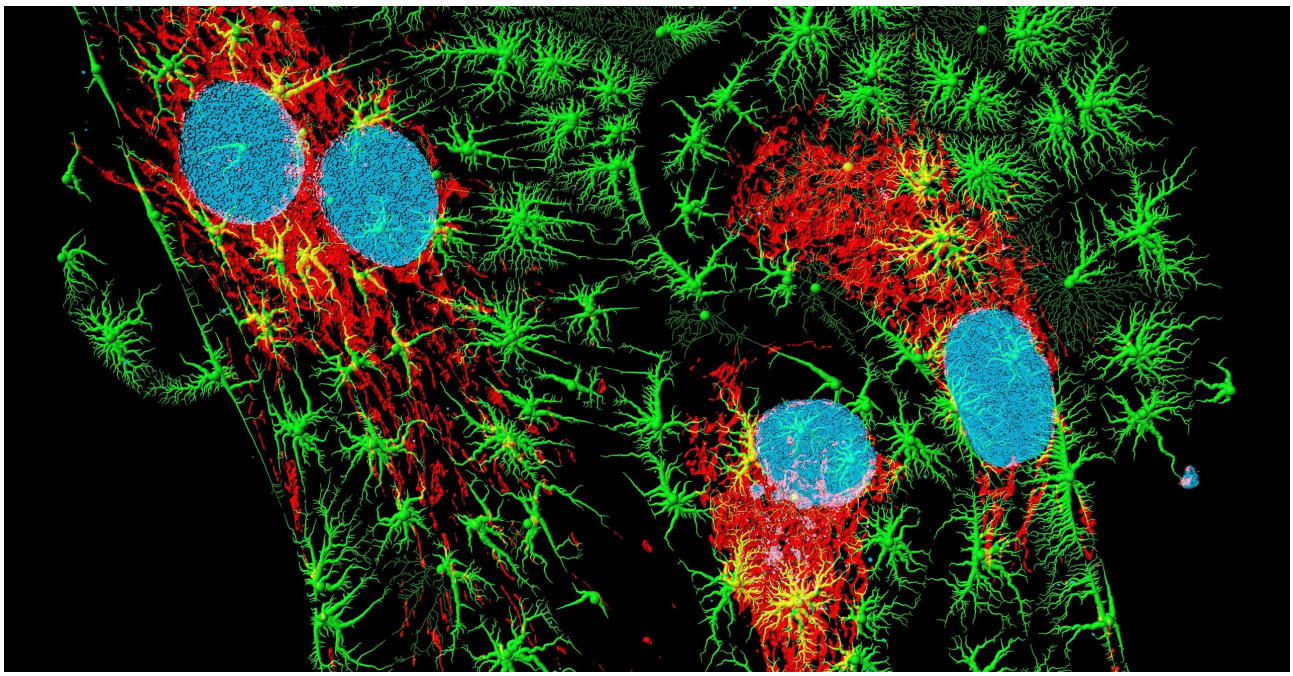

Figure 3: *Bovine pulmonary artery endothelial cells in culture. Blue: nuclei; red: mitochondria; green: microfilaments. Computer generated image from a 3D model based on a confocal laser scanning microscopy using fluorescent marker dyes. Source: Heiti Paves, https://commons.wikimedia.org/wiki/File:Fibroblastid.jpg.*

vel efficitur. Nulla consectetur, turpis quis fringilla tincidunt, mi neque iaculis lectus, vel commodo elit odio non ex. Duis facilisis, purus ac viverra iaculis, turpis lectus ultrices ante, ac vestibulum ligula magna in libero. Etiam tristique maximus lacinia. Vestibulum hendrerit, lacus malesuada laoreet blandit, sapien velit sollicitudin nunc, eu porttitor urna ligula at lorem. Aliquam faucibus eros in fermentum venenatis. Fusce consectetur congue pellentesque. Suspendisse at nisi sit amet est porttitor cursus. Cras placerat faucibus nunc, a laoreet justo dignissim sit amet.

International Support

àáâäãåèéêëìíîïòóôöõøùúûüßýñç
ÀÁÂÄÃÅÈÉÊËÌÍÎÏÒÓÔÖÕØÙÚÛÜÝÑ
SSÇÆ

Links

This is a clickable URL link: LaTeX Templates. This is a clickable email link: vel@latextemplates.com. This is a clickable monospaced URL link: `https://www.LaTeXTemplates.com`.

Discussion

This statement requires citation [1]. This statement requires multiple citations [1, 2]. This statement contains an in-text citation, for directly referring to a citation like so: Jones and Smith [2].

Subsection One

Suspendisse potenti. Vivamus suscipit dapibus metus. Proin auctor iaculis ex, id fermentum lectus dapibus tristique. Nullam maximus eros eget leo pretium dapibus. Nunc in auctor erat, id interdum risus. Suspendisse aliquet vehicula accumsan. In vestibulum efficitur dictum. Sed ultrices, libero nec fringilla feugiat, elit massa auctor ligula, vehicula tempor ligula felis in lectus. Suspendisse sem dui, pharetra ut sodales eu, suscipit sit amet felis. Donec pretium viverra ante, ac pulvinar eros. Suspendisse gravida consectetur urna. Pellentesque vitae leo porta, imperdiet eros eget, posuere sem. Praesent eget leo efficitur odio bibendum condimentum sit amet vel ex. Nunc maximus quam orci, quis pulvinar nibh eleifend ac. Quisque consequat lacus magna, eu posuere tellus iaculis ac. Sed vitae tortor tincidunt ante sagittis iaculis.

Subsection Two

Nullam mollis tellus lorem, sed congue ipsum euismod a. Donec pulvinar neque sed ligula ornare sodales. Nulla sagittis vel lectus nec laoreet. Nulla volutpat malesuada turpis at ultricies. Ut luctus velit odio, sagittis volutpat erat aliquet vel. Donec ac neque eget neque volutpat mollis. Vestibulum viverra ligula et sapien bibendum, vel vulputate ex euismod. Curabitur nec velit velit. Aliquam vulputate lorem elit, id tempus nisl finibus sit amet. Curabitur ex turpis, consequat at lectus id, imperdiet molestie augue. Curabitur eu eros molestie purus commodo hendrerit. Quisque auctor ipsum nec mauris malesuada, non fringilla nibh viverra. Quisque gravida, metus quis semper pulvinar, dolor nisl suscipit leo, vestibulum volutpat ante justo ultrices diam. Sed id facilisis turpis, et aliquet eros.

Subsubsection Example Duis venenatis eget lectus a aliquet. Integer vulputate ante suscipit felis feugiat rutrum. Aliquam eget dolor eu augue elementum ornare. Nulla fringilla interdum volutpat. Sed tincidunt, neque quis imperdiet hendrerit, turpis sapien ornare justo, ac blandit felis sem quis diam. Proin luctus urna sit amet felis tincidunt, sed congue nunc pellentesque. Ut faucibus a magna faucibus finibus. Etiam id mi euismod, auctor nisi eget, pretium metus. Proin tincidunt interdum mi non interdum. Donec semper luctus dolor at elementum. Aenean eu congue tortor, sed hendrerit magna. Quisque a dolor ante. Mauris semper id urna id gravida. Vestibulum mi tortor, finibus eu felis in, vehicula aliquam mi.

Aliquam arcu turpis, ultrices sed luctus ac, vehicula id metus. Morbi eu feugiat velit, et tempus augue. Proin ac mattis tortor. Donec tincidunt, ante rhoncus luctus semper, arcu lorem lobortis justo, nec convallis ante quam quis lectus. Aenean tincidunt sodales massa, et hendrerit tellus mattis ac. Sed non pretium nibh.

Donec cursus maximus luctus. Vivamus lobortis eros et massa porta porttitor. Nam vitae suscipit mi. Pellentesque ex tellus, iaculis vel libero at, cursus pretium sapien. Curabitur accumsan velit sit amet nulla lobortis, ut pretium ex aliquam. Proin eget volutpat orci. Morbi eu aliquet turpis. Vivamus molestie urna quis tempor tristique. Proin hendrerit sem nec tempor sollicitudin.

References

[1] J. M. Smith and A. B. Jones. *Book Title*. 7th. Publisher, 2023.

[2] A. B. Jones and J. M. Smith. "Article Title". In: *Journal title* 13.52 (Mar. 2024), pp. 123–456. DOI: 10.1038/s41586-021-03616-x.