TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**NGUYỄN MINH HOÀNG CHƯƠNG - 52000744**

**BÁO CÁO CUỐI KỲ**

**AUTODECODER VÀ**

**VARIATIONAL AUTODECODER**

**NHẬP MÔN THỊ GIÁC MÁY TÍNH**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024**

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**NGUYỄN MINH HOÀNG CHƯƠNG - 52000744**

**BÁO CÁO CUỐI KỲ**

**AUTODECODER VÀ**

**VARIATIONAL AUTODECODER**

**NHẬP MÔN THỊ GIÁC MÁY TÍNH**

Người hướng dẫn

**TS. PHẠM VĂN HUY**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024**

# LỜI CẢM ƠN

Trước tiên, em xin gửi lời cảm ơn chân thành và lòng biết ơn sâu sắc đến TS. Phạm Văn Huy. Thầy là người đã luôn hỗ trợ và đóng góp tận tình cho chúng em trong suốt quá trình giảng dạy và là nền tảng để tôi hoàn thành môn học Nhập môn Thị giác máy tính.

Tiếp theo, em xin gửi lời cảm ơn đặc biệt đến Khoa Công Nghệ Thông Tin trường Đại học Tôn Đức Thắng vì đã tạo điều kiện cho chúng em được học tập và nghiên cứu môn học này. Khoa đã luôn sẵn sàng chia sẻ các kiến thức bổ ích cùng như chia sẻ các kinh nghiệm, tham khảo tài liệu, giúp ích không chỉ cho việc thực hiện và hoàn thành đề tài nghiên cứu mà còn giúp ích cho việc học tập và rèn luyện trong quá trình thực hành tại trường Đại học Tôn Đức Thắng nói chung.

Cuối cùng, nhờ những góp ý từ thầy cô và các bạn, em sẽ hoàn thành tốt hơn những bài nghiên cứu trong tương lai. Tôi mong quý thầy cô và các bạn cũng như mọi người luôn quan tâm và hỗ trợ em, luôn tràn đầy sức khỏe và bình an.

*TP. Hồ Chí Minh, ngày ... tháng … năm 2024*

*Tác giả*

*(Ký tên và ghi rõ họ tên)*

**CÔNG TRÌNH ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi và được sự hướng dẫn khoa học của TS. Phạm Văn Huy. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong Báo cáo còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung Báo cáo của mình**. Trường Đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày … tháng … năm 20..*

*Tác giả*

*(Ký tên và ghi rõ họ tên)*

# AUTO-ENCODER VÀ VARIATIONAD AUTO-ENCODER

# TÓM TẮT

# AUTO-ENCODER and VARIATIONAD AUTO-ENCODER

**ABSTRACT**

# MỤC LỤC

[LỜI CẢM ƠN i](#_Toc160969901)

[AUTO-ENCODER VÀ VARIATIONAD AUTO-ENCODER iii](#_Toc160969902)

[TÓM TẮT iii](#_Toc160969903)

[AUTO-ENCODER and VARIATIONAD AUTO-ENCODER iv](#_Toc160969904)

[MỤC LỤC v](#_Toc160969905)

[DANH MỤC HÌNH VẼ vii](#_Toc160969906)

[DANH MỤC BẢNG BIỂU viii](#_Toc160969907)

[CHƯƠNG 1: MỞ ĐẦU VÀ TỔNG QUAN ĐỀ TÀI 1](#_Toc160969908)

[1.1 Đặt vấn đề 1](#_Toc160969909)

[1.1.1 Lý do chọn đề tài 1](#_Toc160969910)

[1.1.2 Mục tiêu nguyên cứu 1](#_Toc160969911)

[1.1.3 Đối tượng và phạm vi nguyên cứu 1](#_Toc160969912)

[1.1.4 Phương pháp nguyên cứu 1](#_Toc160969913)

[1.1.5 Ý nghĩa khoa học và thực tiễn của đề tài 1](#_Toc160969914)

[CHƯƠNG 2 : CƠ SỞ LÝ THUYẾT 1](#_Toc160969915)

[2.1 Mô hình Auto-Encoder 1](#_Toc160969918)

[2.1.1 Auto-Encoder là gì? 1](#_Toc160969919)

[2.1.2 Latent space 2](#_Toc160969920)

[2.1.3 Generative model 2](#_Toc160969921)

[2.2 Mô hình Variational Auto-Encoder 2](#_Toc160969922)

[2.2.1 What is it? 2](#_Toc160969923)

[2.2.2 Block Diagram 2](#_Toc160969924)

[2.2.3 Loss function 2](#_Toc160969925)

[2.2.4 So sánh mô hình AE và VAE 4](#_Toc160969926)

[CHƯƠNG 3: MÔ HÌNH ĐƯỢC SỬ DỤNG TRONG BÁO CÁO 6](#_Toc160969927)

[CHƯƠNG 4: KẾT LUẬN 7](#_Toc160969928)

[2.3 Kết quả đạt được 7](#_Toc160969929)

[2.4 Hạn chế của phương pháp giải quyết bài toán 7](#_Toc160969930)

[2.5 Hướng phát triển trong tương lai 7](#_Toc160969931)

[TÀI LIỆU THAM KHẢO 8](#_Toc160969932)

# DANH MỤC HÌNH VẼ

# DANH MỤC BẢNG BIỂU

# CHƯƠNG 1: MỞ ĐẦU VÀ TỔNG QUAN ĐỀ TÀI

## Đặt vấn đề

### Lý do chọn đề tài

AE và VAE có khả năng học biểu diễn nén của dữ liệu, hỗ trợ trong nhiều ứng dụng như giảm chiều dữ liệu, tạo sinh dữ liệu, phát hiện bất thường, và hơn thế nữa.

Phát triển mới trong học sâu các tiến bộ gần đây trong AE và VAE, bao gồm các cải tiến về mô hình và kỹ thuật huấn luyện, mở ra cơ hội nghiên cứu mới.

### Mục tiêu nguyên cứu

### Đối tượng và phạm vi nguyên cứu

### Phương pháp nguyên cứu

### Ý nghĩa khoa học và thực tiễn của đề tài

# CHƯƠNG 2 : CƠ SỞ LÝ THUYẾT



## Mô hình Auto-Encoder

### Auto-Encoder là gì?

Encoder

This defines the approximate posterior distribution q(z|x), which takes as input an observation and outputs a set of parameters for specifying the conditional distribution of the latent representation z. In this example, simply model the distribution as a diagonal Gaussian, and the network outputs the mean and log-variance parameters of a factorized Gaussian. Output log-variance instead of the variance directly for numerical stability.

Decoder

This defines the conditional distribution of the observation q(x|z), which takes a latent sample as input and outputs the parameters for a conditional distribution of the observation. Model the latent distribution prior P(z) as a unit Gaussian.

### Latent space

### Generative model

## Mô hình Variational Auto-Encoder

### What is it?

Variational autoencoder giải quyết vấn đề về non-regularized latent không chính quy trong autoencoder và cung cấp khả năng tạo cho toàn bộ không gian. Encoder trong AE xuất ra các outputs latent vectors. Thay vì xuất ra các vectơ trong không gian latent , bộ mã hóa VAE đưa ra các tham số của phân bố được xác định trước trong không gian tiềm ẩn cho mỗi đầu vào. Sau đó, VAE áp đặt một ràng buộc đối với phân phối tiềm ẩn này, buộc nó phải là phân phối chuẩn(normal distribution). Ràng buộc này đảm bảo rằng không gian tiềm ẩn được chuẩn hóa.

### Block Diagram

The block diagram of VAE can be seen below. During training, the input data x is fed to the encoder function e\_theta(x). Just like AE, the input is passed through a series of layers (parameterized by the variable theta) reducing its dimensions to achieve a compressed latent vector z. However, the latent vector is not the output of the encoder. Instead, the encoder outputs the mean and the standard deviation for each latent variable. The latent vector is then sampled from this mean and standard deviation which is then fed to the decoder to reconstruct the input. The decoder in the VAE works similarly to the one in AE.

### Loss function

The loss function is defined by the VAE objectives. VAE has two objectives:

1. Reconstruct the input
2. Latent space should be normally distributed

Hence the training loss of VAE is defined as the sum of these the reconstruction loss and the similarity loss. The reconstruction error, just like in AE, is the mean squared loss of the input and reconstructed output. The similarity loss is the KL divergence between the latent space distribution and standard gaussian (zero mean and unit variance). The loss function is then the sum of these two losses.

Variational Autoencoder loss function — Image by Author

As mentioned before, the latent vector is sampled from the encoder-generated distribution before feeding it to the decoder. This random sampling makes it difficult for backpropagation to happen for the encoder since we can’t trace back errors due to this random sampling. Hence we use a reparameterization trick to model the sampling process which makes it possible for the errors to propagate through the network. The latent vector z is represented as a function of the encoder’s output.

The reparameterization trick is used to represent the latent vector z as a function of the encoder’s output.

Latent space visualization

The training tries to find a balance between the two losses and ends up with a latent space distribution that looks like the unit norm with clusters grouping similar input data points. The unit norm condition makes sure that the latent space is evenly spread out and does not have significant gaps between clusters. In fact, the clusters of similar-looking data inputs usually overlap in some regions. Below is an example of the latent space generated by training the network on the same MNIST dataset, as was used to visualize the latent space of AE. Note how there are no gaps between clusters and the space resembles the distribution of unit norm.

Latent space of Variational Autoencoder — Image by Author

An important thing to note is that when the latent vector is sampled from the regions with overlapping clusters, we get morphed data. We get a smooth transition between the decoder's output when we sample the latent space moving from one cluster to the other.

### So sánh mô hình AE và VAE

|  |  |  |
| --- | --- | --- |
| Tiêu Chí | Autoencoder (AE) | Variational Autoencoder (VAE) |
| Mục Tiêu | Giảm kích thước và tái tạo dữ liệu. | Biểu diễn dữ liệu trong không gian latent có ý nghĩa và tạo dữ liệu. |
| Khả năng tái tạo dữ liệu | Hạn chế, không gian latent không được khuyến khích để mượt mà hoặc liên tục. | Cao, không gian latent được tổ chức tốt, hỗ trợ tạo sinh dữ liệu mới một cách hiệu quả. |
| Latent Space | Thường không được quy định rõ ràng. | Được điều chỉnh để gần với phân phối xác suất chuẩn, tạo điều kiện thuận lợi cho việc lấy mẫu và nội suy. |
| Ứng Dụng | Nén dữ liệu, giảm kích thước nhưng vẫn giữ được thông tin quan trọng. Detection exeption. Loại Bỏ Nhiễu (Denoising) | New data generation, học không giám sát, tối ưu hóa transform. |
| Khi nào nên sử dụng | Khi cần giảm kích thước của dữ liệu mà vẫn giữ được độ chính xác cao trong việc tái tạo dữ liệu.  Khi cần phát hiện ngoại lệ hoặc nhiễu trong hình ảnh một cách hiệu quả và chính xác. | Khi cần tạo dữ liệu mới và khám phá các biểu diễn phân phối xác suất của dữ liệu.  Khi cần nghiên cứu sâu hơn về cấu trúc và tính chất của dữ liệu, như việc tạo ra các biểu diễn phân hối xác suất của dữ liệu.  Khi cần thực hiện các biến đổi phức tạp hoặc tối ưu hóa dựa trên không gian latent của dữ liệu. |

# CHƯƠNG 3: MÔ HÌNH ĐƯỢC SỬ DỤNG TRONG BÁO CÁO

# CHƯƠNG 4: KẾT LUẬN

## Kết quả đạt được

## Hạn chế của phương pháp giải quyết bài toán

## Hướng phát triển trong tương lai

# TÀI LIỆU THAM KHẢO