

Universidad Veracruzana

FACULTAD DE NEGOCIOS Y TECNOLOGÍAS

VARIATIONAL AUTOENCODERS (VAE)

Servicio Social

Autor:

Fhatima Reyes Alejandre
Juan David Delgado Muñoz

Jueves 29 de Febrero 2024

1 Introducción

El uso de la Inteligencia Artificial (IA) se hace más presente en nuestros días en campos como finanzas, salud, la industria, agricultura, etc. Por lo que entender como funciona esta tecnología emergente es de gran importancia.

En el presente documento trataremos sobre los Codificadores Automaticos Variacionales (Variational Autoencoders, VAEs), los cuales son modelos de aprendizaje no supervisado que tienen como principal función generar nueva información apartir de datos existentes.

1.1 Objetivos

Al finalizar la lectura se espera que puedas comprender lo siguiente:

1. Que son los Autoencoders y como funcionan.
2. Componentes que conforman un Autoencoder (Encoder, Bottleneck, y Decoder)
3. Como funciona un modelo generativo VAE.
4. El Espacio Latente y la importancia de su regulariación.

2 Autoencoders

2.1 ¿Qué son los Autoencoders?

Para poder entender de una mejor manera los modelos generativos VAE, es necesario conocer el concepto de Autoencoders.

Los Autoencoders son un tipo de red neuronal artificial cuya función principal es aprender de manera eficiente codificaciones o patrones de un conjunto de datos no etiquetados mediante la compresión de estos, con la finalidad de reducir la dimensionalidad de las características presentes para posteriormente poder ser reconstruidos, permitiendo la eliminación de ruido en los datos, así como la detección de anomalías.

2.2 ¿Cómo esta conformado un Autoencoder?

Ahora que sabemos cual es el el obojtivo de un autoencoder, veamos las partes que lo conforman.

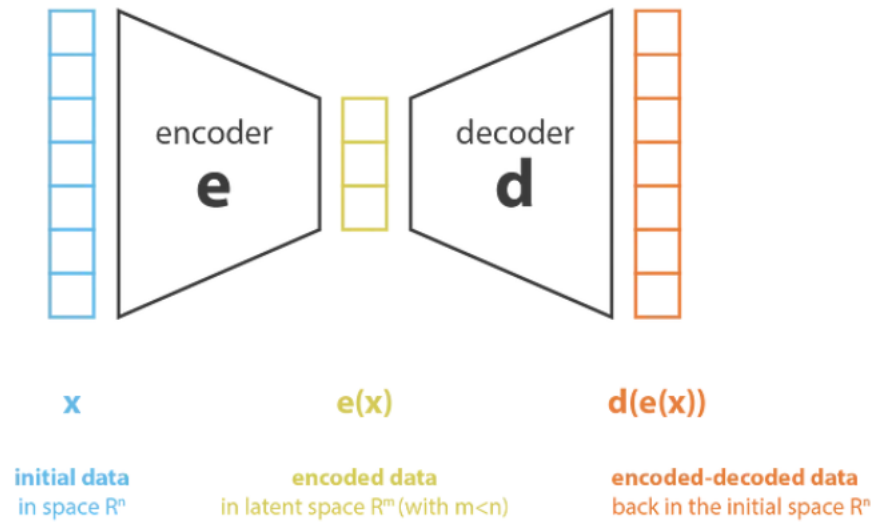


Figure 1: Estructura de un Autoencoder (Rocca, 2019).

Como se observa en la imagen, un Autoencoder esta formado principalmente por un codificador (Encoder), el espacio latente y el decodificador (Decoder), cada uno con una función especifica:

- **Encoder:** Es el proceso mediante el cual se producen nuevas características, las cuales son representaciones de las características de entrada, esta obtención se obtiene mediante selección o extracción. Es importante mencionar que durante este proceso una parte de la información se perdiera, por lo que no será posible su reconstrucción.
- **Espacio Latente:** La reducción dimensional producida al comprimir los datos por parte del encoder genera como resultado el "espacio latente", es aqui donde se encuantra la información más relevante de los datos de entrada.
- **Decoder:** Realiza el proceso inverso de un encoder, toma el dato comprimido para reconstruirlo tratando de obtener como salida el dato de entrada original.

El objetivo de realizar la reducción de dimensionalidad es encontrar un encoder-decoder que permita mantener el máximo de información durante el proceso de codificación, y lograr el mínimo valor para el error de reconstrucción a la hora de decodificar, a este proceso también se le conoce como cuello de botella (bottleneck).

3 VAEs

Por sí solos, los Autoencoders presentan problemas para la generación de nueva información, ya que al comprimir los datos de entrada no se realizan procesos de regularización (Este será un punto tratado más adelante).

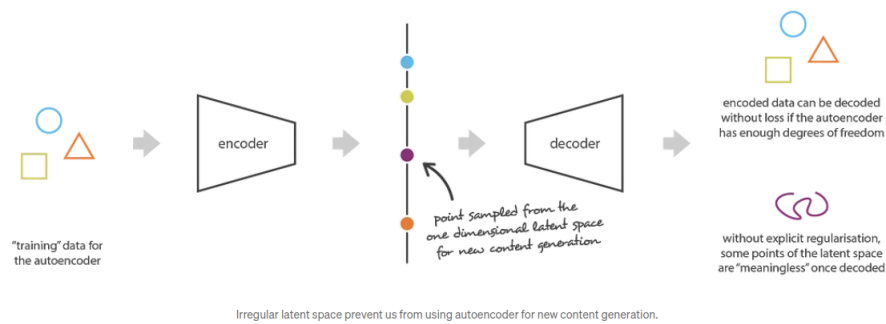


Figure 2: Limitación de un Autoencoder (Rocca, 2019).

Para resolver este problema, es donde entra VAE, El cual es un autoencoder, que permiten una regularización del espacio latente con la finalidad de evitar el sobreentrenamiento (overfitting), garantizando así que existan buenas propiedades para el proceso generativo.

El proceso de aprendizaje de un VAE es el siguiente:

- El decoder codifica la entrada como una distribución de puntos sobre el espacio latente.
- A partir de la distribución obtenida, se toma un punto como muestra.
- La muestra tomada se decodifica y calcula el error de reconstrucción.
- Finalmente, el error es propagado a través de la red.

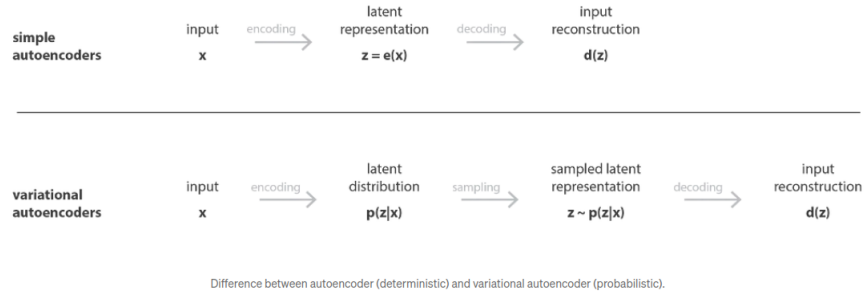


Figure 3: Diferencias entre un Autoencoder y un Variational Autoencoder.

3.1 Regularización del Espacio Latente

Como se menciona en el punto anterior, VAE presenta mejores funciones generativas gracias a la regularización del espacio latente. Esta regularización se basa en dos conceptos principales: La **continuidad** y la **completitud**. Luego entonces, se dice que el espacio latente es **continuo** si dos puntos cercanos entre si generan una decodificación similar. También, el espacio latente tendrá **completitud** si un punto muestreado genera información significativa cuando sea decodificado.

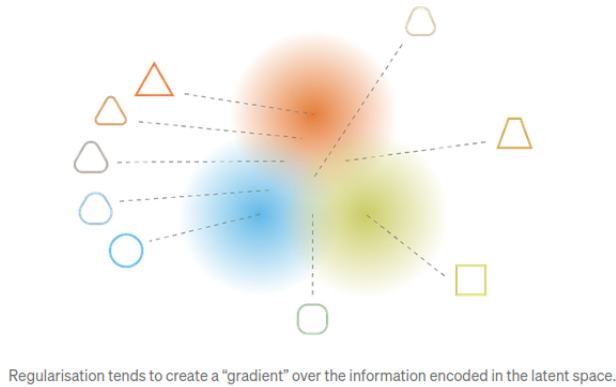


Figure 4: Regularización del espacio latente (Rocca, 2019).

4 Bibliografías

- Rocca, J. (2019, 23 septiembre). Understanding Variational Autoencoders (VAES) - towards data science. Medium. Recuperado 29 de febrero de 2024, de <https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73>
- Sotaquirá, M. (2019, 11 mayo). Autoencoders: explicación y tutorial en Python — Codificando Bits. Codificando Bits. Recuperado 29 de febrero de 2024, de <https://www.codificandobits.com/blog/autoencoders-explicacion-y-tutorial-python/>
- Rivera, M. (2018, diciembre). Autocodificadores Variacionales. Recuperado 29 de febrero de 2024, de http://personal.cimat.mx:8181/mrivera/cursos/aprendizaje_profundo/vae/vae.html