

El espacio latente en modelos generativos: explorando las dimensiones ocultas

Uno de los conceptos fundamentales en los modelos generativos es el **espacio latente**. Este concepto se refiere a una representación de dimensiones ocultas que captura aspectos significativos de los datos. En esencia, el espacio latente es un espacio abstracto en el que se proyectan los datos generados por el modelo. A diferencia del espacio de características original de los datos, el espacio latente carece de una interpretación directa y puede albergar información abstracta y altamente comprimida. Cada punto dentro de este espacio latente representa una configuración particular de las variables latentes del modelo y puede corresponder a una instancia única de los datos generados.

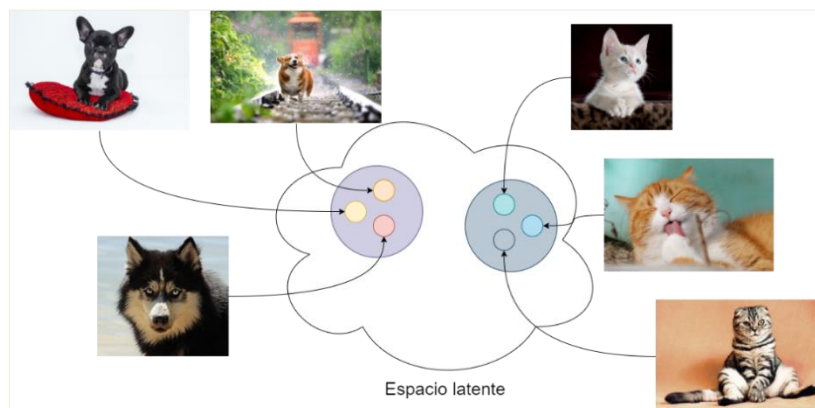


Figura 1. Representación gráfica de la estructura del espacio latente en una base de datos que incluye imágenes de perros y gatos.

Supongamos que disponemos de una base de datos que contiene imágenes de perros y gatos. Cuando deseamos emplear un modelo generativo para crear imágenes a partir de estas dos categorías de datos, el concepto del espacio latente se vuelve fundamental. Este espacio latente debe proporcionar representaciones de la información que permitan distinguir elementos de cada distribución. En un escenario ideal, se visualizaría algo similar a lo representado en la **Figura 1**.

El espacio latente en modelos generativos

Como se puede observar, todas las imágenes de perros se mapean a una región específica en el espacio latente, mientras que las imágenes de gatos ocupan una región separada. Este ejemplo ilustra cómo el espacio latente organiza la información semántica contenida en los datos, de manera que las imágenes que comparten similitudes semánticas se ubican cercanas entre sí. Este enfoque facilita la diferenciación entre las distintas categorías de datos y esencialmente permite al modelo generativo aprender y generar imágenes coherentes y pertinentes a partir de las distribuciones de categorías, en este caso de imágenes de perros y gatos.

Dependiendo del modelo generativo, el espacio latente, a menudo denotado como z , exhibe un comportamiento específico. En el contexto de las Redes Generativas Adversarias (GANs), el espacio latente se utiliza directamente por el generador, elemento de esta arquitectura que se encarga de generar nuevos datos. Al mismo tiempo, el discriminador evalúa si los datos son reales o falsos en un juego estratégico. A medida que este proceso de competencia se repite durante el entrenamiento, el espacio latente se ajusta de manera que los valores dentro de él comienzan a representar características semánticas de los datos. En la **Figura 2** se ilustra una representación gráfica del espacio latente en el contexto de la arquitectura de las GANs.

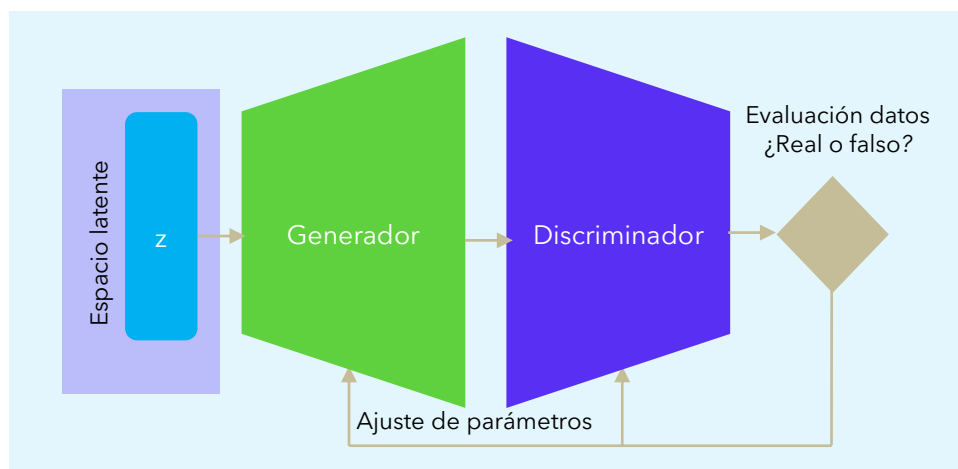


Figura 2. Diagrama de la representación de una arquitectura GAN incluyendo el espacio latente.

En los modelos probabilísticos de difusión, el espacio latente actúa como un intermediario que ayuda a transformar gradualmente los datos originales en nuevos datos. Este espacio se encuentra

en el intermedio del proceso de difusión y, por lo general, tiene una menor dimensión que los datos de entrada.

Dentro de las ventajas del espacio latente se destaca la capacidad para generar muestras nuevas y variadas. Al explorar diferentes regiones del espacio latente, es posible obtener diferentes ejemplos generados. Por ejemplo, en un modelo generativo de imágenes de caras humanas, moviéndose a lo largo de una dimensión específica del espacio latente, se podría cambiar el color del cabello de los rostros generados, mientras que en otra dimensión se podría variar la expresión facial.

El espacio latente también permite la interpolación suave entre diferentes muestras. Al trazar una línea recta en el espacio latente y generar muestras en cada punto a lo largo de esa línea, se pueden obtener transiciones graduales y continuas entre las características correspondientes en los datos generados. Esto es especialmente útil en aplicaciones como la generación de imágenes o la síntesis de voz, donde se desea generar ejemplos que se transformen de manera suave y natural.

La comprensión del espacio latente y la exploración de sus dimensiones ocultas pueden proporcionar conocimientos valiosos sobre los datos y las características que los modelos generativos han aprendido. Al visualizar el espacio latente o analizar las direcciones y los cambios en las dimensiones, los investigadores pueden descubrir regularidades, tendencias y características latentes que subyacen en los datos generados. Sin embargo, la representación del espacio latente puede tener sus desafíos. Por un lado, la interpretación exacta de las dimensiones del espacio latente puede ser difícil, ya que a menudo no se corresponden directamente con características específicas de los datos.

Conclusión

El espacio latente en los modelos generativos es una representación abstracta que captura características ocultas y abstractas de los datos generados. Explorar el espacio latente puede proporcionar conocimientos sobre las características subyacentes y las tendencias en los datos generados. Aunque la interpretación y el análisis del espacio latente pueden presentar desafíos, su comprensión es fundamental para aprovechar todo el potencial en la generación de contenido y la comprensión de los datos.

Bibliografía

Asperti, A., Tonelli, V. Comparing the latent space of generative models. *Neural Comput & Applic* 35, 3155-3172 (2023). <https://doi.org/10.1007/s00521-022-07890-2>



© - **Derechos Reservados:** la presente obra, y en general todos sus contenidos, se encuentran protegidos por las normas internacionales y nacionales vigentes sobre propiedad Intelectual, por lo tanto su utilización parcial o total, reproducción, comunicación pública, transformación, distribución, alquiler, préstamo público e importación, total o parcial, en todo o en parte, en formato impreso o digital y en cualquier formato conocido o por conocer, se encuentran prohibidos, y solo serán lícitos en la medida en que se cuente con la autorización previa y expresa por escrito de la Universidad de los Andes.

De igual manera, la utilización de la imagen de las personas, docentes o estudiantes, sin su previa autorización está expresamente prohibida. En caso de incumplirse con lo mencionado, se procederá de conformidad con los reglamentos y políticas de la universidad, sin perjuicio de las demás acciones legales aplicables.
