

En el campo de la generación de imágenes mediante modelos generativos, existen diversas métricas utilizadas para evaluar la calidad y diversidad de las imágenes generadas. Una de estas métricas ampliamente utilizadas es el Inception Score.

El Inception Score se basa en la idea de emplear una red neuronal convolucional pre-entrenada, como la arquitectura Inception, para analizar y clasificar las imágenes generadas. Al utilizar esta red, se extraen características significativas de las imágenes y se calcula la distribución de las clases correspondientes.

El cálculo del Inception Score implica evaluar tanto la calidad como la diversidad de las imágenes generadas. La calidad se determina a través de la baja entropía en las distribuciones de clases, lo cual indica que las imágenes son clasificadas con mayor seguridad en categorías específicas. Por otro lado, la diversidad se evalúa mediante la alta entropía en las distribuciones de clases, lo que indica una mayor variabilidad en las imágenes generadas.

El Inception Score proporciona una medida cuantitativa de la capacidad del modelo generativo para generar imágenes de alta calidad y diversidad. Sin embargo, es importante tener en cuenta que esta métrica tiene sus limitaciones y críticas. No es una métrica absoluta y puede no capturar completamente la calidad perceptual de las imágenes generadas.

Otras medidas similares al Inception Score son las siguientes:

- SSIM (Structural Similarity Index): El SSIM es una medida sofisticada que evalúa la similitud estructural entre dos imágenes, teniendo en cuenta elementos como el contraste, la luminancia y la estructura. Se basa en la premisa de que las imágenes de alta calidad deben ser perceptualmente similares en términos de su estructura. Un alto valor de SSIM indica una mayor similitud estructural entre las imágenes comparadas.
- LPIPS (Learned Perceptual Image Patch Similarity): El LPIPS es una medida que incorpora la percepción humana al evaluar la similitud entre imágenes. Utiliza una red neuronal pre-entrenada para calcular la distancia entre características perceptuales de las imágenes. Al combinar el conocimiento aprendido por la red neuronal con la sensibilidad humana, el LPIPS proporciona una evaluación refinada de la similitud perceptual entre imágenes.
- FID (Frechet Inception Distance): El FID es una métrica avanzada que utiliza una red neuronal convolucional pre-entrenada, como Inception, para extraer características de las imágenes generadas y reales. Luego, calcula la distancia de Frechet entre las distribuciones de características. Una distancia de Frechet más baja indica una mayor similitud entre las distribuciones y, por lo tanto, una mejor calidad y coherencia de las imágenes generadas.

- **Precision and Recall:** Estas medidas, comúnmente utilizadas en el campo del procesamiento de imágenes, evalúan la capacidad de un modelo generativo para generar imágenes realistas y de alta calidad. La precisión se refiere a la proporción de imágenes generadas correctamente clasificadas como "real" por un clasificador pre-entrenado. El recall se refiere a la proporción de imágenes reales correctamente clasificadas como "real". Un equilibrio óptimo entre alta precisión y alto recall indica una generación de imágenes de alta calidad y realismo.
- **Perceptual Path Length:** Esta medida sofisticada evalúa la suavidad y la coherencia de las transiciones entre imágenes generadas en el espacio latente de un modelo generativo. La longitud del camino perceptual cuantifica la distancia perceptual entre imágenes a lo largo de una trayectoria en el espacio latente. Una longitud de camino perceptual más baja indica transiciones suaves y una mayor coherencia en las características visuales de las imágenes generadas.

Estas medidas adicionales ofrecen diferentes perspectivas y enfoques para evaluar la calidad, la similitud perceptual y la coherencia de las imágenes generadas. Al utilizar una combinación de estas medidas, es posible obtener una evaluación más completa y sofisticada de los modelos generativos y las imágenes que generan.