

Generativní modely: GAN vs VAE

Generativní modely jsou třídou modelů strojového učení, které se nesnaží data pouze klasifikovat (např. „toto je pes“), ale snaží se naučit základní distribuci dat tak, aby dokázaly generovat **nové vzorky**, které vypadají jako ty původní.

Dvěma nejvýznamnějšími architekturami v této oblasti jsou **VAE** (Variational Autoencoders) a **GAN** (Generative Adversarial Networks).

1. VAE (Variational Autoencoders)

VAE vycházejí z klasických **autoenkodérů**, ale přidávají do latentního prostoru (bottlenecku) pravděpodobnostní prvek.

- **Princip:** Místo toho, aby enkodér namapoval vstup na jeden pevný bod, namapuje jej na **pravděpodobnostní rozdělení** (střední hodnotu a rozptyl).
- **Generování:** Nová data vznikají tak, že náhodně vybereme bod z tohoto rozdělení a dekodér jej „přeloží“ do podoby obrázku či zvuku.
- **Výhoda:** Latentní prostor je spojitý a organizovaný. Můžete v něm „cestovat“ a plynule měnit vlastnosti (např. postupně přidávat úsměv na tváři).
- **Nevýhoda:** Výstupy bývají často mírně rozmazané (**blurry**).

2. GAN (Generative Adversarial Networks)

GAN, které představil Ian Goodfellow v roce 2014, fungují na principu **souboje dvou neuronových sítí**.

Architektura GAN:

- **Generátor:** Snaží se vytvořit falešný vzorek (např. obrázek obličeje) z náhodného šumu. Jeho cílem je oklamat Diskriminátor.
- **Diskriminátor:** Funguje jako „policista“ nebo „znalec“. Dostává střídavě pravé obrázky z databáze a falešné obrázky od Generátoru. Jeho úkolem je poznat, co je pravé a co podvrh.

Proces učení (Minimax game):

Během tréninku se obě sítě neustále zlepšují. Generátor se učí vytvářet stále realističtější falzifikáty, zatímco Diskriminátor se učí lépe odhalovat i drobné chyby. Nakonec je Generátor tak dobrý, že Diskriminátor již nedokáže rozdíl poznat.

3. Srovnání GAN vs. VAE

Vlastnost	VAE	GAN
Princip	Statistická rekonstrukce	Souboj dvou sítí (Hra s nulovým součtem)
Kvalita výstupu	Často rozmazané	Velmi ostré a realistické
Stabilita učení	Stabilní, snadno se trénuje	Nestabilní, citlivé na nastavení parametrů
Latentní prostor	Spojité, snadno interpretovatelný	Často nekonzistentní a těžko ovladatelný
Využití	Kompresce, augmentace dat, anomálie	Deepfakes, generování umění, super-rozlišení

[Image comparison of images generated by VAE vs GAN showing blurriness vs sharpness]

4. Praktické aplikace

- **Deepfakes:** Výměna obličejů ve videu (primárně GAN).
- **Image-to-Image translation:** Převod satelitních snímků na mapy nebo černobílých fotek na barevné.
- **Zvýšení rozlišení (Super-Resolution):** Rekonstrukce detailů v nekvalitních fotografiích.
- **Generování molekul:** Návrh nových léků s požadovanými vlastnostmi (často VAE).

5. Co přišlo po nich?

Dnes jsou GAN i VAE doplňovány nebo nahrazovány **Difúzními modely** (Diffusion Models), které stojí za nástroji jako **DALL-E**, **Midjourney** nebo **Stable Diffusion**. Ty fungují na principu postupného odstraňování šumu z obrazu.

Související články:

- [Difúzní modely \(Stable Diffusion\)](#)
- [Neuronové sítě](#)
- [Autoenkodéry](#)

Tagy: *ai neural_networks gan vae generative_models deep_learning*

From:
<https://serviceit.cz/> - IT ENCYKLOPEDIE

Permanent link:
https://serviceit.cz/doku.php?id=it:ai:generative_models

Last update: **2026/01/02 12:44**

