

Autoencoder

Autoencoder je typ umělé **neuronové sítě**, která se učí kopírovat svůj vstup na svůj výstup. Ačkoliv to zní triviálně, síť je navržena tak, aby uprostřed obsahovala „úzké hrdlo“, které ji nutí data komprimovat a naučit se pouze jejich nejpodstatnější vlastnosti (tzv. latentní reprezentaci).

1. Architektura: Kodér a Dekodér

Autoencoder se skládá ze dvou hlavních částí, které jsou symetricky uspořádány kolem středu:

- **Encoder (Kodér):** Tato část sítě postupně zmenšuje počet neuronů v každé vrstvě. Převádí vstupní data (např. obrázek o rozlišení 100×100 pixelů) do nízkorozměrného vektoru.
- **Bottleneck (Latent Space / Code):** Nejužší část sítě. Zde jsou data uložena v nejvíce komprimované podobě. Tento prostor obsahuje „esenci“ dat.
- **Decoder (Dekodér):** Tato část je zrcadlovým obrazem kodéru. Snaží se z komprimovaného vektoru znovu sestavit původní data tak, aby se co nejvíce podobala vstupu.

2. Jak probíhá učení?

Učení probíhá minimalizací **rekonstrukční chyby** (Reconstruction Loss). Síť porovnává svůj výstup se svým původním vstupem.

- Pokud je výstup rozmazaný nebo chybný, algoritmus **Backpropagation** upraví váhy v kodéru i dekodéru.
- Cílem je, aby se síť nenaučila data jen tupě kopírovat, ale aby pochopila jejich vnitřní strukturu.

3. Typy Autoencoderů

Typ	Popis	Využití
Denoising Autoencoder	Na vstup dostane zašuměná data (např. zrnitou fotku) a má za úkol rekonstruovat čistý obraz.	Odstraňování šumu z obrazu a zvuku.
Sparse Autoencoder	Nutí síť aktivovat jen malý počet neuronů v úzkém hrdle.	Extrakce unikátních příznaků.
Variational (VAE)	Místo pevného vektoru se učí pravděpodobnostní rozdělení dat.	Generování nových dat (obrázků, hudby).
Contractive	Je navržen tak, aby byl odolný vůči malým změnám ve vstupních datech.	Velmi stabilní extrakce příznaků.

4. Praktické využití

- **Redukce dimenzionality:** Podobně jako metoda PCA, ale autoencodery dokážou zachytit i nelineární vztahy v datech.
- **Detekce anomálií:** Síť natrénovaná na „normálních“ datech (např. zdravý tlukot srdce) nedokáže přesně rekonstruovat „nenormální“ vstup (arytmii). Vysoká rekonstrukční chyba pak

signalizuje problém.

- **Kompresce dat:** Efektivní ukládání specifických typů dat (např. lékařských snímků).
- **Barvení černobílých fotografií:** Specificky upravené modely se dokážou naučit doplňovat chybějící barevnou informaci.

5. Omezení

Autoencodery jsou **ztrátové** (lossy) a velmi specifické pro daný dataset. Autoencoder vytrénovaný na fotkách psů nebude schopen kvalitně zrekonstruovat fotky aut, protože se naučil pouze „psí“ příznaky.

Zajímavost: Autoencodery byly jedním z prvních nástrojů pro tzv. *Unsupervised Pre-training*, který pomohl odstartovat éru **hlubokého učení** tím, že pomáhal inicializovat váhy ve velmi hlubokých sítích.

[Zpět na AI rozcestník](#)

From:

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

Permanent link:

<https://serviceit.cz/doku.php?id=autoencoder>

Last update: **2025/12/31 14:28**

