

Autoenkodéry (Autoencoders)

Autoenkodér je typ umělé **neuronové sítě**, která se používá pro učení efektivního kódování dat v rámci **učení bez učitele**. Cílem autoenkodéru je naučit se reprezentaci (kompresi) vstupních dat, obvykle za účelem snížení dimenzionality nebo odstranění šumu.

Základní princip spočívá v tom, že síť se snaží na svém výstupu zrekonstruovat přesně to, co dostala na vstupu.

1. Architektura autoenkodéru

Autoenkodér se skládá ze tří hlavních částí:

A. Enkodér (Encoder)

Část sítě, která postupně zmenšuje vrstvy a komprimuje vstupní data do menšího rozměru. Tato část „extrahuje“ nejdůležitější rysy dat.

B. Bottleneck (Hrdlo / Latentní prostor)

Nejužší vrstva sítě. Zde jsou data uložena v nejvíce komprimované podobě (tzv. **latentní reprezentace**). Tato vrstva nutí síť zahodit nepodstatné informace (šum) a ponechat jen to podstatné.

C. Dekodér (Decoder)

Zrcadlová část k enkodéru. Bere komprimovaná data z hrdla a snaží se z nich zrekonstruovat původní vstup v co nejvyšší kvalitě.

2. Jak probíhá učení?

Autoenkodér nepotřebuje lidské štítky (labels). Jako „učitel“ mu slouží samotný vstup.

- **Loss Function (Ztrátová funkce):** Měří rozdíl mezi vstupem (x) a zrekonstruovaným výstupem ($\hat{x}$). Nejčastěji se používá **MSE (Mean Squared Error)**.
- Cílem sítě je minimalizovat tuto chybu rekonstrukce.

$$\text{Loss} = \|x - \hat{x}\|^2$$

3. Typy autoenkodéru a jejich využití

Denoising Autoencoder (Odšumovací)

Do vstupu se záměrně přidá šum (např. náhodné pixely). Síť se trénuje tak, aby na výstupu byl čistý obraz. Tím se učí chápat skutečnou strukturu dat, nikoliv jen kopírovat pixely.

Variational Autoencoder (VAE)

Pokročilý typ, který místo pevných bodů v latentním prostoru modeluje pravděpodobnostní rozdělení.

- **Využití:** Generování nových dat (např. generování lidských tváří, které neexistují).

Sparse Autoencoder

Vynucuje, aby v hrdle bylo aktivních jen několik málo neuronů zároveň. To pomáhá extrahovat unikátní a silné příznaky.

4. Praktické aplikace

- **Redukce dimenzionality:** Modernější a nelineární alternativa k [PCA](#).
- **Detekce anomálií:** Síť trénovaná na normálních datech (např. z motoru) bude mít vysokou chybu rekonstrukce u dat, která vykazují poruchu.
- **Vybarvování obrázků:** Vstupem je černobílý obrázek, cílem je rekonstruovat barevný.
- **Kompres dat:** Efektivní ukládání informací v malém objemu dat.

5. Ukázka v Pythonu (Keras)

```
from keras.layers import Input, Dense
from keras.models import Model

# Definice rozměrů
input_size = 784 # např. 28x28 obrázek
code_size = 32 # rozměr komprese (bottleneck)

# Enkodér
input_img = Input(shape=(input_size,))
encoded = Dense(code_size, activation='relu')(input_img)

# Dekodér
decoded = Dense(input_size, activation='sigmoid')(encoded)

# Sestavení modelu
autoencoder = Model(input_img, decoded)
```

```
autoencoder.compile(optimizer='adam', loss='binary_crossentropy')
```

Související články:

- [Neuronové sítě](#)
- [Učení bez učitele](#)
- [Generativní modely \(GAN vs VAE\)](#)

Tagy: ai neural_networks autoencoders deep_learning unsupervised_learning

From:

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

Permanent link:

<https://www.serviceit.cz/doku.php?id=it:ai:autoencoders>

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

