

# Mély generatív modellek (deep generative models) használata az agy látórendszerében végrehajtott hierarchikus feldolgozás megértésére

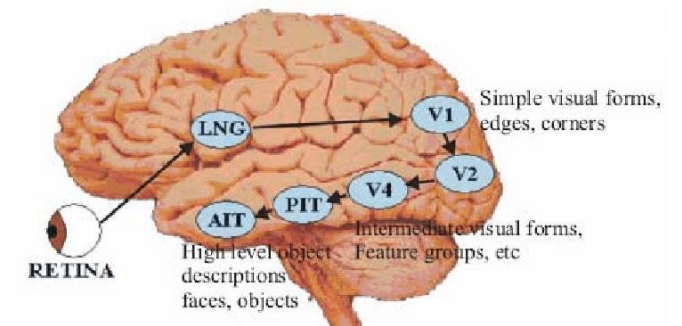
Kun Mátyás

Témavezetők: Csikor Ferenc, Orbán Gergő

Komputációs Rendszerszintű Idegtudomány Kutatócsoport, Komputációs Tudományok Osztálya, Wigner FK

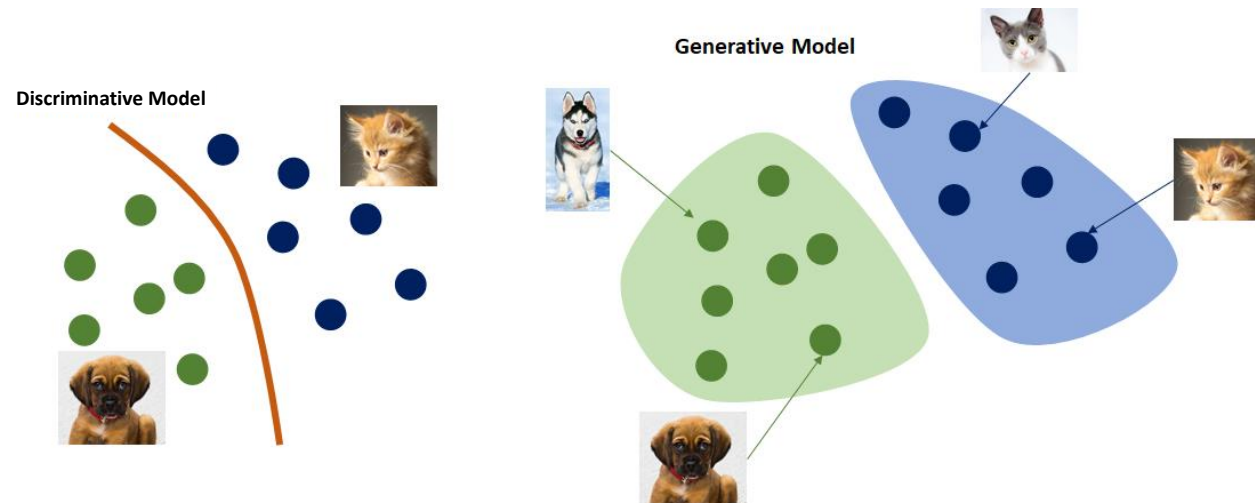
# A feladatom ismertetése

- A látórendszer egyes részeinek működését és azok szerepét szeretném jobban megismerni egy gépi tanuló modell segítségével
- A feladatom tehát egy olyan modell tanítása és implementálása, ami a lehető legjobban hasonlít az agyban megtalálható látórendszerhez
- Éppen ezért természetes képeket fogunk adni a modellnek, ezáltal jobban megtudjuk vizsgálni, hogy milyen reprezentációkat alakít ki egy ilyen rendszer
- Fontos megjegyezni még, hogy csupán a két főbb részét vizsgálom ennek a hierarchikus rendszernek, V1-et és V2-t



# A modellem megértéséhez szolgáló háttér

- Az *unsupervised learning* gépi tanulási ágon belül, egy olyan modellt használunk, ami *probabilisztikus* eszközöket használ és látens változók segítségével határozza meg az eredeti adatstruktúra eloszlását
- Az ilyen modellek tanítása során olyan  $\theta$  paramétereket akarunk megtalálni, melyek felhasználásával  $\log(p_{\theta}(x)) \approx p(x)$
- Generatív modellek ezen belül olyanok, amik a megjósolt eloszlásokból mintavételezés segítségével az eredeti adatpontokhoz hasonlókat tudnak létrehozni

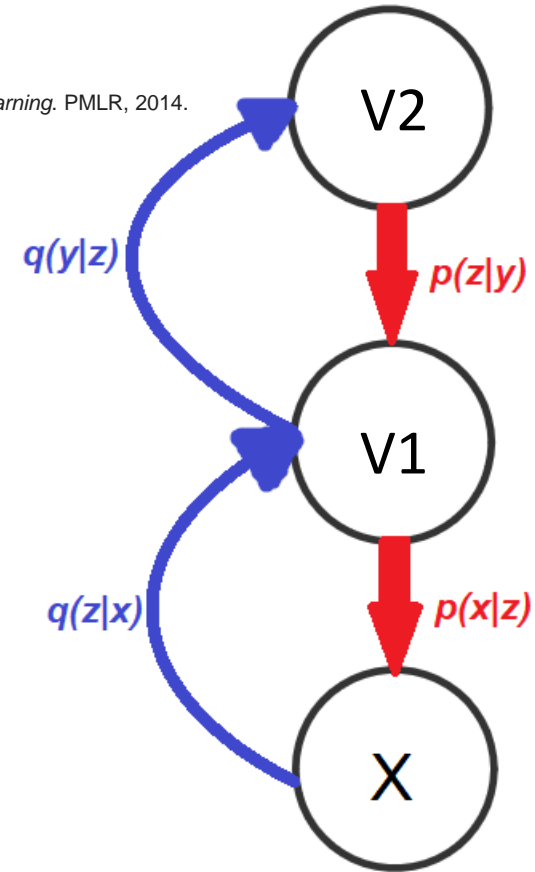


# Modellem, a variational autoencoder

Kingma, Diederik P., and Max Welling. "An introduction to variational autoencoders." *arXiv preprint arXiv:1906.02691* (2019).

Rezende, Danilo Jimenez, Shakir Mohamed, and Daan Wierstra. "Stochastic backpropagation and approximate inference in deep generative models." *International conference on machine learning*. PMLR, 2014.

- Annak érdekében, hogy az eddig ismertett modell tanítása során keletkező problémákat kezelhetővé tudjuk alakítani, szükségünk lesz egy közelítő inferencia modellre  $q(z|x)$ -re (ugyanis  $p(z|x)$  nem számolható ki hatékonyan)
- Ebben az esetben a variációs posteriort szeretnénk úgy optimalizálni, hogy  $q(z|x) \approx p(z|x)$  legyen
- Az általam használt variational autoencoder két inferencia és két generatív leképezést tartalmaz, hierarchikusan vannak egymásra építve



# Tanulási folyamata a VAE-nak

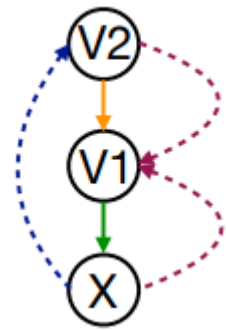
- Mivel a célunk, hogy a modell által prediktált  $\log(p_\theta(x))$  (*log marginal likelihood*) közel azonos legyen az eredeti  $p(x)$ -hez, jobban meg kell értenünk, hogy milyen tagokból épül fel:
- Egy Kullback-Leibler divergencia (*KL*) tagból és egy Evidence lowerbound tagból (*ELBO*)
- Tanítás során a  $\log(p_\theta(x))$  értékét szeretnénk maximalizálni, melyet közelítőleg az ELBO maximalizálásával tudjuk megtenni

$$\log(p(x)) = \text{KL}[q(y, z|x) || p(y, z|x)] + \text{ELBO}$$

$$\text{ELBO} = E_{q(z|x)}[\log(p(x|z))] - E_{q(z|x)}[\text{KL}[q(y|z) || p(y)]] - E_{q(y|z)}[\text{KL}[q(z|x) || p(z|y)]]$$

# Összefoglalás

- Az ELBO-t akkor is többféleképp tudjuk particionálni, ha a generatív modellt rögzítjük
- Az általam használt particionálás a Chain inference ELBO, ennek egy alternatív particionálása a Top-Down inference ELBO, amivel a kutatócsoport foglalkozik:
- Azért használunk különböző fajta particionálásokat, mert mindegyik más fajta megfigyeléseket vetíthet előre
- A modell implementálását most fogom befejezni
- Implementálás után elkezdem tanítani azt, hogy a későbbiekben már a fontosabb összefüggésekre tudjak fényt deríteni



Top Down:  $q(V1, V2|X) = q(V1|V2, X) \cdot q(V2|X)$

Chain:  $q(V1, V2|X) = q(V2|V1) \cdot q(V1|X)$

Köszönöm a figyelmet!