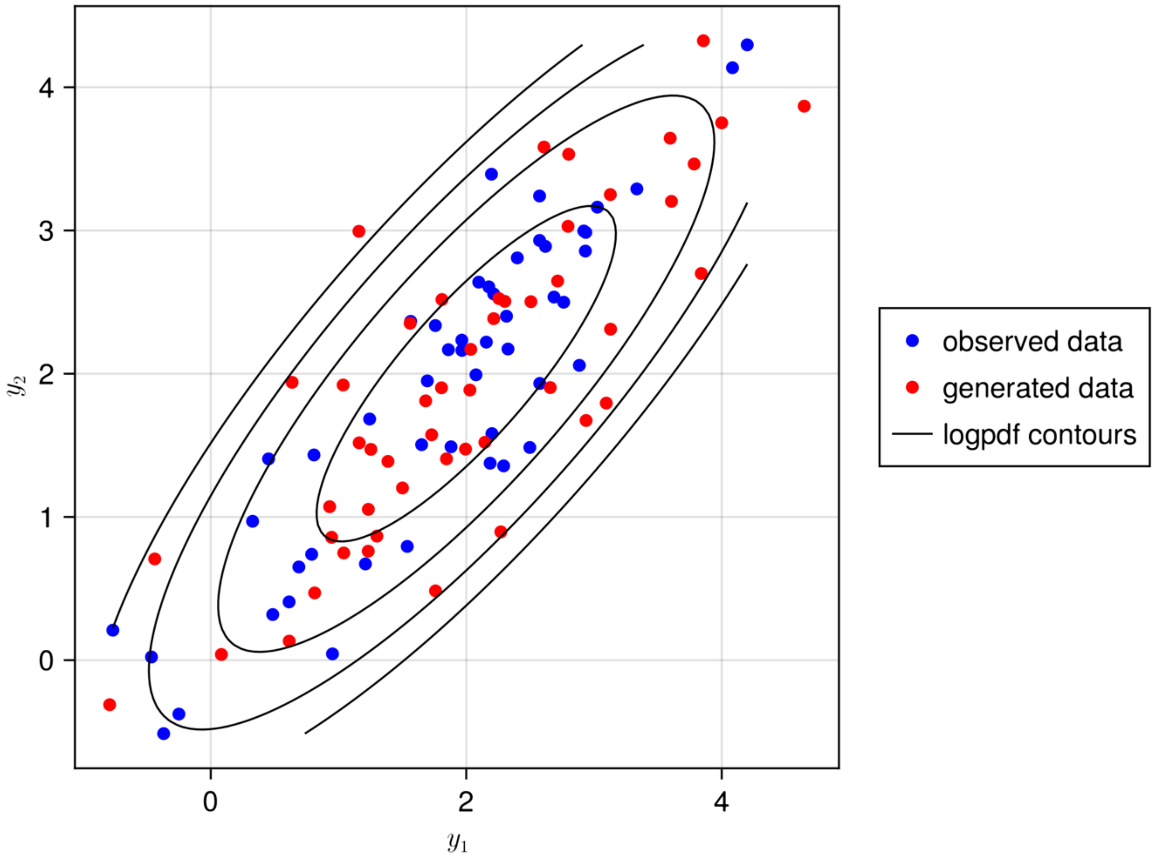
DeepPumas
Embedding models

Niklas Korsbo

Generative AI

Goal:
Generated data instinguishable in distribution from real data



DeepPumas

www.thispersondoesnotexist.com

How is that done?

Data is a mix of

- Observed quantities (pixel intensities)
- Unobserved quantities (face, smiling, …)

We humans learn to extract the unobserved
GenAI needs to do that too.

DeepPumas

How is that done?

Data is a mix of

- Observed quantities ($\boldsymbol{y}$)
- Unobserved quantities ($\boldsymbol{z}$ – "latent variables")

DeepPumas

Generative models

- Definitions
  - $\boldsymbol{z}$: latent variables of dimension d
  - $\boldsymbol{y}$: observed response/data
  - $\boldsymbol{y_g}$: generated/simulated/synthetic response/data
- Model

$$\boldsymbol{y_g} = \boldsymbol{f}(\boldsymbol{z}) + \boldsymbol{\epsilon}$$
$$\boldsymbol{z} \sim \text{Normal}(0, \boldsymbol{I}_{d \times d})$$
$$\epsilon_i \sim \text{Normal}(0, \sigma^2)$$

- Objective: choose $\boldsymbol{f}$ such that the distribution of $\boldsymbol{y_g}$ is close to the distribution of the observed data $\boldsymbol{y}$

DeepPumas

Conditional generative models

- Definitions
  - $\boldsymbol{z}$: latent variables of dimension d
  - $\boldsymbol{x}$: observed covariates
  - $\boldsymbol{y}$: observed response
  - $\boldsymbol{y_g}$: generated/simulated/synthetic response
- Model

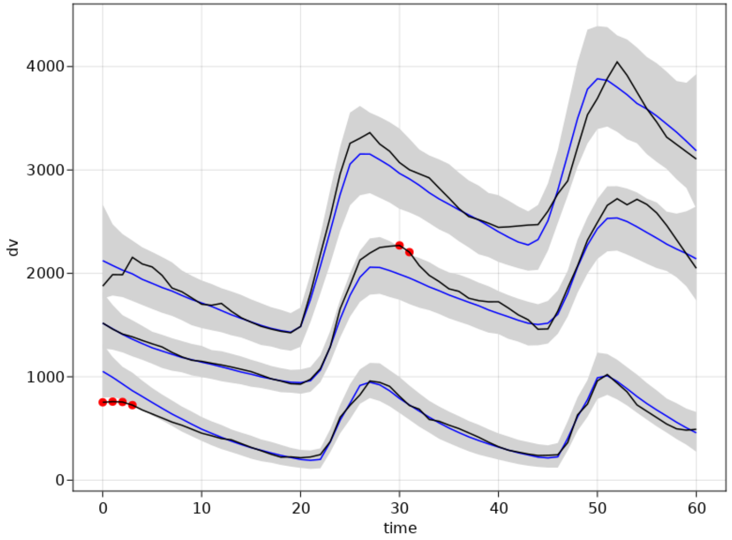
$$\boldsymbol{y_g} = \boldsymbol{f}(\boldsymbol{z}, \boldsymbol{x}) + \boldsymbol{\epsilon}$$
$$\boldsymbol{z} \sim \text{Normal}(0, \boldsymbol{I}_{d \times d})$$
$$\epsilon_i \sim \text{Normal}(0, \sigma^2)$$

- Objective: choose $\boldsymbol{f}$ such that the conditional distribution of $\boldsymbol{y_g} \mid \boldsymbol{x}$ is close to the conditional distribution of the observed data $\boldsymbol{y} \mid \boldsymbol{x}$

NLME is Generative AI!

- Definitions
  - $\boldsymbol{\eta}$: latent variables of dimension d and covariance matrix $\boldsymbol{\Omega}$
  - $\boldsymbol{x}$: observed covariates
  - $\mathbf{dv}$: observed response
  - $\mathbf{dv_g}$: generated/simulated/synthetic response
- Model

$$\mathbf{dv_g} = \boldsymbol{f_\theta}(\boldsymbol{\eta}, \boldsymbol{x}) + \boldsymbol{\epsilon}$$
$$\boldsymbol{\eta} \sim \text{Normal}(0, \boldsymbol{\Omega})$$
$$\epsilon_i \sim \text{Normal}(0, \sigma^2)$$



- Objective: choose $\boldsymbol{f_\theta}$ such that the conditional distribution of $\mathbf{dv_g} \mid \boldsymbol{x}$ is close to the conditional distribution of the observed data $\mathbf{dv} \mid \boldsymbol{x}$

Latent variables = Random effects

An Introduction to Variational Autoencoders

Diederik P. Kingma, Max Welling

$$p_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{z}) = p_{\boldsymbol{\theta}}(\mathbf{z}) p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{z}) \tag{1.14}$$

where $p_{\boldsymbol{\theta}}(\mathbf{z})$ and/or $p_{\boldsymbol{\theta}}(\mathbf{x}|\mathbf{z})$ are specified. The distribution $p(\mathbf{z})$ is often called the *prior distribution* over $\mathbf{z}$, since it is not conditioned on any observations.

The marginal distribution over the observed variables $p_{\boldsymbol{\theta}}(\mathbf{x})$, is given by:

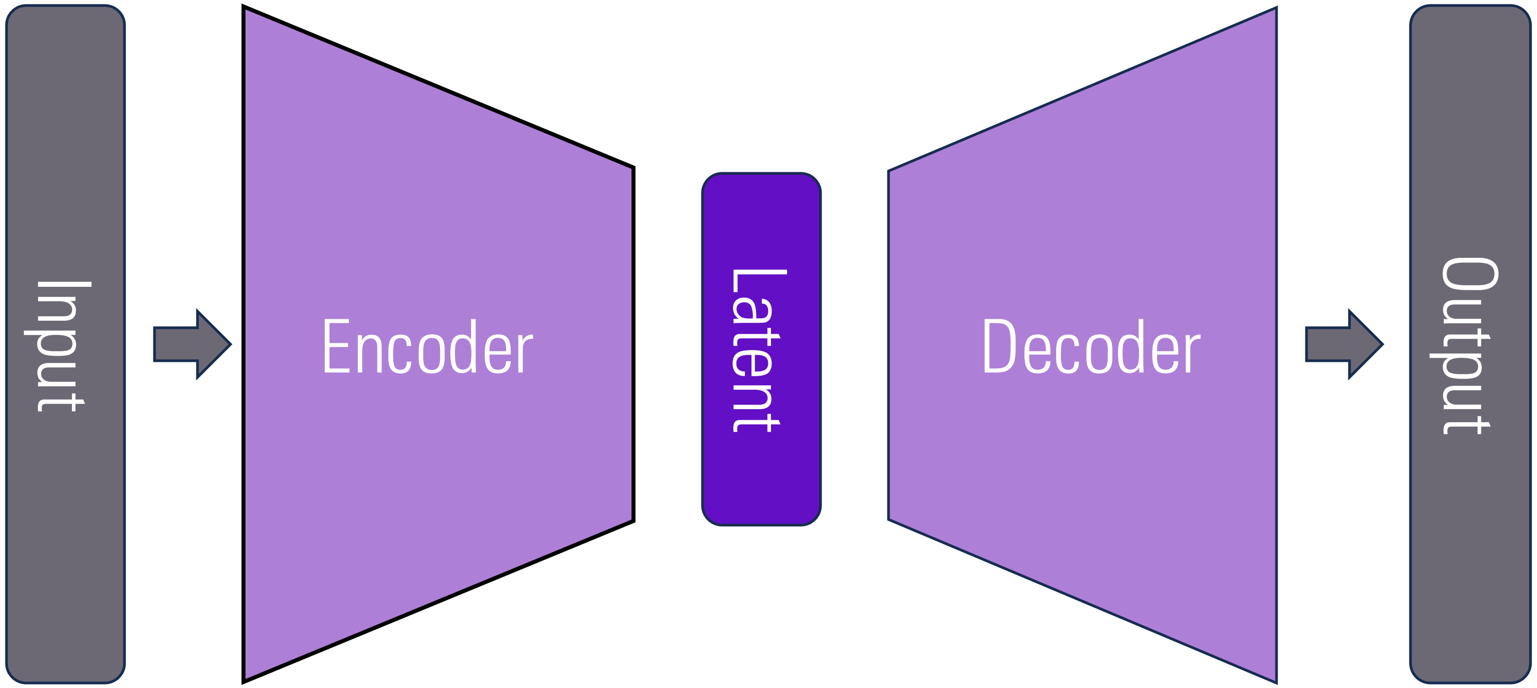
$$p_{\boldsymbol{\theta}}(\mathbf{x}) = \int p_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{z}) \, d\mathbf{z} \tag{1.13}$$

This is also called the (single datapoint) *marginal likelihood* or the *model evidence*, when taken as a function of $\boldsymbol{\theta}$.

Maximize the marginal likelihood.

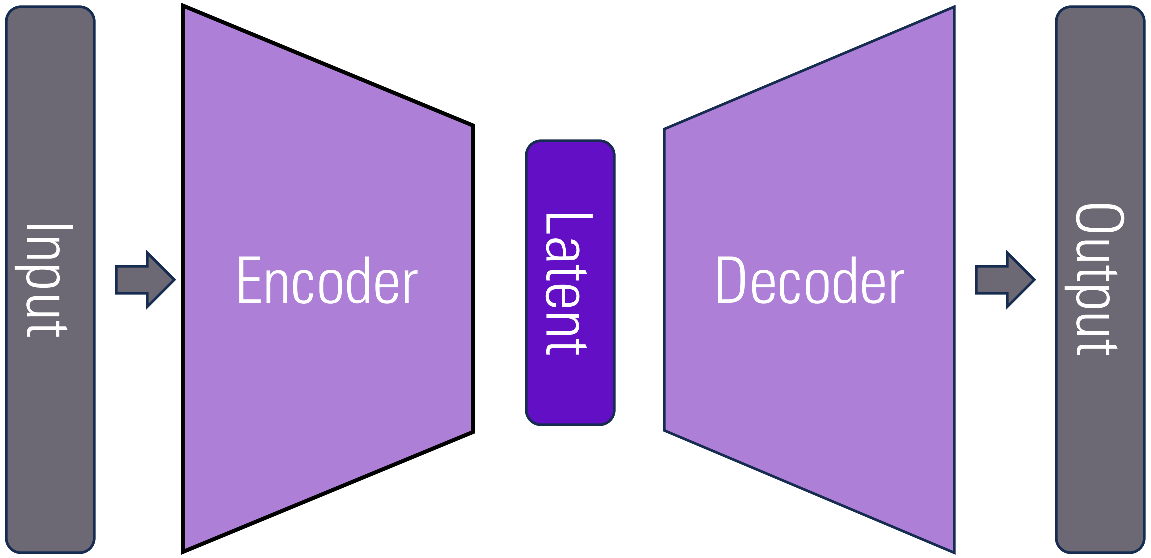
Marginalize over $\boldsymbol{z}$

Generative AI – Typical anatomy



DeepPumas

NLME as GenAI



Input: Time series of outcomes/concentrations
Output: Time series of outcomes/concentrations

Decoder: The structural NLME (parameter transforms, dynamics, etc.)
Encoder: ???

DeepPumas

What do “latent variables” represent?

NLME:

- Interpretation from structural constraints: $CL = tvCL \cdot \exp(\eta_{cl})$.
    - These interpretations can fool us!
- Structural constraints not necessary: $NN\left(\frac{Central}{Vc},\ \eta\right)$

DeepPumas

What do “latent variables” represent?

Image GenAI:

For an image:

- Not pixel-by-pixel intensity

Rather:

- What objects are in the image
- What are the characteristics of the objects
- What are they doing
- What’s the style

What do “latent variables” represent?

Text GenAI:

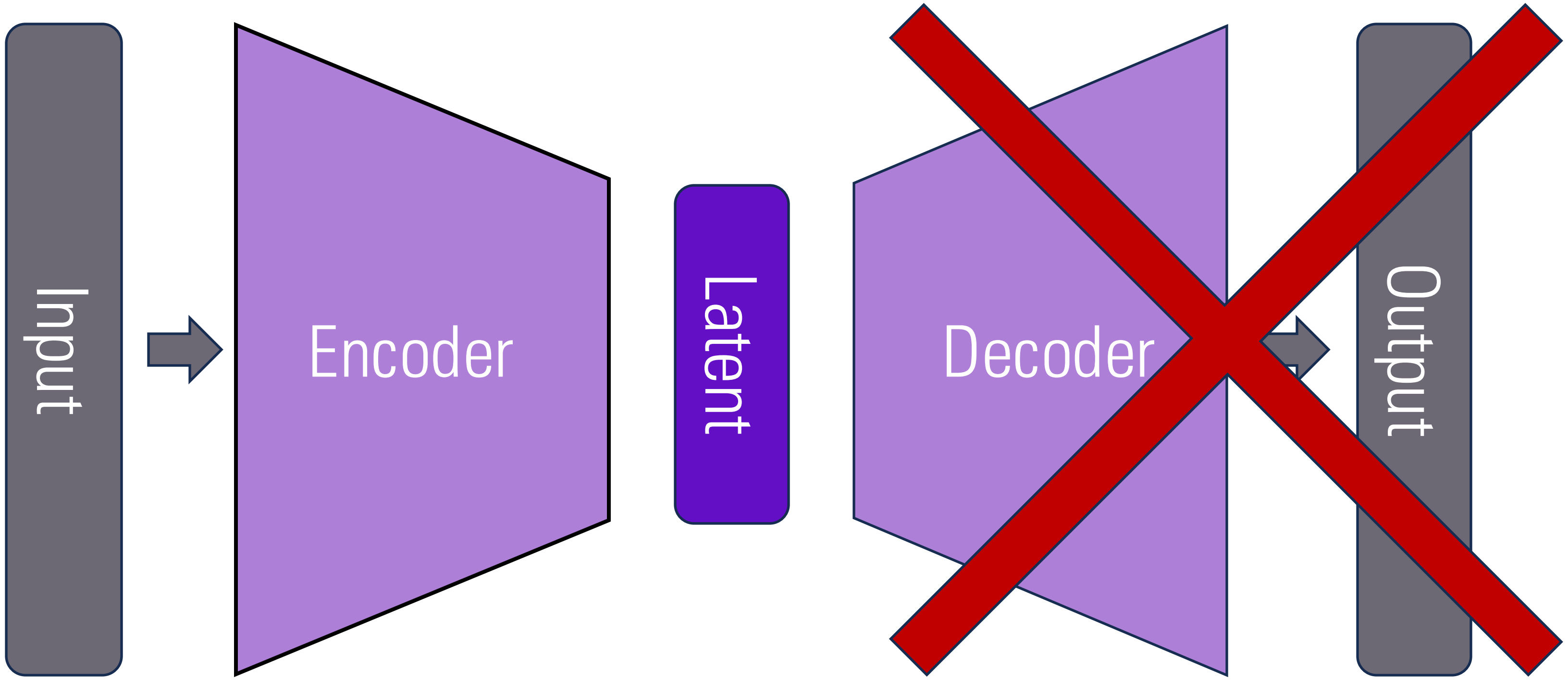
- Not words

rather:

- Sentiment
- Conveyed information
- Writing style
- Language
- …

DeepPumas

Embedding models

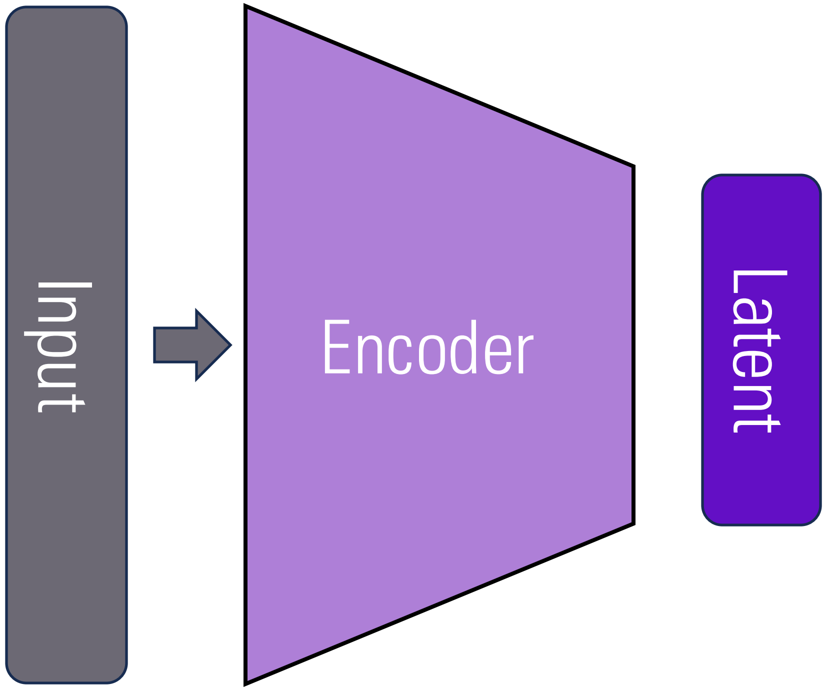


Extract meaningful, information-dense features from the data

https://huggingface.co/spaces/mteb/leaderboard

DeepPumas

A unified "API" for complex data



Different data modalities (images, text, etc) all map to embeddings - simple vectors of numbers.

DeepPumas