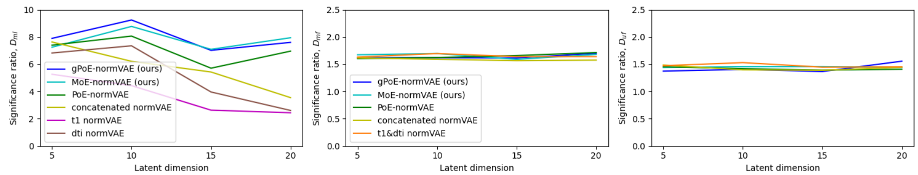
Supplementary Materials

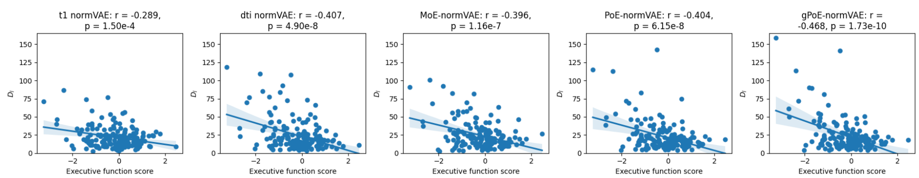
Table 1: α weightings for the gPoE-normVAE model with L_{dim}=10.

modality	latent 0	latent 1	latent 2	latent 3	latent 4	latent 5	latent 6	latent 7	latent 8	latent 9
T1	0.422716	0.319826	0.704999	0.4191	0.345487	0.520485	0.464519	0.698466	0.156988	0.632687
DTI	0.577284	0.680174	0.295001	0.5809	0.654513	0.479515	0.535481	0.301534	0.843012	0.367313

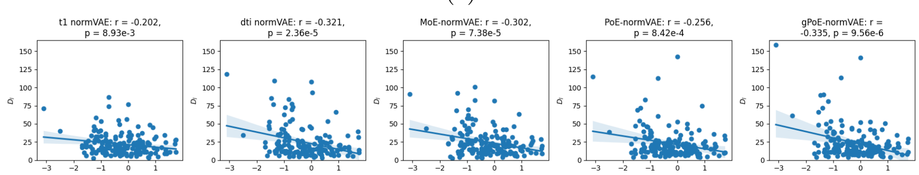


(a)

Fig. 1: (a) Significance ratio calculated from D_{ml}, and D_{mf} and D_{uf} for the UK Biobank.



(a)



(b)

Fig. 2: Pearson correlation between D_{ml} and (a) executive function and (b) memory scores for the T1 normVAE, DTI normVAE, MoE normVAE, PoE normVAE, and gPoE normVAE models applied to the ADNI dataset.

Product and generalised product of Gaussian's. In this section, we provide the derivation for the parameters of the Product of Experts (PoE) and generalised Product of Experts (gPoE) Gaussian product distributions.

Proof. The probability density of a Gaussian distribution is given by:

$$p = K \exp\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\} = K \exp\left\{\mathbf{x}^T \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}^{-1}\mathbf{x}\right\} \quad (1)$$

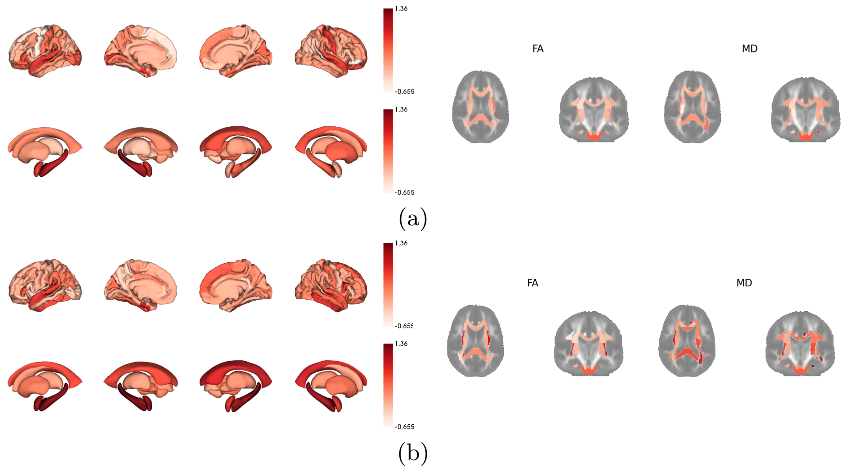


Fig. 3: (a) Average D_{uf} using the gPoE-normVAE (L_{dim}=10) model for the LMCI and (b) AD cohort. The left-hand plots show T1 features and the right-hand plots show DTI features.

where K is a normalisation constant. The product of N Gaussian distributions is a Gaussian of the form:

$$\prod_{n=1}^{N} p_n \propto \exp\{\mathbf{x}^T \sum_{n=1}^{N} \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\mu}_n - \frac{1}{2}\mathbf{x}^T (\sum_{n=1}^{N} \boldsymbol{\Sigma}_n^{-1})\mathbf{x}\} \tag{2}$$

where $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} = \sum_{n=1}^{N} \boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n$ and $\boldsymbol{\Sigma}^{-1} = \sum_{n=1}^{N} \boldsymbol{\Sigma}_n^{-1}$. Thus the product Gaussian has a mean $\boldsymbol{\mu} = \frac{\sum_{n=1}^{N} \boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n}{\sum_{n=1}^{N} \boldsymbol{\Sigma}_n^{-1}}$ and covariance $\boldsymbol{\Sigma} = (\sum_{n=1}^{N} \boldsymbol{\Sigma}_n^{-1})^{-1}$. If we have isotropic Gaussian distributions $p_n = \mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\sigma}_n^2 \mathbf{I})$, the parameters of the product Gaussian become $\boldsymbol{\mu} = \frac{\sum_{n=1}^{N} \boldsymbol{\mu}_n / \boldsymbol{\sigma}_n^2}{\sum_{n=1}^{N} 1/\boldsymbol{\sigma}_n^2}$ and $\boldsymbol{\sigma}^2 = \frac{1}{\sum_{n=1}^{N} 1/\boldsymbol{\sigma}_n^2}$.

Similarly, for a weighted product of N Gaussian distributions, the product Gaussian has the form:

$$\prod_{n=1}^{N} p_n^{\alpha_n} \propto \exp\{\mathbf{x}^T \sum_{n=1}^{N} \alpha_n \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\mu}_n - \frac{1}{2}\mathbf{x}^T (\sum_{n=1}^{N} \alpha_n \boldsymbol{\Sigma}_n^{-1})\mathbf{x}\}. \tag{3}$$

For isotropic Gaussian distributions, the weighted product Gaussian has mean $\boldsymbol{\mu} = \frac{\sum_{n=1}^{N} \boldsymbol{\mu}_n \boldsymbol{\alpha}_n / \boldsymbol{\sigma}_n^2}{\sum_{n=1}^{N} \boldsymbol{\alpha}_n / \boldsymbol{\sigma}_n^2}$ and variance $\boldsymbol{\sigma}^2 = \sum_{n=1}^{N} \frac{1}{\boldsymbol{\alpha}_n / \boldsymbol{\sigma}_n^2}$.