Probabilistic Embedding Model for Latent Feature Learning

A probabilistic latent variable model was built on PCA-reduced data to learn a compact, non-linear representation of high-dimensional gene expression profiles. This model is a type of neural network containing an encoder and decoder with an entropy-limited latent layer of variables (where , and , the number of PCA features). This architecture generates an embedding that preserves information from the high-dimensional input within a lower-dimensional space.

The encoder network, defined as , maps from the input space to the latent embedding . The decoder, , reconstructs the input from the latent space. The primary objective is to minimize the expected squared Euclidean (L2) distance between the input and its reconstruction:

Here, and are the learnable parameters of the encoder and decoder, respectively, and represents the reconstructed input. The L2 loss, denoted , captures the total reconstruction error across all input dimensions. Explicitly, this corresponds to:

Unlike conventional encoders, the probabilistic embedding model (PEM) represents each sample as a probability distribution in latent space. This formulation captures biological variability and uncertainty in gene expression.

The input matrix (based on 500 PCs) is passed to the encoder , which outputs a mean vector and a variance vector :

The decoder reconstructs the input from a sampled latent vector . The full objective function includes a regularization term:

The reconstruction term ensures fidelity to the input, while the KL divergence regularizes the latent distribution toward a standard normal prior. After training, the learned latent variables were used for downstream analysis, including gene importance scoring using Integrated Gradients and pathway enrichment analysis.