Data pre-processing

GatorSC takes the scRNA-seq gene expression matrix as input. The scRNA-seq data are represented as a two-dimensional matrix, where rows correspond to cells, and columns to genes. Each cell is associated with an available label. The matrix is processed using the Scanpy toolkit. Preprocessing involves normalization and logarithmic transformation of the gene expression matrix.

Implementation and parameter settings

GatorSC was implemented in Python (v3.9.12) with PyTorch (v1.11.2). For each scRNA-seq dataset, we randomly split cells into training/validation/test sets with a ratio of 80%/10%/10%. For hierarchical graph modeling, we used two attention heads and initialized the learnable thresholds as $\varphi^{(1)} = 0.05$, $\varphi^{(2)} = 0.02$, and $\varphi^{(3)} = 0.03$; the hop number was set to 2. For adaptive fusion of multi-level graph representations, the hidden dimension of both the GCN and MLP was set to 225, and the number of expert networks was set to 3. For unified self-supervision, the Bernoulli deletion and addition probabilities were both set to 0.4, and the maximum path length was set to 2. The temperature was set to $\tau = 0.25$. The loss scaling coefficients were set to $\alpha_1 = 0.55$, $\alpha_2 = 0.53$, and $\alpha_3 = 0.62$; $\beta_1 = 0.90$ and $\beta_2 = 0.85$; $\gamma_1 = 0.76$ and $\gamma_2 = 0.89$; and $\lambda = 0.68$. We trained the model using the Adam optimizer with a learning rate of 10^{-3} and a batch size of 256. For clustering, we applied K-means (scikit-learn default settings) to the learned embeddings, setting the number of clusters to the ground-truth number of cell types for each dataset. For other clustering baselines, we used the default hyperparameters provided by the corresponding packages. All hyperparameters were selected via grid search based on validation performance. All experiments were conducted on a workstation equipped with an NVIDIA RTX 4090 GPU (24 GB memory).