Bridge Neural Network Training Curriculum

Phase 1: Supervised Learning

Train model to recognize knowledge boundaries using labeled examples

$Loss = \lambda_1 L^{LM}Loss + \lambda_2 BridgeDetectionLoss$

Training:
- Base Model
- Bridge Detector

Metrics:
- Language Accuracy
- Bridge Activation

Phase 2: Bridge Detection

Focus on accurate activation of bridge neurons at knowledge boundaries

$Loss = 0.5\lambda_1 L^{LM}Loss + 3\lambda_2 BridgeDetectionLoss$

Training:
- Bridge Detector
- Limited Base Model

Metrics:
- K.B. Precision
- K.B. Recall

Phase 3: Bridge Retrieval

Train query encoder to generate effective neural query representations

$Loss = \lambda_1 L^{LM}Loss + \lambda_3 QueryQualityLoss$

Training:
- Query Encoder
- External Service

Metrics:
- Retrieval Precision
- Query-Doc Alignment

Phase 4: Integration (RL-based)

Optimize model's ability to incorporate retrieved information

$Loss = ActorLoss + \lambda_4 KLDivergence$

Training:
- Response Integrator
- End-to-end System

Metrics:
- Factual Accuracy
- Integration Quality