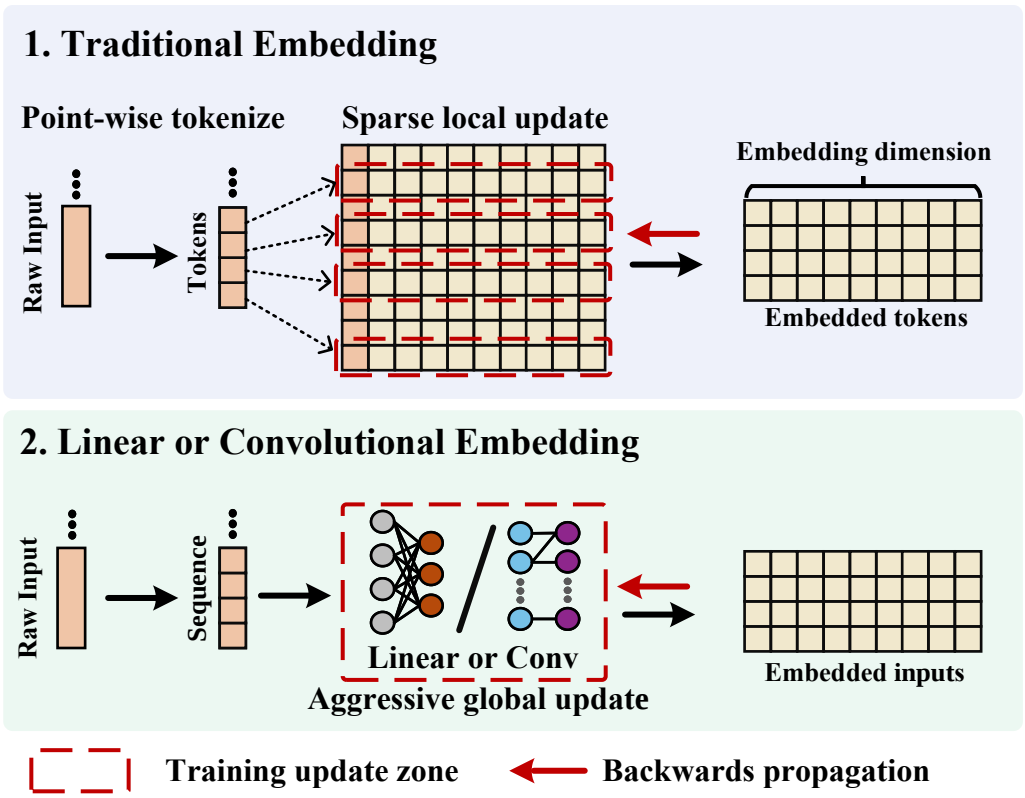
1. Traditional Embedding
Point-wise tokenize
Sparse local update
Embedding dimension
Raw Input
Tokens
Embedded tokens
2. Linear or Convolutional Embedding
Raw Input
Sequence
Linear or Conv
Aggressive global update
Embedded inputs
Training update zone
Backwards propagation