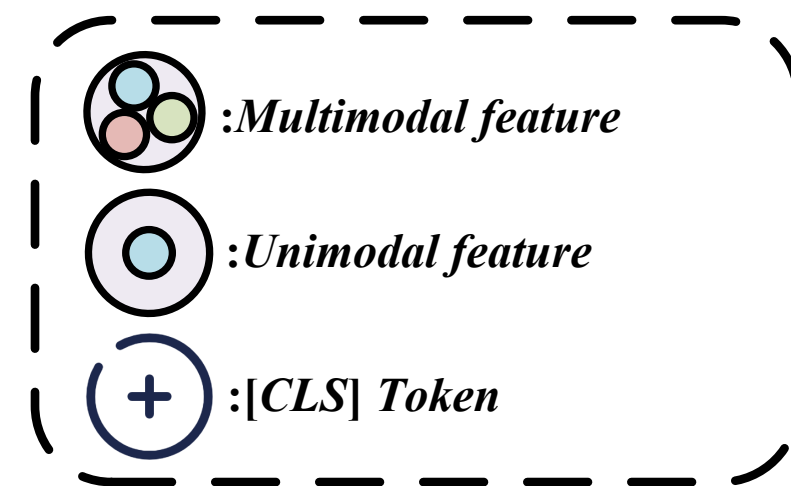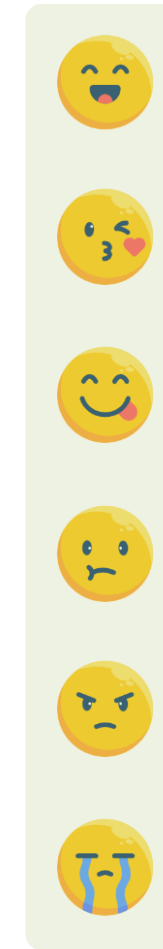**Multimodal sentiment analysis based on two-step hierarchical attention and self-supervised label generation**

→

Two-step Hierarchical Attention and Orthogonal Information Enhancement Model

→

⊙ : *Multimodal feature*

⊙ : *Unimodal feature*

⊕ : *[CLS] Token*

Update

constraint   *constraint*   *interference*

Feature extraction module

Two-step hierarchical attention

Orthogonal information enhancement module

Unimodal self-supervised label generation module

Sentiment prediction module

$$\| x^T \cdot y \|_2^F$$

Calculate   *Update*

z

y   x

Y

Z   x