| y_{ans} | Answer Token | $\overline{r_i}$ | Passage Relevance Label | $I(\cdot)$ | Indicator Function | $H(\cdot)$ | Entropy Function | $M_k(r_i\|q)$ | Relevance Score Predicted by k-th External Model |
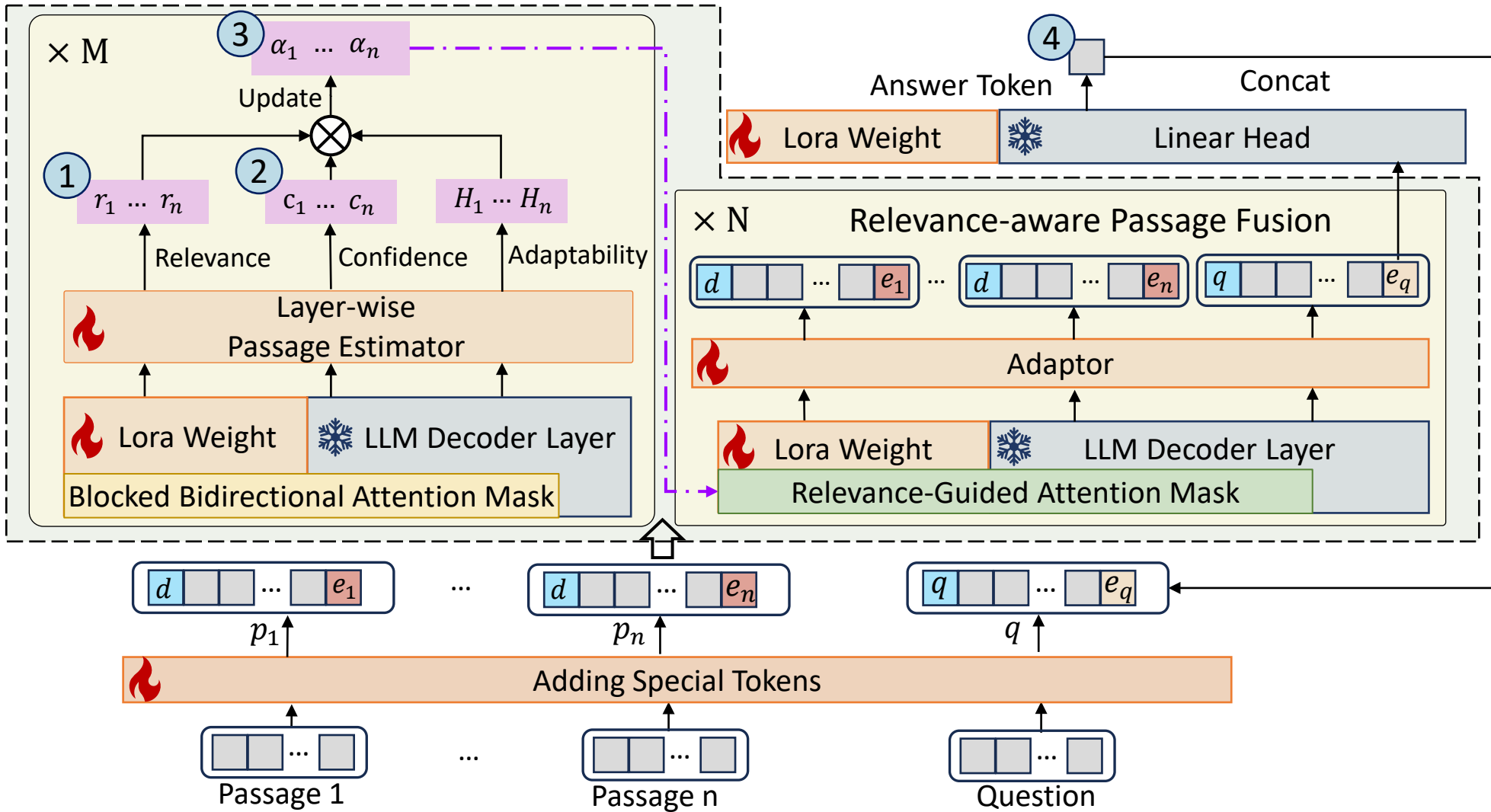
① Relevance Loss $\mathcal{L}_{\text{relevance}} = -\frac{1}{n}\sum_{i=1}^{n}\overline{r_i}\log r_i + (1-\overline{r_i})\log(1-r_i)$

② Confidence Loss $\mathcal{L}_{\text{confidence}} = -\sum_{i=1}^{n}c_i - \left[\frac{1}{K}\sum_{k=1}^{K}I(M_k(r_i\|q) == \overline{r_i})\right]$

③ Diversity Loss $\mathcal{L}_{diversity} = -H(\alpha_1, ..., \alpha_n)$

④ Language Model Loss $\mathcal{L}_{LM} = -log\, p(y_{\text{ans}}|q, p_1, ... p_n)$