

### Cross entropy loss on user's new data

$$\min_{\theta_{f_{QA}}, \theta_K, \theta_P, \theta_{f_c}} \mathcal{L}(f(X + P_T), y)$$

Diagram illustrating the components of the cross entropy loss on user's new data:

- $\theta_{f_{QA}}$ : Query adapter
- $\theta_K$ : Selected Keys
- $\theta_P$ : Selected Prompts
- $\theta_{f_c}$ : Classifier
- $X + P_T$ : Prompt Offset-ed Input
- $y$ : Ground Truth Label

### Query-Key clustering loss

$$\max_{\theta_{f_{QA}}, \theta_K} \lambda \sum_{i \in (s_i)_{i=1}^T} \gamma(q, K_i)$$

Diagram illustrating the components of the Query-Key clustering loss:

- $\lambda$ : Weight
- $\sum_{i \in (s_i)_{i=1}^T}$ : Sum over selected keys
- $\gamma(q, K_i)$ : Cosine Similarity between Query and Selected keys
- $q$ : Query
- $K_i$ : Selected keys