

(a) Influence of two key hyper-parameters for model performance.

(b) Attention heatmap on ABCD.