

1. Make a brief introduction about a variant of Transformer.

2. Briefly explain why adding convolutional layers to Transformer can boost performance.
