ML HW4 Report

B09901066 謝承修

1. **Make a brief introduction about a variant of Transformer. (2 pts)**

   Conformer is a variant of Transformer. It combines the best of the Transfomer models and convolution neural networks (CNN), good at capturing content-based global interactions and able to exploit local features effectively.

2. **Briefly explain why adding convolutional layers to Transformer can boost performance. (2 pts)**
   Since Transformer only focus on the global context, lack of exploiting local features effectively, while convolutional layer, is able to utilize receptive fields to capture features, which can ammend the weakness of self-attention.