**Machine Learning 2023 Spring Homework 5 Report**

生醫電資所 R11945071 莊聰賢

1. **Visualize the similarity between different pairs of positional embedding and briefly explain the result. Additionally, attach the code that you used for visualization.**

positional embedding就是把位置嵌入向量Visualize，例如繪製成heatmap、散點圖等，以便清楚地觀察不同位置embed之間的similarity、模式或其他特點。

下圖呈現出不同pairs of positional embedding之間的similarity，而neighboring words之間的關係會比較緊密，因此decoder所得到的similarity也會比較高。

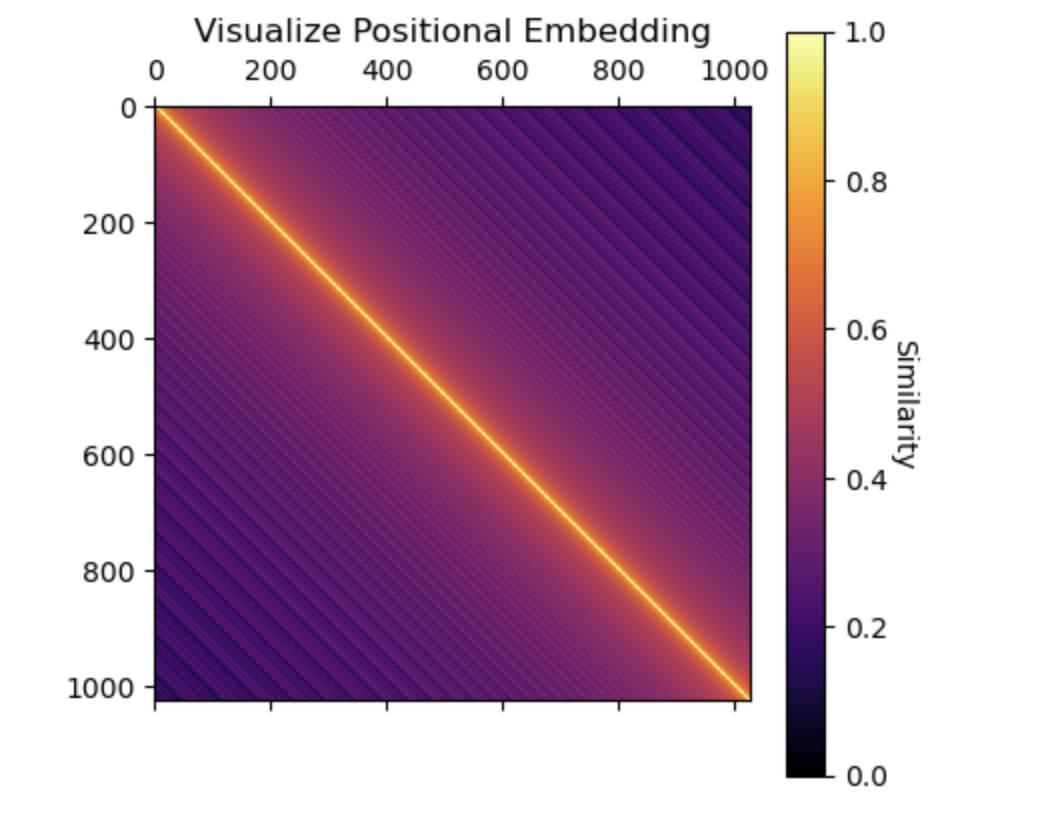


Fig1. Similarity between different pairs of positional embedding.

下圖為該次visualize使用的程式碼，主要參考pytorch寫出，figuresize定在(10,10)，matshow參數的部分選用cmap=”infero”，並加上colorbar協助理解顏色與similarity之間的關聯性。

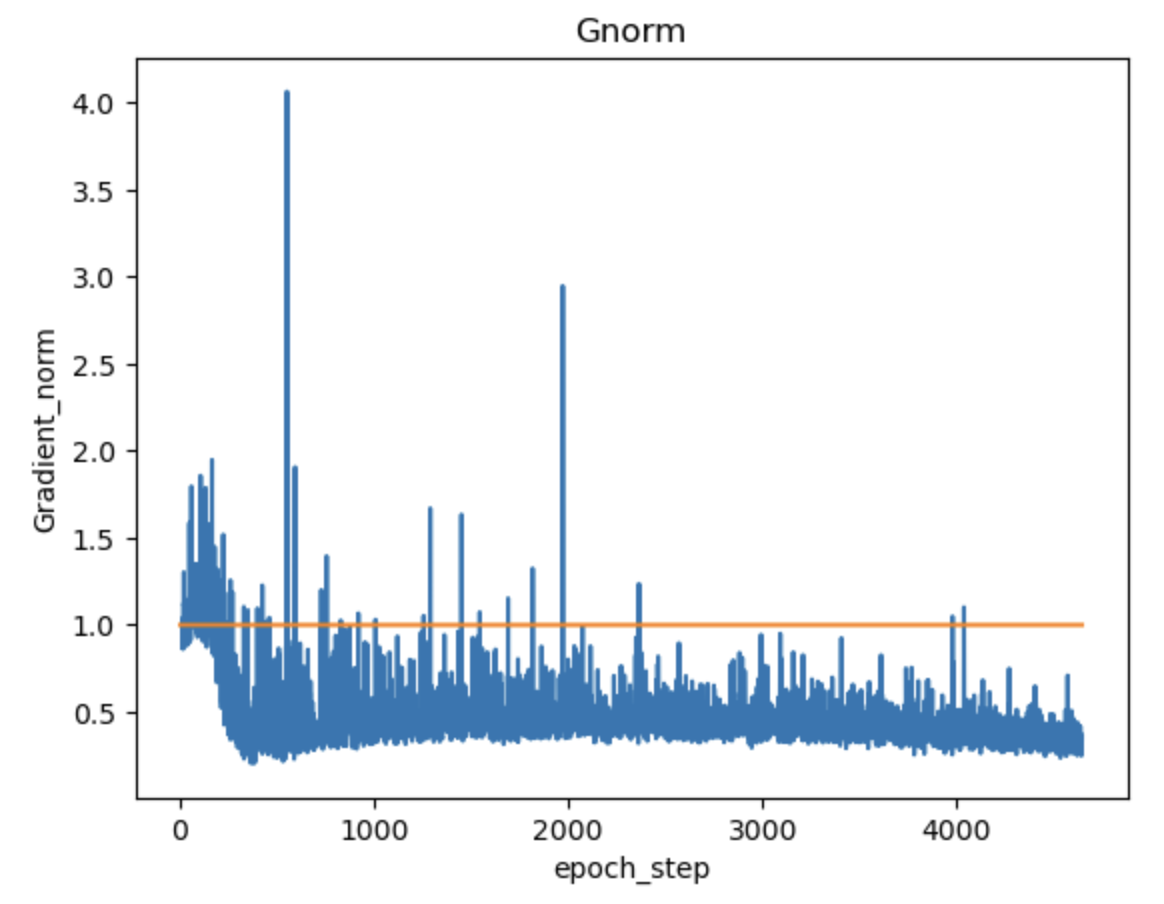


Fig2. code that I used for visualization.

1. Clip gradient norm and visualize the changes of gradient norm in different steps. Circle two places with gradient explosion.

下圖為使用本次通過strong 之model(without back translation)所得到的gradient norm結果，通過strong的epochs為30，而此圖為跑5epochs所得到的結果，可以由此圖得知多半在最一開始時發生所謂的gradient explosions，後期較少發生。

Fig2. Gradient norm (5epochs)



此次模型參數:

