Similarity Matrix

```python
import matplotlib.pyplot as plt
output = []
print(pos_emb)
for i in range(np.shape(pos_emb)[0]):
#      output.append([])
    tmp = torch.tensor(pos_emb[i])
    output.append(np.array(nn.functional.cosine_similarity(tmp, pos_emb)))
    if i==1:

        print(output)
# print(type(output))
# output = float(output)
plt.imshow(output, cmap = 'hot', interpolation = 'nearest')
plt.colorbar()
plt.title("Similarity Matrix")
plt.show()
```
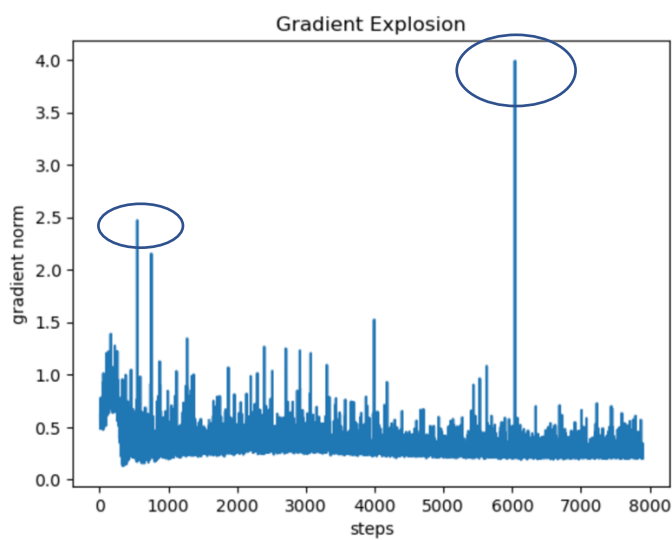
As we can see, the yellow line have largest similarity. It happen where the distance between two positional embedding is close. When the difference of positional embedding increases, the similarity decrease.

Circle two explosion in 10 epoch (use sample code with get_rate).