Contrastive Learning

: The goal of contrastive loss is to encourage the model to learn to differentiate between similar and dissimilar pairs of examples. In contrastive loss, the model is trained to minimize the distance between pairs of similar examples and maximize the distance between pairs of dissimilar examples. The loss function is defined as the average of the distances between pairs of similar examples and the maximum margin between pairs of dissimilar examples.

Training Relphormer for KG representation learning. The overall Relphormer optimization procedure is illustrated in Algorithm 1. During training, we jointly optimize masked knowledge loss and contextual contrastive constrained objects. The $\mathcal{L}_{contextual}$ can be viewed as a constraint term to the whole loss $\mathcal{L}_{all}$ as follows:

$$\mathcal{L}_{all} = \mathcal{L}_{MKM} + \lambda \mathcal{L}_{contextual} \quad (7)$$

where λ is a hyper-parameter, $\mathcal{L}_{MKM}$ and $\mathcal{L}_{contextual}$ are mask knowledge and contextual loss.

For example, this model (Relphormer) has two modules(pipelines). One of the modules gathers textual information and the other module aggregates structure information in graph. L_{all} is the total loss of the model, L_{MKM} is texture module loss and $L_{Contextual}$ is structural module loss.

Ultimately, we need to minimize the total loss. It means, if we choose the appropriate lambda (it's like a weight), we can define the all the losses. If we can make neural networks for predicting future climate change, we can define the losses because we have many datasets about climate change, renewable energy, energy source consumption.

For instance, total loss is for climate change (temperature etc..), one of the losses is for renewable energy generation and the other losses is co2 emission ratio by fossil fuel.

Total Loss = Climate change loss

L1 = Renewable Energy Generation loss

L2 = CO2 emissions by fossil fuel loss

Total Loss (Cost) = L1 + wL2

(We can set the 'w' as a hyperparameter or learnable parameter using neural net. Furthermore, we can easily guess that the L1 and L2 have inverse relationship. If L1 goes up, L2 goes down. We are not comparing the whole ratio of generation. We are comparing the losses. Ex) L1 -wL2 = Cost)

(If we use Renewable Energy Consumption dataset, L1 and L2 have symmetric relation. Ex) L1 + wL2 = Cost)