EN.520.650 HW5

Qihua Gong

1. The basic reason of model collapse happens is $\min_G \max_D V(G,D) \neq \max_D \min_G V(G,D)$. under the main reason there are many situation to cause the collapse. For example, in objective G, it creat images to fool D $\nabla\theta_g \frac{1}{m}\sum_{i=1}^{m} \log(1-D(G(z^{(i)})))$. Another example is when G is trained without update to D or mode collapses to a single point. In a conclusion, the collapse most happens in: 1. real-life data is multimodal
   2. when few modes generated
   3. samples generated by two different GANs

   so, the way to mitigate this problem is to: 1. use more characteristic and diverse datasets
   2. use situable data to generate situable modes
   3. don't change GAN models in one training.

2. (ii)

3. (v) Fast gradient sign method

4. Class activation map is a explanation method used for CNN. In the network, the stack of fully connected layers at the very end of the model has been replaced by layer GAP. It averages the activations of each feature map and concatenates these averages and output them as a vector. Then a weight sum of this vector is fed to the final softmax loss layer. In Grad-CAM, it usually let the gradients of any target concept store and flow into the final convolutional layer. Then, compute an importance score based on the gradients and produce a coarse localization map highlighting the important regions in the image for predicting that concept.

   So, Grad-CAM does not require a particular CNN architecture. Grad-CAM is a generalization of CAM, a method that does not require using a particular architecture. However, the CAM requires an architecture that applies global average pooling to the final convolutional feature maps, followed by a single fully connected layer that produces the predictions.

4. The Bayesian network is a probabilistic graphical version that represent a hard and rapid of variables and their conditional dependencies through a direct acyclic graph. It can use to encompass prediction, anomaly detection, diagnostics. For the reason bayesian is a step by step probailistic model, so it's compact, flexible and iberpretable illustration of a joint possibility distribution. We can check the causal relationship between variables. The Bayesian network also has the advantage of being easy and mathematically consistent across multiple accuracy records and surprising assets. It's as good as the reliability of its initial information.

The Deep learning network is a model to mimic human thought by mixing inputs, weights and biases. Deep learning network combine these together to correctly recognize, classify and describe devices in statistics. It aim to obtain algorithm knowledge very similar to synthetic neural network and mimic the statistical processing of the mind. It has many hidden layer and a number of enter and out placed layers. So, we can't know the process in the neural layers and the algorithm complexity usually very high. The algorithm efficiency not very good (low interpretability) but it can precise detect and learning. Deep gaining knowledge of makes use of synthetic neural networks to carry out state of heart computations on large quantities of statistics. It's a shape of gadget reading that works primarily based totally on the shape and characteristic of the human thoughts.

5. I think there are two ways.

The first is to balance the performance and network complexity of the deep learning network. The complexity here also determines the interpretability of the model, such as reducing hidden network layers. The so-called balance is actually to have a clear understand of the model to be processed and the things to learn, so that to build a suitable model to meets that needs. For example, when you classify or learning a simple things, you can reduce the complex network structure for better interpretability. For learning objects with complex features that need to be actually identified, we need to design complex network models to achieve their expected learning performance.

The second approch is to add parameters to the network layer to intuitively control its interpretability or use algorithms to monitor it. There are studies which suggest that such top-down connections in order to generate lower features of images starting from higher level representations. Such a mechanism can explain the creation of vivid imagery, dreaming as well as the disambiguating effect on the interpretation of local imagine regions by providing contextual prior information from previous frames.