Deep Learning
Assignment 4
10 points

1. [2 marks]
Briefly explain how dropout can be regarded as simulating an ensemble of neural network architectures.

The purpose of ensembles of neural network architectures is to reduce overfitting. Dropout is used to simulate this by randomly dropping out nodes during training. Dropout is a simple way to prevent overfitting. It has the effect that a layer with a different number of nodes to the prior layer or say different view of architecture. Dropout can force nodes within a layer to be more robust. Since the outputs of a layer under dropout are randomly subsampled, it has the effect of thinning the network during training.

2. [3 marks]
List three different methods for regularizing an autoencoder.

Sparse Autoencoders: each unit only activates to a certain type of inputs, not all of them.
Denoising Autoencoders: it use the corrupted input by some form of noise.
Contractive Autoencoder: regularizing by penalizing derivatives.

3. [2 marks]
What is the Energy function for these architectures?
a. Boltzmann Machine
b. Restricted Boltzmann Machine

a. $E(x) = -x^T Ux - b^T x$, where U is the "weight" matrix of model parameters and b is the vector of bias parameters. Formally, we decompose the units x into two subsets: the visible units v and the latent (or hidden) units h. The energy function becomes$E(v, h) = -v^T Rv - v^T W\, h - h^T Sh - b^T v - c^T h$.
b. $E(v, h) = -b^T v - c^T h - v^T W\, h$,

4. [3 marks]
Briefly explain in words the difference between a
Variational Autoencoder and a Generative Adversarial Network (GAN)

Variational autoencoders are capable of both compressing data and synthesizing data like a GAN. But GANs generate data in fine, granular detail, images generated by VAEs tend to be more blurred. VAEs can be used with discrete inputs, while GANs can be used with discrete latent variables. GAN is explicitly set up to optimize for generative tasks.