Student ID: yz12120

Mark Scheme							
Section	Question	Mark Breakdown	Marks Got	Marks Max	Marker Comments	Question Statement	
1 - VAE	1.1a	Module - Layers - MLP	6	6	Nice, but you are using sequential blocks in a rather strange fashion - only partially chaining together rather than completely	Implement a VAE - fill the blanks of nn.Module class + specify sensible hyp.param	
		Module - Layers - CNN (subsume mlp)	3	3			
		Module - Layers - BatchNorm / Dropout etc.	1	1			
		Module - Encode - Basics	3	3			
		Module - Encode - LogVar instead of Var	1	1			
		Module - Reparam - Correctly Sampling	2	2			
		Module - Forward	2	2			
		HypParam - Sensible (beta in next part)	2	2			
		Qualitative	4	5			
	1.1b	Loss with Reco and correct KL	1	1	Very nice. It is not entirely clear to me what you mean with point 1., and you are not rounding the MNIST digits, so the assumption w.r.t bernoulli could have warrant some extra discussion	Choose a suitable loss and describe your choice of RECO term	
		Loss with correct Reco and beta	1	1			
		Discussion of reco term - need to mention rescaling inputs/outputs appropriately	2	3			
	1.2a	Plots showing train and test lost terms	2	3	Seeing plots for multiple beta values would have been ideal	Results - Plot losses, investigate effect of beta	
	1.2b	Reconstructions of test set images and a few samples	1	1	Nice discussion but a) it is not clear how your "Thus the KL-divergence should increase..." follows directly from the statement of the KL divergence. For example, what would you see it beta = 1000?. b) I am not sure that you have understood posterior collapse correctly, as you do not explain it and it is quite difficult to infer what you think it is	Show reconstructions of test set, and samples	
		Each bullet point addressed	2	3		Discuss (inc. posterior collapse), visulize reconstructions	
		Qualitative	1	2			
	1.3a	Explain presence / absence of clusters	2	2	Good discussion, but we were hoping you would investigate T-SNE somewhat more thoroughly (w.r.t perplexity / geometry preserving properties etc.)	Perform T-SNE on learned representations. Discuss the results - effects of beta, clustering, outliers, boundaries, relability of T-SNE	
		Explain effect of KL and Beta on Disentangling clusters	3	3			
		Explain outliers / boundaries between clusters	1	1			
		Reflect on reliability of conclusions from T-SNE; second mark for additional plots	0	1			
	1.3b	Perform at least one interpolation	2	2	Good. In general, be careful about assuming that trajectories on a T-SNE plot can be straightforwardly linked to linear interpolation in the original high dimensional space	Interpolate between classes in latent space. Discuss: qualitative character of interp, relevance to T-SNE plot	Part Total 44 /50
		Discuss obervation from interpolation	1	1			
		Relate interpolation to T-SNE	1	1			
2 - GAN	2.1a	HypParam - Sensible	2	2	I think you did pretty well! Certainly, there are many sophisticated methods (WGAN loss and conditioning) which might have been easier to improve performance with than fiddling learning rates etc.	Implement a GAN starting with the DCGAN architecture.	
		Module - Generator-constructor + forward	3	3			
		Module - Discriminator-constructor + forward	3	3			
	2.1b	Train - Discriminator	6	6			
		Train - Generator	6	6			
	2.1c	note: please feel free to interpolate scores	7	10		Visualise the generated images	See Ref Samples
	2.1d	Detail 3 engineering features introduced. e.g. batchnorm, label smoothing, ReLU, #layers	3	4		Discuss your final architecture along with the experiments which led to it.	
		Qualitative	4	6			
	2.2	G and D loss plot - remove mark if no legend	1	1	Good discussion here, though linking the shapes of the loss curves more explicitly to the minmax loss would have been ideal :)	Plot the generator and discriminator loss curves - discuss whether sensible/expected	
		Comparision between plots and theoretical optimum/zero sum game	3	4			
	2.3	Description of mode collapse	1	1	Very nice discussion here :)	Provide a discussion on whether you noticed any mode collapse, what this behaviour may have been attributed to and explain what you did in order to combat it if it was a problem. If you didn't experience mode collapse still detail methods to combat it.	Part Total 43 /50
		Qualitative. note: If the students states that mode collapse was not a problem check the generated images to make sure that this is the case. If mode collapse was an issue yet nothiing was done to combat it remove 2 marks	4	4			
	Penalty Marks	Transformations etc. not provided below		-1	Good work!		
		Models do not load properly		-1			
		Model provided is poor compared to claimed results (VAE and GAN)		-1			
		GAN Exceeds maximum parameter count		-1			CW Total 87 /100

Student ID: yz12120