

# Homework 1: Autoregressive Models

Name: YOUR NAME

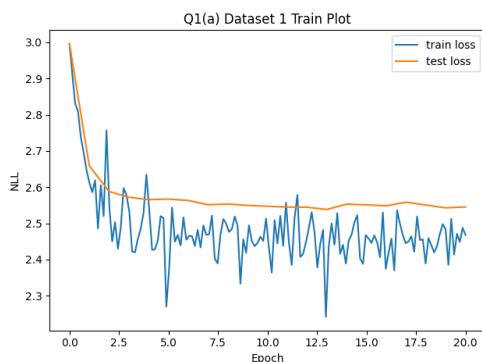
Student ID: YOUR STUDENT ID

**Deliverable:** This PDF write-up by **Tuesday February 7th, 23:59pm**. Your PDF should be generated by simply replacing the placeholder images of this LaTeX document with the appropriate solution images that will be generated automatically when solving each question. The solution images are automatically generated and saved using the accompanying IPython notebook. Your PDF is to be submitted into Gradescope. This PDF already contains a few solution images. These images will allow you to check your own solution to ensure correctness. Submit this PDF, your iPython notebook, and any other code you wrote to Gradescope!

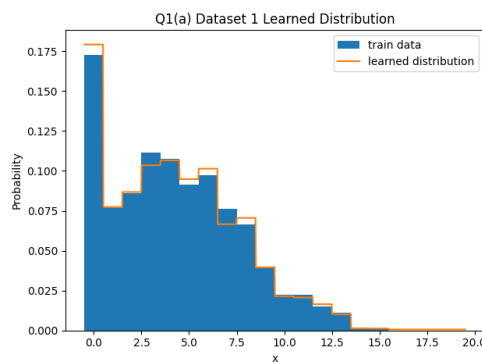
## Question 1: 1D Data

### (a) [10pt] Fitting a Histogram

Final test loss for dataset 1: 2.562 nats / dim



(a) Dataset 1: Training curve



(b) Dataset 1: Learned distribution

Final test loss for dataset 2: **FILL IN HERE** nats / dim

Placeholder

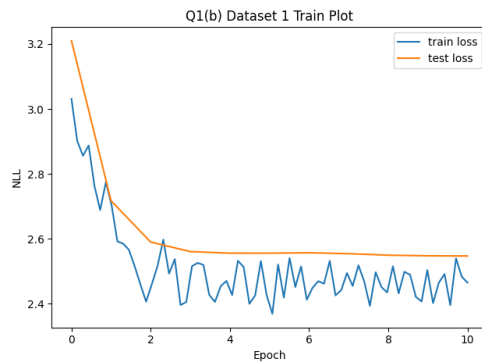
Placeholder

(a) Dataset 2: Training curve

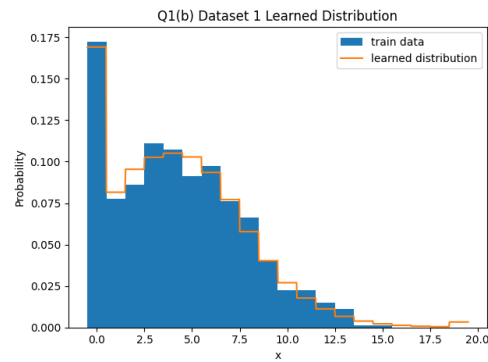
(b) Dataset 2: Learned distribution

**(b) [10pt] Fitting Discretized Mixture of Logistics**

Final test loss for dataset 1: 2.555 nats / dim



(a) Dataset 1: Training curve



(b) Dataset 1: Learned distribution

Final test loss for dataset 2: **FILL IN HERE** nats / dim

Placeholder

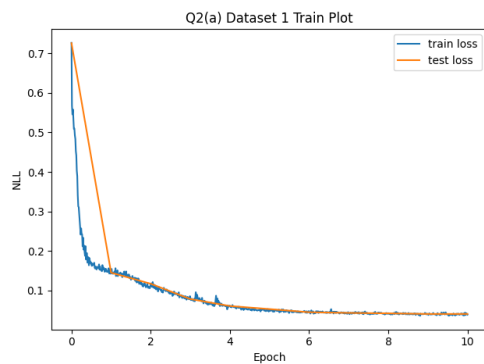
Placeholder

(a) Dataset 2: Training curve

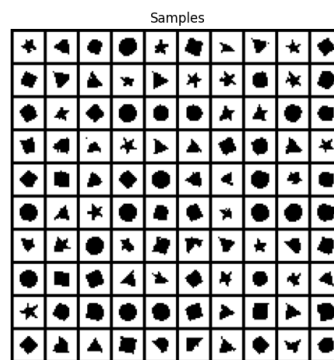
(b) Dataset 2: Learned distribution

**Question 2: PixelCNNs****(a) [15pt] PixelCNNs on Shapes and MNIST**

Final test loss for dataset 1: 0.0420 nats / dim



(a) Dataset 1: Training curve



(b) Dataset 1: Samples

Final test loss for dataset 2: **FILL IN HERE** nats / dim

Placeholder

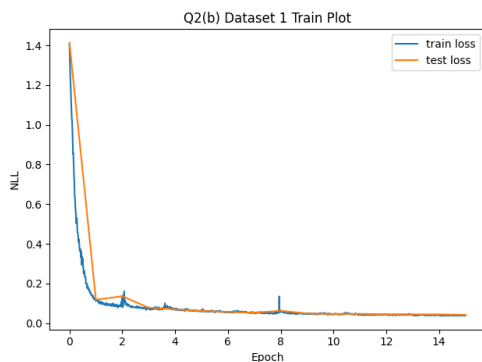
Placeholder

(a) Dataset 2: Training curve

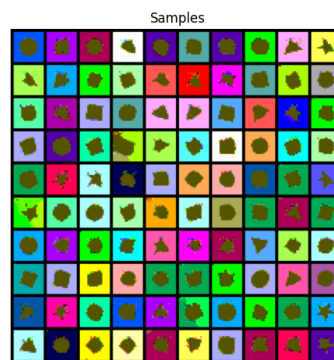
(b) Dataset 2: Samples

## (b) [15pt] PixelCNN on Colored Shapes and MNIST: Independent Color Channels

Final test loss for dataset 1: 0.0444 nats / dim



(a) Dataset 1: Training curve



(b) Dataset 1: Samples

Final test loss for dataset 2: **FILL IN HERE** nats / dim

Placeholder

Placeholder

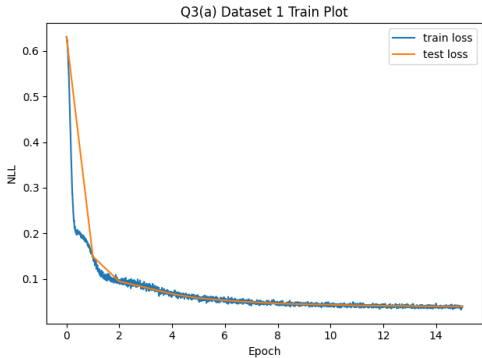
(a) Dataset 2: Training curve

(b) Dataset 2: Samples

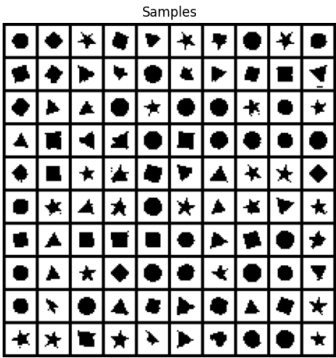
Question 3: Causal Transformer - iGPT

(a) [15pt] Autoregressive Transformer on Shapes and MNIST

Final test loss for dataset 1: 0.0397 nats / dim



(a) Dataset 1: Training curve



(b) Dataset 1: Samples

Final test loss for dataset 2: **FILL IN HERE** nats / dim

Placeholder

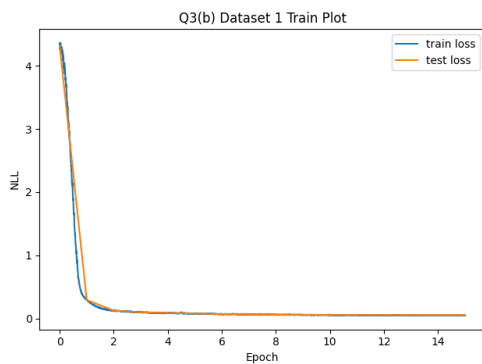
Placeholder

(a) Dataset 2: Training curve

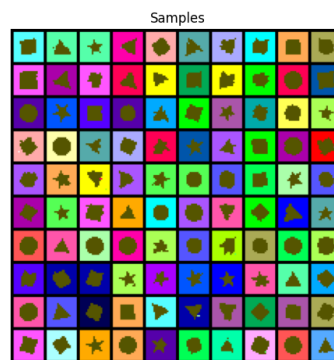
(b) Dataset 2: Samples

## (b) [15pt] Autoregressive Transformer on Colored Shapes and MNIST

Final test loss for dataset 1: 0.0541 nats / dim



(a) Dataset 1: Training curve



(b) Dataset 1: Samples

Final test loss for dataset 2: **FILL IN HERE** nats / dim

Placeholder

Placeholder

(a) Dataset 2: Training curve

(b) Dataset 2: Samples

(c) [15pt] **K,V Caching for Improved Inference**

# Placeholder

Figure 13: Dataset 2: Inference Speed

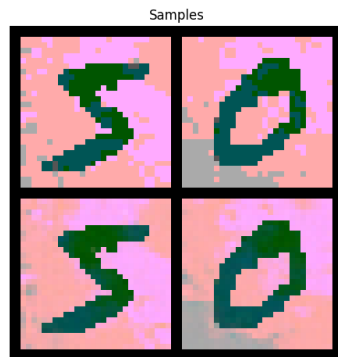
Placeholder

Placeholder

(a) Dataset 2: Samples (no caching)

(b) Dataset 2: Samples (caching)



**Question 4: Causal Transformer - Tokenized Images****(a) [5pt] Image Quantization**

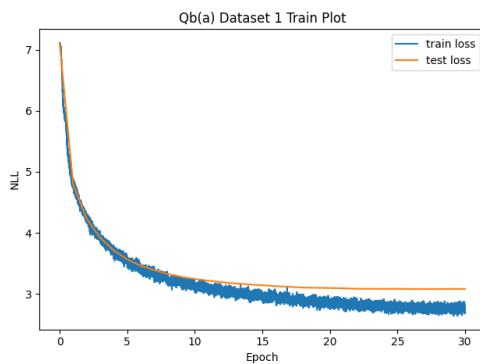
(a) Dataset 1: Quantized Examples

Placeholder

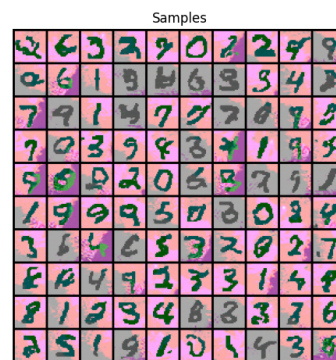
(b) Dataset 2: Quantized Examples

**(b) [15pt] Autoregressive Transformer on Colored Shapes and MNIST with Vector Quantization**

Final test loss for dataset 1: 3.083 nats / dim



(a) Dataset 1: Training curve



(b) Dataset 1: Samples

Final test loss for dataset 2: **FILL IN HERE** nats / dim

Placeholder

Placeholder

(a) Dataset 2: Training curve

(b) Dataset 2: Samples

**Question 5: Causal Transformer - Text**

**(a) [20pt] Modeling Text**

Final test loss: **FILL IN HERE** nats / dim

Placeholder

Placeholder

(a) Training curve

(b) Text samples

**Question 6: Causal Transformer - Multimodal**

**(a) [20pt] Multimodal Text and Image Generation**

Final test loss: **FILL IN HERE** nats / dim

Placeholder

Placeholder

(a) Training curve

(b) Image conditioned samples

Placeholder

Placeholder

(c) Text conditioned samples

(d) Unconditional samples