Supplementary Slides

2

Loss Function

- In the experiment results, categorical (multinomial) cross entropy was used as the loss function. Suppose there are m classes and n samples, let p_{ij} denote the probability that sample i belongs to class j (note that $\sum_j p_{ij} = 1$), and $y_{\mathrm{ij}} = \begin{cases} 1 & \text{sample } i \text{ belongs to class } j \\ 0 & \text{otherwise} \end{cases}$.
- Categorical cross entropy is calculated by:

$$\mathcal{L} = -\frac{1}{n}\sum_{i=i}^{n}\sum_{j=1}^{m}\left(y_{ij}\log\left(p_{ij}\right)\right)$$

3

Loss Function Example

- Consider the following trinary classification example (Prob[x] is the normalized predicted probability that the sample belongs to class x using softmax function):

Sample no.	Ground Truth	Prediction	Prob[A]	Prob[B]	Prob[C]
1	**A**	A	**0.95**	0.01	0.04
2	**A**	B	**0.05**	0.55	0.40

$$\mathcal{L} = -\frac{1}{2}(\log_2 0.95 + \log_2 0.05) = 2.2$$

4

Loss Function Example

- If a sample can belong to more than one class (for explanation, not used in the experiment):
  - Prob[x] shows the probability that the sample belongs to class x
  - Probability is not normalized
  - Assume threshold = 0.5

Sample no.	Ground Truth	Prediction	Prob[A]	Prob[B]	Prob[C]
1	**A**	A	**0.9**	0.2	0.4
2	**A, B**	B, C	**0.1**	**0.7**	0.55

$$\mathcal{L} = -\frac{1}{2}\left(\log_2 0.9 + (\log_2 0.1 + \log_2 0.7)\right) = 1.99$$

5

Anomaly Detection Results

MLP (200 epochs)	ReLU	ELU	LeakyReLU
Test Accuracy	0.9727	0.9462	0.9672
Test Loss	0.1185	0.3053	0.1892
Train Accuracy	0.9996	0.9960	0.9991
Train Loss	0.0010	0.0131	0.0017

CNN (50 epochs)	ReLU	ELU
Test Accuracy	0.9681	0.9599
Test Loss	0.1068	0.2025
Train Accuracy	0.9969	0.9918
Train Loss	0.0119	0.0223

6

Comparing MLP with CNN