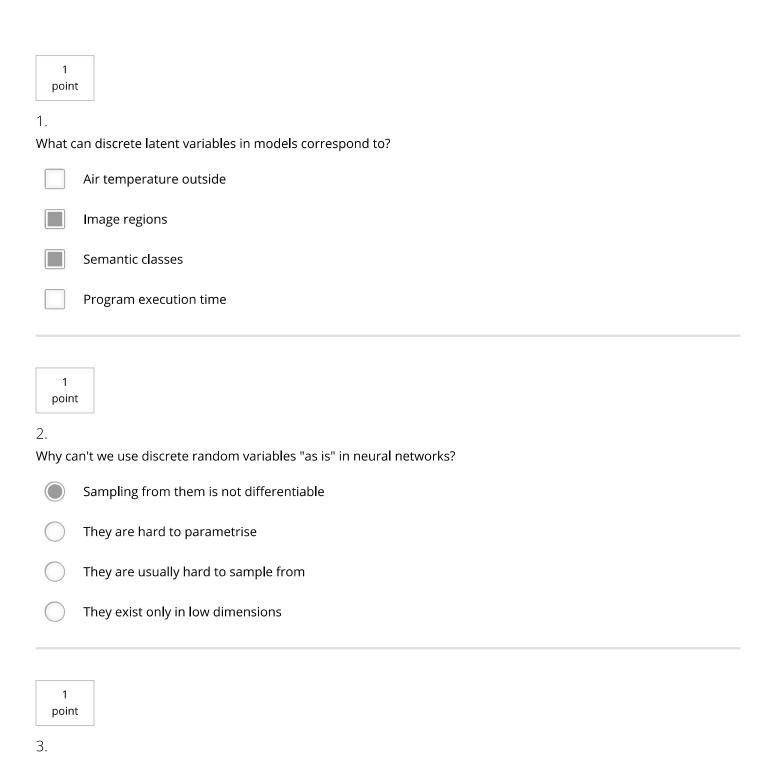
Categorical Reparametrization with Gumbel-Softmax





Select Gumbel(0, 1) distribution

 $\operatorname{Gumbel}(\mathrm{x};0,1)\propto e^{-(x+e^{-x})}$, $x\in(-\infty,\infty)$

 $\operatorname{Gumbel}(\mathrm{x};0,1)\propto e^{-x}$, $x\in(0,\infty)$

Gumbel(x; 0, 1) $\propto x(1-x)$, $x \in (0,1)$ Categorical Reparametrization with Gumbel-Softmax

Quiz, 9 que $\operatorname{\textbf{Strums}} \mathrm{bel}(\mathrm{x};0,1) \propto x e^{-x}$, $x \in (0,\infty)$

1 point

4

What is a simple way get sample g from $\operatorname{Gumbel}(0,1)$

- $u \sim U(0,1), g \sim e^u$
- $u \sim U(0,1), g \sim e^{-u}$
- $u \sim U(0,1), g \sim -log(-log(u))$
- $igg(u \sim U(0,1), g \sim -log(u) igg)$

1 point

5.

In a Gumbel-Max trick to draw samples from $\mathrm{Cat}(\pi)$, one can compute $z=\mathrm{one_hot}(arg\max_i[g_i+\log\pi_i])$, where g_i are i.i.d. samples from Gumbel(0, 1) distrubtion. How can we approximate $\mathrm{one_hot}(arg\max_i[\ldots])$ operation in a differentiable manner?

- By directly sampling from categorical distribution
- By sorting indices of $[g_i + \log \pi_i]$ and taking the index of a maximum value
- By sampling from Gumbel(0, 1) distribution
- igcup By using Softmax operation over $[g_i + \log \pi_i]$

1 point

6

For which values of $\alpha \$ is it possible to compute derivatives of Gumbel Softmax distribution $\partial y/\partial \pi$?

- au>0
- $\tau \leq 0$

Categorical Reparametrization with Gumbel-Softmax

Quiz, 9 questions τ

1 point
7.
How does parameter $ au$ influence the behaviour of Gumbel-Softmax?
As $ au$ approaches to 0 , samples are far from one-hot
lacksquare As $ au$ approaches to 0 , samples are close to one-hot
As $ au$ approaches to 0 , variance of the gradient becomes small
As $ au$ approaches to 0 , variance of the gradient becomes large
1 point 8.
What is a Straight-through Gumbel-Softmax estimator?
z is sampled from Gumbel(0, 1), gradient is computed from discrete estimation
z is sampled from discrete distribution, gradient is computed from discrete estimation
z is sampled from discrete distribution, gradient is computed from continuous estimation
z is sampled from continuous distribution, gradient is computed from discrete estimation
1 point 9.
How can one deal with unlabeled data for classification problem?
Use VAE with continuous latent variables
Use reparametrization trick
Marginalize over all classes

Categorical Reparametrization with Gumbel-Softmax

Quiz, 9 questions	
I, Jiadai Zhao , understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.	
Learn more about Coursera's Honor Code	
Submit Quiz	



