Scalable Machine Learning and Deep Learning - Review Questions 4

Deadline: December 1, 2019

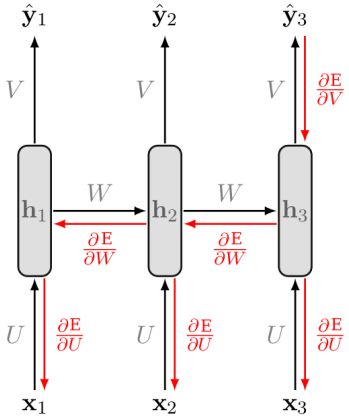
1. What's the vanishing problem in RNN?

---

2. Explain the impact of different gates in LSTM?

---

3. Assume the error of the following network is $\mathtt{E} = \mathtt{E}^{(1)} + \mathtt{E}^{(2)}$, then compute the $\frac{\partial \mathtt{E}}{\partial \mathtt{u}}$.



---

4. Assume we have a stacked autoencoder with three hidden layers $\mathtt{h}_1$, $\mathtt{h}_2$, and $\mathtt{h}_3$, in which each layer applies the following functions respectively, $\mathbf{h_1} = \mathtt{f}_1(\mathbf{x})$, $\mathbf{h_2} = \mathtt{f}_2(\mathbf{h}_1)$, and $\mathbf{h}_3 = \mathtt{f}_3(\mathbf{h}_2)$, and the output of the network will be $\mathbf{y} = \mathtt{f}_4(\mathbf{h}_3)$. Do you think if it is a good autoencoder if it generates $\mathtt{f}_4(\mathtt{f}_3(\mathtt{f}_2(\mathtt{f}_1(\mathbf{x})))) = \mathbf{x}$ for all input instances $\mathbf{x}$. How can we improve it?

---

5. How does Gibbs sampling work? When do we need to use Gibbs sampling?

---

6. How do you tie weights in a stacked autoencoder? What is the point of doing so?