Scalable Machine Learning and Deep Learning - Review Questions 2

Deadline: November 15, 2020

1. **0.5 point.** Which of the following is/are true about individual tree in Random Forest?
   (a) Individual tree is built on a subset of the features.
   (b) Individual tree is built on all the features.
   (c) Individual tree is built on a subset of instances.
   (d) Individual tree is built on full set of instances.

---

2. **0.5 point.** Ensemble model estimators (such as Random Forest) in Spark have a parameter called `featureSubsetStrategy`. What does it do?

---

3. **1 point.** Explain why the entropy becomes zero when all class partitions are pure?

---

4. **1 point.** Explain why the Gini impurity becomes zero when all class partitions are pure?

---

5. **0.5 point.** Assume a feedforward neural network with one hidden layer, in which the output of the hidden units and output units are computed by functions $\mathbf{h} = \mathtt{f}(\mathbf{x})$ and $\mathbf{out} = \mathtt{g}(\mathbf{h})$, respectively. Show that if we use linear functions in $\mathtt{f}$ and $\mathtt{g}$, e.g., $\mathbf{h} = \mathtt{f}(\mathbf{x}) = \mathbf{w}_1^\intercal \mathbf{x}$ and $\mathbf{out} = \mathtt{g}(\mathbf{h}) = \mathbf{w}_2^\intercal \mathbf{h}$, then the feedforward network as a whole would remain a linear function of its input.

---

6. **0.5 point.** What's the problem of using `step` function as an activation function in deep feedforward neural networks?

---

7. **1 point.** Compute the value of $\mathtt{w}_2$ and $\mathtt{w}_8$ after the first iteration of the backpropagation in the following figure. Assume all the neurons use the ReLU activation function and we use squared error function as the cost function. In this figure, red and orange colors indicate the initial values of the weights and biases, while the numbers in blue show the input and true output values.