✓ **Congratulations! You passed!**

Go to next item

Grade received 100%   **To pass** 80% or higher

Interpretability

Total points 3

1. Interpretation methods can be grouped based on (Check all that apply):   **1 / 1 point**

   ☑ Whether they are intrinsic or post-hoc.

   > ✓ **Correct**
   > Nice job! One way of grouping interpretability methods is by whether the model is intrinsically interpretable or treated as a black box, and external tools are used to analyze it.

   ☑ Whether they are local or global.

   > ✓ **Correct**
   > You've got it. They can also be grouped by whether they generate interpretations which are local or global.

   ☐ Whether they are complex or basic.

   ☑ Whether they are model specific, or model agnostic.

   > ✓ **Correct**
   > Keep it up! Model specific methods are limited to certain model types, while model agnostic methods are applied to any model after it is trained.

2. One key aspect that helps improve interpretability is the presence of monotonic features.   **1 / 1 point**

   ○ False

   ◉ True

   > ✓ **Correct**
   > That's right! This matches our domain knowledge for many features in many kinds of problems. When we are trying to understand a model result, if the features are monotonic, it matches our intuition about the world's reality, which we are trying to model.

3. Many classic models are intrinsically interpretable models, such as the transparent, intuitive, and relatively easy to understand neural networks.   **1 / 1 point**

   ○ True

   ◉ False

   > ✓ **Correct**
   > Absolutely right! NNs' complex architecture makes them "black boxes" when you try to interpret them.