Considerations

ICE plots

Post-hoc interpretability method	• Used AFTER training the machine learning model
Local interpretation	• Analyses the relationship between feature and target for each observation (I'd argue that they are global as well).
Model agnostic	• Especially useful for opaque estimators

Advantages

- ICE, like PDP, accurately represent how the feature influences the prediction when features are not correlated.

- ICE, unlike PDP, can help us unmask interactions or heterogeneous relationships.

Limitations

When features are correlated, the interpretation of ICE is tricky.

Animal	Area
Polar bear	Artic
Tiger	India
Koala	Australia
Kangaroo	Australia

Limitations

By assigning arbitrary values to all observations we could be creating impossible data points.

Animal	Area	Area
Polar bear	Artic	Australia
Tiger	India	Australia
Koala	Australia	Australia
Kangaroo	Australia	Australia

Limitations

- Can realistically be used to examine a handful of features.
- It does not scale.
- It does not output an “importance” value.
- If we plot too many observations it becomes harder to analyze.

Train In Data

THANK YOU

www.trainindata.com