Highlight Note

Model Interpretability in Azure Machine Learning

As machine learning becomes increasingly integral to decisions that affect health, safety, economic wellbeing, and other aspects of people's lives, it's important to be able to understand how models make predictions; and to be able to explain the rationale for machine learning based decisions while identifying and mitigating bias.

The range of machine learning algorithm types and the nature of how machine learning model training works make this a hard problem to solve, but model interpretability has become a key element of helping to make model predictions explainable.

The *Interpret-Community* Package

Model interpretability in Azure Machine Learning is based on the open source **Interpret-Community** package, which is itself a wrapper around a collection of *explainers* based on proven and emerging model interpretation algorithms, such as **Shapely Additive Explanations (SHAP)** and **Local Interpretable Model-agnostic Explanations (LIME)**

More Information: For more information about the **Interpret-Community** package, see the project GitHub repository at **https://github.com/interpretml/interpret-community/**. For details about its implementation in Azure Machine Learning, see **Model interpretability in Azure Machine Learning** in the Azure Machine Learning documentation.