

higher \leftrightarrow lower

output value

base value

0.00317 0.0 **0.01** 0.02296 0.06004 0.1479 0.3206 0.562 0.7771 0.9046 0.9626 0.9859

petal length (cm) = 5.1

petal width (cm) = 2.4

sepal width (cm) = 2.8

sepal length (cm) = 5.8

Figure: Explanation of the probability for the class Setosa for a flower from the Iris dataset. The classifier is an SVM with radial basis function and pairwise coupling².

Efficiency: $\underbrace{f(\mathbf{x})}_{\text{prediction}} - \underbrace{\mathbb{E}_{\text{Pr}}[f(\mathbf{X})]}_{\text{base value}} = \sum_{i=1}^d \phi_i(f, \mathbf{x}, \text{Pr}).$

Note that the Shapley explanation is ran in the *logit* domain!