

WHY INTERPRETABILITY?

- ML: huge potential to aid decision-making process due to its predictive performance
- ML models are black boxes, e.g., XGBoost, RBF SVM or DNNs
→ too complex to be understood by humans
- Some applications are "learn to understand"
- When deploying ML models, lack of explanations
 - ❶ hurts trust
 - ❷ creates barriers

→ Many disciplines with required trust rely on traditional models, e.g., linear models, with less predictive performance

