Interpretable ML for CPAMS

A significant trajectory of additive manufacturing (AM) technologies is cyber-physical AM systems (CPAMS) that seamlessly integrate computer-aided design models and physical AM processes.

The future growth of CPAMS is negatively impacted by shape deviations.

Machine learning (ML) algorithms can enable automated modeling of shape deviations based on point cloud data collected from CPAMS (Ferreira, Sabbaghi, and Huang, 2018).

Of critical importance in ML for CPAMS is interpreting the inputs effectively considered to be relevant, and their learned associations.

Current interpretability methods for ML have several limitations.

- Focus on single predictions and small domain regions.
- Ineffectiveness in the case of non-monotonic relationships.
- Inability to assess interactions between inputs.

Aim: Global ML Interpretability for CPAMS

Informative ML of CPAMS requires comprehensive interpretability of the learned complex relationships between inputs and shape deviations.

We developed a new predictive comparison methodology for global interpretability in ML that can identify relevant inputs and their relationships.

Notations and Definitions

Standard Average Predictive Comparison $\Delta_u$

$$\Delta_u = \int_{\mathbb{R}^d} \mathbb{E}_p \left( \frac{\mathbb{E}(y|u,v,x) - \mathbb{E}(y|u^0,v,x)}{\mathbb{E}(y|u,v,x)} \right) p(q) dq$$

Average Magnitude Predictive Comparison $\Delta_{\text{mag}}(u)$

$$\Delta_{\text{mag}}(u) = \int \frac{\mathbb{E}(y|u^0,v,x) - \mathbb{E}(y|u,v,x)}{(u - u^0)^2} p(q) dq$$

Average Root Integral of Squares Predictive Comparison $\Lambda_u$

$$\Lambda_u = \sqrt{\int \left( \mathbb{E}(y|u^0,v,x) - \mathbb{E}(y|u,v,x) \right)^2 p(q) dq}$$

Conditional and Interaction Comparisons

Average Conditional Predictive Comparison $\Delta_{u|z}$

$$\Delta_{u|z} = \int_{\mathbb{R}^d} \mathbb{E}_p \left( \frac{\mathbb{E}(y|u, z, v, x) - \mathbb{E}(y|u^0, z, v, x)}{\mathbb{E}(y|u, z, v, x)} \right) p(q) dq$$

Average Two-Factor Interaction Predictive Comparison $\Delta_{u \times z}$

$$\Delta_{u \times z} = \int_{\mathbb{R}^d} \mathbb{E}_p \left( \frac{\mathbb{E}(y|u, z, v, x) - \mathbb{E}(y|u^0, z, v, x)}{\mathbb{E}(y|u, z, v, x)} \right) p(q) dq$$

Application: Interpreting BELM for SLA

ML Algorithm: Bayesian Extreme Learning Machines (BELM)

Interpretations of BELM Inputs for Process A

Interpretations of BELM Inputs for Process B

Broader Impacts

Our new predictive comparison methodology helps to address the objective of obtaining insightful interpretations for the inputs of complex, black box ML algorithms for CPAMS.

A broader impact of our methodology is more informative and global understanding of the effects of inputs in CPAMS, with the potential of immediate practical application for a large community of AM users.

Our methodology possesses a broad scope of application to any ML algorithm that yields predictions of outcomes and enables the construction of a distribution capturing the inferential uncertainty for $\beta$.

Acknowledgments

The work of Raquel De Souza Borges Ferreira and Arman Sabbaghi was supported by the U.S. National Science Foundation under Grant No. CMMI-1544841 as part of the NSF/DHS/DOT/NASA/NIH Cyber-Physical Systems Program.

References
